



# COMPOSITIO MATHEMATICA

## On Gaussian primes in sparse sets

Jori Merikoski

Compositio Math. **161** (2025), 181–243.

doi: [10.1112/S0010437X24007632](https://doi.org/10.1112/S0010437X24007632)



FOUNDATION  
COMPOSITIO  
MATHEMATICA



LONDON  
MATHEMATICAL  
SOCIETY  
EST. 1865





# On Gaussian primes in sparse sets

Jori Merikoski

## ABSTRACT

We show that there exists some  $\delta > 0$  such that, for any set of integers  $B$  with  $|B \cap [1, Y]| \gg Y^{1-\delta}$  for all  $Y \gg 1$ , there are infinitely many primes of the form  $a^2 + b^2$  with  $b \in B$ . We prove a quasi-explicit formula for the number of primes of the form  $a^2 + b^2 \leq X$  with  $b \in B$  for any  $|B| = X^{1/2-\delta}$  with  $\delta < 1/10$  and  $B \subseteq [\eta X^{1/2}, (1-\eta)X^{1/2}] \cap \mathbb{Z}$ , in terms of zeros of Hecke  $L$ -functions on  $\mathbb{Q}(i)$ . We obtain the expected asymptotic formula for the number of such primes provided that the set  $B$  does not have a large subset which consists of multiples of a fixed large integer. In particular, we get an asymptotic formula if  $B$  is a sparse subset of primes. For an arbitrary  $B$  we obtain a lower bound for the number of primes with a weaker range for  $\delta$ , by bounding the contribution from potential exceptional characters.

## Contents

<b>1</b>	<b>Introduction</b>	<b>182</b>
1.1	Overview of the proof	185
1.2	Notation and conventions	187
<b>2</b>	<b>Lemmas</b>	<b>188</b>
2.1	Introducing finer-than-dyadic smooth weights	188
2.2	Elementary estimates	189
2.3	Sieve bounds	189
2.4	Poisson summation	190
2.5	Fourier expansions	191
2.6	Large sieve bounds	192
2.7	Zeros of Hecke $L$ -functions	192
2.8	Character sums	193
<b>3</b>	<b>Set-up and statement of the quasi-explicit formula</b>	<b>196</b>
3.1	Reduction to bounded $\lambda_b$	198
3.2	Reduction from $\lambda_b$ to $\mathbf{1}_B$	198
3.3	Reduction to $B$ in a short interval	198
<b>4</b>	<b>Type I information</b>	<b>199</b>

Received 19 January 2024, accepted in final form 24 September 2024.

*2020 Mathematics Subject Classification* 11N32 (primary), 11N36 (secondary).

*Keywords:* Gaussian prime numbers; sieve methods; Siegel zeros; Hecke  $L$ -functions; large sieve; character sums.

© The Author(s), 2025. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original article is properly cited. *Compositio Mathematica* is

© Foundation Compositio Mathematica.

<b>5</b>	<b>Proof of Theorem 3.1</b>	<b>203</b>
5.1	Proof of Proposition 5.1	208
<b>6</b>	<b>Type II information: preliminaries</b>	<b>209</b>
6.1	Sketch of the argument	210
6.2	An approximation for $\beta_n$	212
6.3	Partitioning the Type II sum	217
6.4	Proof of Lemma 6.1	218
<b>7</b>	<b>Type II information: proof of Proposition 6.7</b>	<b>222</b>
7.1	Cauchy–Schwarz	222
7.2	Proof of Lemma 7.1	223
7.3	Proof of Lemma 7.2	232
7.4	Proof of Lemma 7.3	232
7.5	Proof of Lemma 7.4	235
<b>8</b>	<b>Type II information: proof of Proposition 6.8</b>	<b>236</b>
<b>9</b>	<b>Proof of Theorem 3.3</b>	<b>237</b>
<b>10</b>	<b>Proof of Theorem 3.2</b>	<b>238</b>
10.1	No zeros $\beta > 1 - \varepsilon_1 / \log X$	238
10.2	Zero $\beta > 1 - \varepsilon_1 / \log X$ and $\Omega(B_1) \leq \Omega(B)/2$	240
<b>11</b>	<b>Proof of Theorem 1.6</b>	<b>240</b>
<b>12</b>	<b>Proof of Theorem 1.10</b>	<b>241</b>

## 1. Introduction

The distribution of prime numbers in sparse sets is a central topic in modern analytic number theory. A key motivating question is Landau’s fourth problem, which asks if there are infinitely many prime numbers of the form  $n^2 + 1$ . This is far beyond the current methods as the set is very sparse: the number of integers up to  $X$  of this form is of order  $X^{1/2}$ .

As an approximation to Landau’s question much attention has been given to Gaussian prime numbers  $p = a^2 + b^2$  with  $b$  restricted to some specific sparse set  $B$ . A major breakthrough was achieved by Friedlander and Iwaniec [FI98b] who proved that there are infinitely many primes of the form  $a^2 + b^4$ , that is, with  $B$  being the set of squares. Following this, there have been many variants where  $b$  is drawn from a sparse set  $B$ , for instance, the papers of Heath-Brown and Li, Pratt, and the current author [HBL17, Mer22, Pra20]. We also point out the results of Heath-Brown [HB01], Li [Li21], and Maynard [May20] for primes in other polynomial sequences, where Li’s result has the record for the sparsest polynomial sequence with primes, with size of the set being  $X^{43/67+\varepsilon}$ .

Notably, all of the above-mentioned results exploit heavily the structure of the specific sparse set  $B$ , leaving open the question of what can be said about an arbitrary sparse set  $B$ . In this direction Fouvry and Iwaniec [FI97] proved that one can take  $B$  with density  $(\log X)^{-C}$  for any  $C > 0$  and establish an asymptotic formula for the number of primes  $a^2 + b^2 \leq X$  with  $b \in B$ . Using the argument of Fouvry and Iwaniec one would be required to improve upon the famous Siegel–Walfisz theorem to reach sparser sets  $B$ . In comparison, our main result obtains unconditionally a power saving in the density of  $B$  for the first time.

**THEOREM 1.1.** *There is some (computable)  $\delta > 0$  such that the following holds. If  $B$  is a set of integers with  $|B \cap [0, Y]| \gg Y^{1-\delta}$  for all  $Y \gg 1$ , then there are infinitely many primes of the form  $a^2 + b^2$  with  $b \in B$ .*

This result is a corollary of the following lower bound for the number of such primes, which is weaker than expected by a factor of  $\gg_\varepsilon X^{-\varepsilon}$ .

**THEOREM 1.2.** *There is some (computable)  $\delta > 0$  such that the following holds for any small  $\eta > 0$ . For all sufficiently large  $X$  and for all  $B \subseteq [\eta X^{1/2}, (1 - \eta)X^{1/2}] \cap \mathbb{Z}$  with  $|B| \geq X^{1/2-\delta}$  we have, for any  $\varepsilon > 0$ ,*

$$\sum_{p=a^2+b^2 \leq X} \mathbf{1}_B(b) \gg_\varepsilon X^{1/2-\varepsilon} |B|.$$

*Deduction of Theorem 1.1 from Theorem 1.2.* Let  $\delta > 0$  be small and let  $B$  be a set of integers with  $|B \cap [0, Y]| \gg Y^{1-\delta}$  for all  $Y \gg 1$ . Then by the pigeonhole principle for any  $\varepsilon > 0$  for any sufficiently large  $Y$ , there is some  $Y_1 \in [Y^{1-\delta-\varepsilon}, Y]$  such that  $B_1 := B \cap [Y_1/4, Y_1/2]$  satisfies  $|B_1| \geq Y_1^{1-\delta-\varepsilon}$ . By Theorem 1.2 with  $X = Y_1^2$  and a trivial upper bound for  $a^2 + b^2 \leq X^{1-4\varepsilon}$  we have

$$\begin{aligned} \sum_{X^{1-4\varepsilon} < p=a^2+b^2 \leq X} \mathbf{1}_{B_1}(b) &= \sum_{p=a^2+b^2 \leq X} \mathbf{1}_{B_1}(b) - \sum_{a^2+b^2 \leq X^{1-4\varepsilon}} \mathbf{1}_{B_1}(b) \\ &\gg_\varepsilon X^{1/2-\varepsilon} |B_1|. \end{aligned}$$

Therefore, for all large  $Y$  there exists a prime number  $p = a^2 + b^2 \in (Y^{2-2\delta-8\varepsilon}, Y^2]$  with  $b \in B$ , so that, in particular, there are infinitely many primes of the form  $a^2 + b^2$  with  $b \in B$ .  $\square$

*Remark 1.3.* A back-of-the-envelope estimate shows that it should be possible to establish Theorem 1.2 for some  $\delta \in (1/20, 1/10)$  but we have not checked this as it depends on the numerical constants  $c_1, c_2, c_3$  in Lemmas 2.13 and 2.14 as well as optimization of Theorem 3.1: this would require a separate lengthy optimization similar to the arguments in [HB92]. It is possible to generalize our results to general binary quadratic forms instead of  $a^2 + b^2$  by adapting ideas from [LSX20].

*Remark 1.4.* The lower bound in Theorem 1.2 is the best that can be hoped for in general with the current technology. In fact, improving the lower bound in Theorem 1.2 to the correct order of magnitude would imply the non-existence of Siegel zeros by a suitable application of Theorem 3.3. The implied constant in the lower bound  $\gg_\varepsilon$  in Theorem 1.2 is ineffective but the result can be made effective with  $\varepsilon = \delta$  for some very small  $\delta > 0$  by using the class number formula, so that Theorem 1.1 is effective in all aspects.

*Remark 1.5.* The argument we give works for  $\eta = X^{-\varepsilon}$  for some small  $\varepsilon > 0$ . It is possible to extend the proof of Theorem 1.2 to handle sets  $B \subseteq [0, X^{1/2}]$ . The possibility that  $B$  or the variable  $a$  is restricted to a narrow interval  $[0, X^{1/2-\delta'}]$  for some  $\delta' \leq \delta$  adds only a technical problem, namely, all of the Gaussian primes  $b + ia$  counted lie in a very narrow sector. See Remark 7.7 for further details on how to modify our argument.

In the case that  $B$  is unbiased, we are able to obtain an asymptotic formula similar to [FI97] for  $\delta < 1/10$ . Let

$$\rho(d) := |\{\nu \in \mathbb{Z}/d\mathbb{Z} : \nu^2 + 1 \equiv 0 \pmod{d}\}|$$

and define

$$\omega(b) := \prod_{p|b} \left(1 - \frac{\rho(p)}{p}\right)^{-1}.$$

Then the following result is a corollary of our quasi-explicit formula (cf. Theorem 3.3).

**THEOREM 1.6.** Let  $\eta > 0$  be small. Let  $C'' > 0$  be large compared with  $C' > 0$  which is large compared with  $C > 0$ . Let  $\lambda_b$  be complex coefficients with  $|\lambda_b| \leq X^{o(1)}$ , supported on  $[\eta X^{1/2}, (1 - \eta)X^{1/2}] \cap \mathbb{Z}$ , and satisfying

$$\sum_b |\lambda_b| \geq X^{2/5+\varepsilon}.$$

Suppose that for all  $(\log X)^{C''} < q \leq X^{\delta+\eta}$  we have

$$\sum_{b \equiv 0 \pmod{q}} |\lambda_b| \leq \frac{1}{(\log X)^{C'}} \sum_b |\lambda_b|. \quad (1.1)$$

Then

$$\sum_{a^2+b^2 \leq X} \lambda_b \Lambda(a^2+b^2) = \frac{4}{\pi} \sum_{\substack{a^2+b^2 \leq X \\ (a,b)=1}} (\lambda_b \omega(b) + O_{C,C',C''}(|\lambda_b|(\log X)^{-C})).$$

*Remark 1.7.* With more work the assumptions that  $|\lambda_b| \leq X^{o(1)}$  and  $\sum_b |\lambda_b| \geq X^{2/5+\varepsilon}$  may be replaced (here and later) by

$$\|\lambda_b\|_1 \geq X^{1/5+\varepsilon} \|\lambda_b\|_2,$$

where  $\|\lambda_b\|_p := (\sum_b |\lambda_b|^p)^{1/p}$ . By adapting Harman's sieve [Har07] one can get a correct order lower bound for  $1/2 - \delta$  for some  $\delta \in (1/8, 1/10)$ .

*Remark 1.8.* We note that  $\rho(p) = 1 + \chi_4(p)$  with  $\chi_4$  being the unique non-trivial character to modulus 4 and

$$\frac{4}{\pi} = \prod_p \left(1 - \frac{\rho(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-1}. \quad (1.2)$$

Also, for all  $b$ ,

$$\sum_{\substack{a^2+b^2 \sim X \\ (a,b)=1}} \omega(b) \sim \sum_{a^2+b^2 \sim X} \frac{\varphi(b)}{b} \omega(b) = \sum_{a^2+b^2 \sim X} \prod_{p|b} \left(1 - \frac{\rho(p)}{p}\right)^{-1} \left(1 - \frac{1}{p}\right),$$

so that the main term is of the same form as in [FI97, Theorem 1].

The assumption (1.1) in Theorem 1.6 is very mild since  $C''$  can be taken to be large compared with  $C'$  and, thus, (1.1) applies to most instances that come up in nature. For example, we immediately get as a corollary [Mer22, Theorem 3] for large  $k$  and a weak version of [Pra20, Theorem 1.1] where the base is taken to be large with  $B$  being the set of numbers missing one digit.

In particular, we get an asymptotic formula if  $B$  is a sparse subset of primes.

**COROLLARY 1.9.** Let

$$B \subseteq [\eta X^{1/2}, (1 - \eta)X^{1/2}] \cap \mathbb{P}$$

with  $|B| \geq X^{2/5+\varepsilon}$ . Then, for any  $C > 0$ ,

$$\sum_{a^2+b^2 \leq X} \mathbf{1}_B(b) \Lambda(a^2+b^2) = \frac{4}{\pi} \sum_{a^2+b^2 \leq X} \mathbf{1}_B(b) + O_C\left(\frac{X^{1/2}|B|}{(\log X)^C}\right).$$

The assumption (1.1) in Theorem 1.6 may be replaced by a wide zero-free region for Hecke  $L$ -functions (cf. §2.7 for the relevant notation). For the statement, we let  $\mathcal{M}(u)$  denote the smallest integer  $m$  with  $u|m$ .

**THEOREM 1.10.** *For all  $C > 0$ , there is some  $C' > 0$  such that the following holds. Let  $\lambda_b$  be complex coefficients with  $|\lambda_b| \leq X^{o(1)}$ , supported on  $[\eta X^{1/2}, (1-\eta)X^{1/2}] \cap \mathbb{Z}$ , and satisfying*

$$\sum_b |\lambda_b| \geq X^{2/5+\varepsilon}.$$

*Assume that Hecke  $L$ -functions  $L(s, \xi_k \chi)$  on  $\mathbb{Q}(i)$  with  $|k| \leq X^\eta$  and modulus  $\mathcal{M}(u) \leq X^{1/10+\eta}$  have no zeros in the region*

$$\sigma > 1 - \frac{C' \log \log X}{\log X}, \quad |t| \leq X^\eta. \quad (1.3)$$

Then

$$\sum_{a^2+b^2 \leq X} \lambda_b \Lambda(a^2+b^2) = \frac{4}{\pi} \sum_{\substack{a^2+b^2 \leq X \\ (a,b)=1}} (\lambda_b \omega(b) + O_C(|\lambda_b|(\log X)^{-C})).$$

### 1.1 Overview of the proof

Our goal is to estimate

$$\sum_{p=a^2+b^2 \sim X} \mathbf{1}_B(b), \quad (1.4)$$

where  $|B| = X^{1/2-\delta}$ . It is immediately apparent that potential Siegel zeros can cause a problem, since if  $\chi_1 \in (\widehat{\mathbb{Z}/q_1\mathbb{Z}})^\times$  is an exceptional character (necessarily quadratic) to modulus  $q_1 \leq X^\delta$  and if  $B \subseteq q_1\mathbb{Z}$ , then for  $b \in B$

$$\mu(a^2+b^2) \approx \chi_1(a^2+b^2) = \chi_1(a^2) = 1,$$

implying that the sequence  $a^2+b^2$  is biased towards numbers with an even number of prime factors. In this case we would expect the main term for (1.4) to be multiplied essentially by  $L(1, \chi_1)$ . Note that the function

$$b+ia \mapsto \chi(a^2+b^2)$$

defines a Dirichlet character on  $(\mathbb{Z}[i]/q_1\mathbb{Z}[i])^\times$ .

Similarly as in the proof of Linnik's theorem [Lin44], our argument splits into two cases depending on whether there is a Siegel zero or not. We choose a small parameter  $\varepsilon_1 > 0$  and we will take  $\delta$  to be small in terms of  $\varepsilon_1$ . If there is a zero

$$\beta_1 \geq 1 - \frac{\varepsilon_1}{\log X},$$

then we can use a similar argument as in [FI10, Chapter 24.2] to give a lower bound for primes of the form  $a^2+b^2$ , using just Type I information. This works for a certain fixed  $\varepsilon_1$  once  $\delta > 0$  is sufficiently small.

For simplicity, let us then assume in this sketch that we are in the situation of Theorem 1.10, that is, for Dirichlet characters  $\chi$  to moduli  $u \in \mathbb{Z}[i]$  with  $\mathcal{M}(u) \leq X^{\delta+\eta}$  Hecke  $L$ -functions on  $\mathbb{Q}(i)$  have no zeros in the wider region (1.3). By Vaughan's identity, evaluating (1.4) is reduced to estimating

$$\text{Type I sums: } \sum_{d \leq D} \alpha(d) \sum_{\substack{a^2+b^2 \sim X \\ a^2+b^2 \equiv 0 \pmod{d}}} \mathbf{1}_B(b) \quad (1.5)$$

and for  $MN = X$

$$\text{Type II sums: } S(\alpha, \beta) := \sum_{\substack{m \sim M \\ n \sim N}} \alpha(m) \beta(n) \sum_{mn=a^2+b^2} \mathbf{1}_B(b), \quad (1.6)$$

where  $\alpha, \beta$  denote bounded coefficients.

For the Type I sums (1.5) the argument goes back to [FI97] and we get an asymptotic formula for  $D = X^{1-\delta-o(1)}$  by applying Poisson summation for the free variable  $a$  and the quadratic large sieve (cf. Lemma 2.12). For small  $\delta$  the exponent of distribution approaches  $1 - o(1)$ , so that we only need very little parity breaking Type II information.

For the Type II sums (1.6) it suffices to consider the case when  $\beta = \mu$ , the Möbius function. Similar to [FI98b], by unique factorization in  $\mathbb{Q}(i)$  we essentially have, for  $w, z \in \mathbb{Z}(i)$ ,

$$a^2 + b^2 = mn = |\bar{w}z|^2 = |b + ia|^2$$

and get

$$S(\alpha, \mu) = \sum_{\substack{|w|^2 \sim M \\ |z|^2 \sim N}} \alpha(|w|^2) \mu(|z|^2) \mathbf{1}_B(\operatorname{Re}(\bar{w}z)).$$

After using Cauchy–Schwarz similarly to [FI98b] we essentially need to show cancellation in

$$\begin{aligned} & \sum_{|z_1|^2, |z_2|^2 \sim N} \mu(|z_1|^2) \mu(|z_2|^2) \sum_{\substack{b_1, b_2 \in B \\ b_1 z_2 \equiv b_2 z_1 \pmod{\operatorname{Im}(z_2 \bar{z}_1)}}} 1 \\ & =: \sum_{|z_1|^2, |z_2|^2 \sim N} \mu(|z_1|^2) \mu(|z_2|^2) T_B(z_1, z_2). \end{aligned}$$

The goal is to evaluate the sum  $T_B(z_1, z_2)$  with a main term  $M_B(z_1, z_2)$  and then bound

$$\sum_{|z_1|^2, |z_2|^2 \sim N} \mu(|z_1|^2) \mu(|z_2|^2) M_B(z_1, z_2) \quad (1.7)$$

by catching the oscillations from  $\mu(|z|^2)$ . All of the previous works [FI98b, HBL17, Mer22, Pra20] rely on specific analytic and arithmetic properties of the set  $B$  to evaluate the sum  $T_B(z_1, z_2)$ .

For a general sparse set  $B$  we first factor out  $b_0 = (b_1, b_2)$  to get, assuming for simplicity that  $b_0 | \operatorname{Im}(z_2 \bar{z}_1)$ ,

$$T_B(z_1, z_2) = \sum_{b_0} \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ b'_1 z_2 \equiv b'_2 z_1 \pmod{\operatorname{Im}(z_2 \bar{z}_1)/b_0}}} 1.$$

It is crucial to carefully track the dependency on  $b_0$  since it is possible that a large subset of  $B \times B$  has a large common factor  $b_0$ , for instance, if  $B$  is in  $q\mathbb{Z}$  for some fixed  $q \leq X^\delta$ .

To evaluate  $T_B(z_1, z_2)$  use Dirichlet characters to expand the congruence  $b'_1 z_2 \equiv b'_2 z_1 \pmod{\operatorname{Im}(z_2 \bar{z}_1)/b_0}$ . The contribution from Dirichlet characters with a large conductor  $d \geq X^{\delta+\eta}/b_0$  may be bounded by using the large sieve for multiplicative characters (Lemma 2.11). The Dirichlet characters with a small conductor  $d < X^{\delta+\eta}/b_0$  give the main term  $M_B(z_1, z_2)$  and we can bound (1.7) provided that for any  $f = db_0 \leq X^{\delta+\eta}$  and  $a \in (\mathbb{Z}/f\mathbb{Z})^\times$ , we have

$$\sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ z_2 \equiv a z_1 \pmod{f}}} \mu(|z_1|^2) \overline{\mu(|z_2|^2)} \ll \frac{N^2}{f^2 (\log X)^C}.$$

This follows by standard arguments (e.g. using Heath-Brown's identity for  $\mu$ ) from the assumption (1.3), provided that  $N$  is sufficiently large compared with  $f$ , say,  $N > X^\eta f^3$ . We give a more detailed sketch of the argument for the Type II sums in § 6.1.

Without assuming the zero-free region (1.3), for the Type II sums we need to extract the contribution from the potential exceptional characters before applying Cauchy-Schwarz to  $S(\alpha, \mu)$ . Denoting  $\mu_z := \mu(|z|^2)$ , we write

$$\mu_z = \mu_z^\# + \mu_z^b,$$

where  $\mu_z^\#$  is an approximation for  $\mu_z$  and  $\mu_z^b$  is a function which is balanced along arithmetic progressions (cf. (1.8)). Importantly,  $\mu_z^\#$  must be simple enough so that  $S(\alpha, \mu^\#)$  can be evaluated using only Type I information.

Let  $\chi_j$  for  $j \leq J = (\log X)^{O(1)}$  denote the worst of the possible exceptional characters  $\chi_j$  of conductors  $u_j \in \mathbb{Z}[i]$  with  $\mathcal{M}(u_j) \leq X^{\delta+\eta}$  and denote the normalized correlation of two coefficients  $\alpha, \beta$  by

$$\mathcal{C}(\alpha, \beta) := \left( \sum_{|z|^2 \sim N} \alpha_z \overline{\beta_z} \right) \left( \sum_{|z|^2 \sim N} |\beta_z| \right)^{-1}.$$

Then, for our approximation, we essentially choose

$$\mu_z^\# := \sum_{j \leq J} \overline{\chi_j}(z) \mathcal{C}(\mu, \overline{\chi_j}).$$

The idea is similar in spirit to the dispersion method of Drappeau [Dra15] which also takes into account contributions from multiple characters. This construction is also motivated by the prime number theorem of Gallagher [Gal70, Theorem 7] and its use by Montgomery and Vaughan to get a power saving for the exceptional set in the binary Goldbach problem [MV75]. For the function  $\mu^b := \mu_z - \mu_z^\#$ , we can then unconditionally show that for any  $f \leq X^{\delta+\eta}$ ,  $N > X^\eta f^3$ , and  $a \in (\mathbb{Z}/f\mathbb{Z})^\times$

$$\sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ z_2 \equiv az_1 (f)}} \mu_{z_1}^b \overline{\mu_{z_2}^b} \ll \frac{N^2}{f^2 (\log X)^C}. \quad (1.8)$$

Taking into account the bias  $S(\alpha, \mu^\#)$  from the exceptional characters for the Type II sums, we obtain the quasi-explicit formula Theorem 3.3.

The paper is structured as follows. In § 2 we present some basic lemmas. In § 3 we split the proof of Theorem 1.2 into two cases depending on the existence of a Siegel zero (Theorems 3.1 and 3.2) and state the quasi-explicit formula Theorem 3.3. In § 4 we evaluate Type I sums and in § 5 we give a proof of Theorem 1.2 under the assumption that a Siegel zero does exist. In §§ 6, 7, and 8 we estimate Type II sums. In §§ 9, 11, and 12 we deduce our main theorems.

## 1.2 Notation and conventions

Our notation and conventions are as follows.

- $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}, \mathfrak{f}, \mathfrak{m}, \mathfrak{n}, \mathfrak{p}$ : Ideals of  $\mathbb{Z}[i]$ , reserving  $\mathfrak{p}$  for prime ideals.
- $u, v, w, z$ : Gaussian integers.
- $(z, w)$ : Greatest common primary divisor.
- $\chi, \psi, u$ : Dirichlet characters  $\chi, \psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times$ ,  $u \in \mathbb{Z}[i]$ .
- $\xi_k(z)$ : Character  $\xi_k(z) = (z/|z|)^k = e^{ik \arg z}$ .



- $\mathcal{N}\mathbf{n}$ : Norm of an ideal,  $\mathcal{N}(z) = |z|^2$ .
- $\mathcal{M}(z)$ : Smallest integer  $m$  with  $z|m$ ,  $\mathcal{N}(z)^{1/2} \leq \mathcal{M}(z) \leq \mathcal{N}(z)$ .
- $k, \ell, m, n, p, q, r, s$ : Integers, reserving  $p$  for a prime number.
- $\nu_j$ : Small power of  $X$ ,  $\nu_j = X^{-\eta_j}$ .
- $\eta$ : A generic small constant.
- $C$ : A generic large constant.
- $F$ : A smooth compactly supported function.
- $H_N(\mathbf{n})$ : Indicator of  $\mathcal{N}\mathbf{n} \in [N, N(1 + \nu_2)]$  for  $\nu_2 = X^{-\eta_2}$ .
- $W$ : Denotes  $X^{1/(\log \log X)^2}$ .
- $P(W)$ : Denotes  $\prod_{p < W} p$ .

## 2. Lemmas

### 2.1 Introducing finer-than-dyadic smooth weights

Here we describe a device that allows us to partition a sum smoothly into finer-than-dyadic intervals. Let  $\nu \in (0, 1/10)$  be small (we will use  $\nu = X^{-\varepsilon}$  or  $\nu = \log^{-C} X$ ), and fix a non-negative  $C^\infty$ -smooth function  $F$  supported on  $[1 - \nu, 1 + \nu]$  and satisfying

$$|F^{(j)}| \ll_j \nu^{-j}, \quad j \geq 0 \quad \text{and} \quad \int_{1/2}^2 F(1/t) \frac{dt}{t} = \nu.$$

Suppose that we want to introduce a smooth partition to bound a sum of the form  $\sum_{n \leq N} f_n$ . We can write (using a change of variables  $t \mapsto t/n$ )

$$\begin{aligned} \sum_{n \leq N} f_n &= \frac{1}{\nu} \sum_{n \leq N} f_n \int_{1/2}^2 F(1/t) \frac{dt}{t} = \frac{1}{\nu} \sum_{n \leq N} \int_{1/2}^{2N} f_n F(n/t) \frac{dt}{t} \\ &= \frac{1}{\nu} \int_{1/2}^{2N} \sum_{n \leq N} f_n F(n/t) \frac{dt}{t}, \end{aligned}$$

so that

$$\left| \sum_{n \leq N} f_n \right| \leq \frac{1}{\nu} \int_{1/2}^{2N} \left| \sum_{n \leq N} f_n F(n/t) \right| \frac{dt}{t}.$$

Hence, at the cost of a factor  $\nu^{-1} \log N$ , it suffices to consider sums of the form

$$\sum_{n \leq N} f_n F(n/t)$$

for  $t \leq 2N$ . Naturally, if the original sum is  $\sum_{n \sim N} f_n$  (dyadic  $n$ ), then it suffices to consider  $\sum_{n \sim N} f_n F(n/t)$  for  $t \asymp N$  at a cost  $\nu^{-1}$ . The effect is the same as with the usual smooth partition of unity except that we did not need to construct explicitly the partition functions  $F$ . We use the notation  $F_t(n) := F(n/t)$ .

We also have the following variant on  $\mathbb{R}/2\pi\mathbb{Z}$ . Fix  $\nu_1 = X^{-\eta_1}$  for some small  $\eta_1 > 0$  and let  $G: \mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{C}$  be a non-negative  $C^\infty$ -smooth function supported on  $[-\nu_1, \nu_1]$ , satisfying

$$|G^{(j)}| \ll_j \nu_1^{-j}, \quad j \geq 0 \quad \text{and} \quad \int G(\theta) d\theta = \nu_1.$$

Then, for a sum over Gaussian integers, we may write

$$\sum_{z \in \mathbb{Z}[i]} f_z = \frac{1}{\nu_1} \int_{\mathbb{R}/2\pi\mathbb{Z}} \sum_{z \in \mathbb{Z}[i]} f_z G(\arg z - \theta) d\theta$$

to obtain a smooth finer-than-dyadic partition in terms of  $\arg z$ .

## 2.2 Elementary estimates

LEMMA 2.1. *We have*

$$\sum_{1 \leq a \leq A} (a, d) \leq \tau(d)A.$$

*Proof.* Denoting  $c = (a, d)$  and  $a = ca'$ , we have

$$\sum_{1 \leq a \leq A} (a, d) \leq \sum_{c|d} c \sum_{1 \leq a' \leq A/c} 1 \leq \tau(d)A. \quad \square$$

In handling the weights  $\mathbf{1}_{(n, P(W))=1}$  we can use the following standard bound for exceptionally smooth numbers, which effectively gives a version of the fundamental lemma of the sieve (see, for instance, [Ten15, Chapter III.5, Theorem 1]).

LEMMA 2.2. *For any  $2 \leq Z \leq Y$ , we have*

$$\sum_{\substack{n \sim Y \\ P^+(n) < Z}} 1 \ll Ye^{-u/2},$$

where  $u := \log Y / \log Z$ .

We also require the following elementary bound (see, for instance, [FI98a, Lemma 1]).

LEMMA 2.3. *For every square-free integer  $n$  and every  $k \geq 2$ , there exists some  $d|n$  such that  $d \leq n^{1/k}$  and*

$$\tau(n) \leq 2^k \tau(d)^k.$$

From this we get the more general version.

LEMMA 2.4. *For every integer  $n$  and every  $k \geq 2$ , there exists some  $d|n$  such that  $d \leq n^{1/k}$  and*

$$\tau(n) \leq 2^{k^2} \tau(d)^{k^3}.$$

*Proof.* Write  $n = b_1 b_2^2 \cdots b_{k-1}^{k-1} b_k^k$  with  $b_1, \dots, b_{k-1}$  square-free, by letting  $b_k$  be the largest integer such that  $b_k^k | n$ , so that  $n/b_k^k$  is  $k$ -free and splits uniquely into  $b_1 b_2^2 \cdots b_{k-1}^{k-1}$  with  $b_j$  square-free. We have

$$\tau(n) \leq \tau(b_1) \tau(b_2)^2 \cdots \tau(b_k)^k.$$

By Lemma 2.3 for all  $j \leq k-1$  there is some  $d_j | b_j$  with  $d_j \leq b_j^{1/k}$  and  $\tau(b_j) \leq 2^k \tau(d_j)^k$ . Hence, for  $d = d_1 \cdots d_{k-1} b_k$ , we have

$$d \leq (b_1 \cdots b_{k-1})^{1/k} b_k \leq (b_1 \cdots b_{k-1} b_k^k)^{1/k} \leq n^{1/k}$$

and

$$\tau(n) \leq (2\tau(d_1) \cdots \tau(d_{k-1}) \tau(b_k))^{k^2} \leq 2^{k^2} \tau(d)^{k^3}. \quad \square$$

## 2.3 Sieve bounds

The following version of the fundamental lemma of the sieve follows from applying [FI10, Theorem 6.9] to the real and the imaginary parts of  $a_n$ .

LEMMA 2.5 (Fundamental lemma of the sieve). *Let  $\mathcal{A} = (a_n)$  be a sequence of complex coefficients and let  $\kappa \geq 0$ ,  $Z \geq 2$ , and  $D \geq Z^{9\kappa+2}$ . Suppose that for some  $\mathbb{X}$  and for some real-valued multiplicative function  $g(d)$ , we have, for all square-free  $d$ ,*

$$\sum_{n \equiv 0 \pmod{d}} a_n = g(d)\mathbb{X} + r_d$$

and suppose that, for all  $p$ , we have  $0 \leq g(p) < 1$ . Suppose that for some  $K > 1$  we have for all  $W < Z$

$$\prod_{W \leq p < Z} (1 - g(p))^{-1} \leq K \left( \frac{\log Z}{\log W} \right)^\kappa.$$

Use the notation

$$V(Z) := \prod_p (1 - g(p)).$$

Then, for some bounded coefficients  $\lambda_d$  depending only on  $\kappa$ , we have

$$\sum_{(n, P(Z))=1} a_n = \mathbb{X}V(Z)(1 + O_{\kappa, K}(e^{-s})) + O \left( \left| \sum_{\substack{d < D \\ d|P(Z)}} \lambda_d r_d \right| \right).$$

*Remark 2.6.* The fact that  $\lambda_d$  depend only on  $\kappa$  and not the sequence  $a_n$  will be important for us since we apply the same sieve to several sequences  $a_n^{(b)}$  indexed by  $b \in B$  and then use the summation over  $b \in B$  while bounding the remainder  $\sum_{b \in B} \left| \sum_{d < D} \lambda_d r_d^{(b)} \right|$  (cf. Proposition 4.1). In our set-up, the functions  $g^{(b)}(d)$  and  $V^{(b)}(Z)$  also depend on  $b$  with  $g^{(b)}(d) = \mathbf{1}_{(d, 2b)=1} \rho(d)/d$ . Thus, we have

$$\sum_{b \in B} V^{(b)}(Z) = V(z) \sum_{b \in B} \prod_{\substack{p|b \\ 2 < p < Z}} (1 - \rho(p)/p)^{-1},$$

where the product over  $p|b, 2 < p < Z$  may be completed to all prime factors  $p|b, p \neq 2$  with a negligible error term if  $Z$  is not too small.

## 2.4 Poisson summation

The following lemma gives a truncated version of the Poisson summation formula.

**LEMMA 2.7** (Poisson summation). *Let  $F$  be as in § 2.1 for some  $\nu \in (0, 1/10)$  and denote  $F_N(n) := F(n/N)$ . Let  $x \gg 1$  and let  $q \sim Q$  be an integer. Let  $\varepsilon > 0$  and denote*

$$H := \nu^{-1}(QN)^\varepsilon Q/N.$$

*Then, for any  $A > 0$ ,*

$$\sum_{n \equiv a \pmod{q}} F_N(n) = \frac{N}{q} \hat{F}(0) + \frac{N}{q} \sum_{1 \leq |h| \leq H} \hat{F}\left(\frac{hN}{q}\right) e_q(-ah) + O_{A, \varepsilon, F}((QN)^{-A}),$$

*where  $\hat{f}(h) := \int f(u) e(hu) du$  is the Fourier transform.*

*Proof.* By the usual Poisson summation formula we have

$$\sum_{n \equiv a \pmod{q}} F_N(n) = \frac{N}{q} \sum_h \hat{F}\left(\frac{hN}{q}\right) e_q(-ah).$$

For  $|h| > H$ , we have by integration by parts  $j \geq 2$  times

$$\hat{F}\left(\frac{hN}{q}\right) = \int F(u) e(uhN/q) du \ll_j \nu^{-j+1} (hN/q)^{-j} \ll_j \nu (QN)^{-j\varepsilon} (h/H)^{-2},$$

which gives the result.  $\square$

We will also need the two-dimensional Poisson summation formula.

LEMMA 2.8. Let  $F: \mathbb{R}^2 \rightarrow \mathbb{C}$  be a  $C^\infty$ -smooth compactly supported function such that

$$|F^{(j,k)}| \ll \nu_1^{-j} \nu_2^{-k}, \quad \int_{\mathbb{R}^2} |F| \ll \nu_1 \nu_2$$

and let  $F_N(\mathbf{n}) := F(n_1/N_1, n_2/N_2)$ . Then

$$\sum_{\mathbf{n} \in \mathbb{Z}^2} F_N(\mathbf{n}) = \sum_{\mathbf{h} \in \mathbb{Z}^2} \widehat{F_N}(\mathbf{h})$$

and

$$\widehat{F_N}(\mathbf{h}) \ll_C \nu_1 \nu_2 N_1 N_2 (1 + \nu_1 |h_1|/N_1)^{-C} (1 + \nu_2 |h_2|/N_2)^{-C}.$$

## 2.5 Fourier expansions

We have the following lemma on Mellin transforms, where the construction of the non-negative majorant  $\tilde{F}$  is from the proof of [IK04, Lemma 7.1] (cf. the function  $g(y)$ , with  $x_m = \log m$ ).

LEMMA 2.9. Fix a non-negative  $C^\infty$ -smooth function  $F$  supported on  $[1 - \nu, 1 + \nu]$  and satisfying

$$|F^{(j)}| \ll_j \nu^{-j}, \quad j \geq 0 \quad \text{and} \quad \int_{1/2}^2 F(1/t) \frac{dt}{t} = \nu.$$

Then, for any  $c \in \mathbb{R}$ ,

$$F(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \dot{F}(s) x^{-s} ds,$$

where the Mellin transform is

$$\dot{F}(s) := \int_0^\infty F(x) x^s \frac{dx}{x} \ll_{C,\sigma} \nu (1 + \nu |s|)^{-C}.$$

Furthermore,  $|\dot{F}(it)|$  has a smooth majorant  $\tilde{F}(t)$  such that, for  $m, n \sim M$ ,

$$\left| \int \tilde{F}(t) (m/n)^{it} dt \right| \ll \mathbf{1}_{|m-n| \leq \nu M}.$$

Similarly, we have the following lemma on Fourier series, which will be applied with the characters

$$e^{ik \arg z} = \left( \frac{z}{|z|} \right)^k$$

to expand smooth weights  $G(\arg z)$ .

LEMMA 2.10. Let  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$  and let  $G$  be a bounded smooth function supported on  $[\theta - \nu, \theta + \nu]$  and satisfying

$$|G^{(j)}| \ll_j \nu^{-j}, \quad j \geq 0 \quad \text{and} \quad \int_{\mathbb{R}/2\pi\mathbb{Z}} G(\alpha) d\alpha = \nu.$$

Then

$$G(\alpha) = \sum_k \check{G}(k) e^{ik\alpha}$$

with

$$\check{G}(k) := \int_{\mathbb{R}/2\pi\mathbb{Z}} F(\alpha) e^{-ik\alpha} d\alpha \ll_C \nu(1 + \nu|k|)^{-C}.$$

Furthermore, there is a majorant  $\tilde{G}(k)$  of  $|\check{G}(k)|$  such that

$$\left| \sum_k \tilde{G}(k) e^{ikx} \right| \ll \mathbf{1}_{|x| \leq \nu}.$$

## 2.6 Large sieve bounds

For Type II sums we need the multiplicative large sieve inequality of Bombieri and Davenport (cf., for instance, [FI10, (9.52)]).

LEMMA 2.11. *For any complex numbers  $\gamma_n$ , we have*

$$\sum_{d \leq D} \frac{d}{\varphi(d)} \sum_{\chi(d)}^* \left| \sum_{n \leq N} \gamma_n \chi(n) \right|^2 \ll (D^2 + N) \sum_{n \leq N} |\gamma_n|^2.$$

For Type I sums we need the large sieve inequality for roots of quadratic congruences (cf. [FI97] and, in particular, [FI05, Lemma 14.4] for this variant with the twist by  $\bar{q}$ ).

LEMMA 2.12. *Let  $q \geq 1$ . For any complex numbers  $\gamma_n$  we have*

$$\sum_{\substack{d \sim D \\ (d,q)=1}} \sum_{\nu^2+1 \equiv 0 \pmod{d}} \left| \sum_{n \leq N} \gamma_n e_d(\nu n \bar{q}) \right|^2 \ll (qD + N) \sum_{n \leq N} |\gamma_n|^2.$$

## 2.7 Zeros of Hecke $L$ -functions

We say that a Gaussian integer  $z$  is primary if  $z \equiv 1 \pmod{2(1+i)}$ , so that every odd ideal of  $\mathbb{Z}[i]$  has a unique primary generator. Note that this definition is multiplicative. We extend any function  $\psi: \mathbb{Z}[i] \rightarrow \mathbb{C}$  to a function on odd ideals by defining  $\psi(\mathfrak{a}) := \psi(z)$  if  $z$  is the primary generator of  $\mathfrak{a}$ . For  $k \in \mathbb{Z}$ , we let  $\xi_k$  denote the character

$$\xi_k(z) := \left( \frac{z}{|z|} \right)^k = e^{ik \arg z},$$

which controls the angular distribution of  $z$ . For a Dirichlet character  $\chi \in (\widehat{\mathbb{Z}[i]/u\mathbb{Z}[i]})^\times$  with a modulus  $u \in \mathbb{Z}[i] \setminus \{0\}$  we define the Hecke  $L$ -function by

$$L(s, \xi_k \chi) := \sum_{\substack{\mathfrak{a} \subseteq \mathbb{Z}[i] \\ (\mathfrak{a}, 2)=1}} \frac{\xi_k(\mathfrak{a}) \chi(\mathfrak{a})}{(\mathcal{N}\mathfrak{a})^s} = \sum_{\substack{z \in \mathbb{Z}[i] \setminus \{0\} \\ z \text{ primary}}} \frac{\xi_k(z) \chi(z)}{|z|^{2s}}.$$

For a modulus  $u$  we define

$$L_u(s, \xi_k) := \prod_{\chi \in (\widehat{\mathbb{Z}[i]/u\mathbb{Z}[i]})^\times} L(s, \xi_k \chi).$$

We have the following lemmas, where all the constants are effectively computable. We will not need the Deuring–Heilbronn zero repulsion as we deal with the case of a Siegel zero via a different method. The first lemma is classical (cf., for instance, [IK04, Chapter 5]).

LEMMA 2.13 (Zero-free region, Landau–Page). *There is a constant  $c_1 > 0$  such that the function  $L_u(s, \xi_k)$  has at most one zero in the region*

$$\sigma > 1 - \frac{c_1}{\log(|u|(2+|t|)(2+|k|))}.$$

*If such a zero exists, then it is real and simple,  $k=0$ , and it is a zero of some  $L(s, \chi_1)$  with a quadratic character  $\chi_1$ .*

We let  $N^*(\alpha, T, K, Q)$  denote the number of zeros of  $L(s, \xi_k \chi)$  with primitive characters  $\chi$ ,  $|k| \leq K$ , and modulus  $|u|^2 \leq Q$  with  $\sigma \geq \alpha$ ,  $|t| \leq T$ . The following lemma is a generalization of [Gal70, Theorem 6] to Gaussian integers.

LEMMA 2.14 (Log-free zero-density estimate). *There is some constants  $c_2, c_3 > 0$  such that*

$$N^*(\alpha, T, K, Q) \leq c_3(Q^2 KT)^{c_2(1-\alpha)}.$$

As a corollary to the zero-free region (Lemma 2.13) we get the following lemma, by taking  $\delta > 0$  small enough in terms of  $c_1$  and for two different moduli  $u_1, u_2$  applying Lemma 2.13 with  $u = u_1 u_2$ .

LEMMA 2.15. *Let  $\delta > 0$  be sufficiently small in terms of  $c_1$ . Then there is at most one modulus  $|u_1| \leq X^{2\delta}$  with a primitive character  $\chi_1$  such that  $L(s, \xi_k \chi_1)$  has a zero  $\beta_1 \geq 1 - \frac{1}{\sqrt{\delta} \log X}$ . If such a zero exists, then it is real and simple,  $k=0$ , and it is a zero of some  $L(s, \chi_1)$  with a real character  $\chi_1$ .*

The following lemma is proved by the same argument as in [MV07, (11.7)].

LEMMA 2.16. *Let  $\chi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times$  for some  $|u|^2 \leq Q$  and let  $|k| \leq Q$ . If  $L(s, \xi_k \chi)$  has no zeros counted by  $N^*(\alpha, T, K, Q)$ , then for all  $\sigma > (1+\alpha)/2$ ,  $|t| \leq T$ ,  $|k| \leq K$  we have*

$$|L(s, \xi_k \chi)|^{-1} \ll \log Q.$$

## 2.8 Character sums

We need the following lemma for computing sums over primitive characters.

LEMMA 2.17. *For any  $a$  we have*

$$\left| \sum_{\chi(d)}^* \chi(a) \right| \leq (a-1, d),$$

where the sum extends over primitive Dirichlet characters of  $(\mathbb{Z}/d\mathbb{Z})^\times$ .

*Proof.* By the Chinese remainder theorem we have

$$S(a; d) := \sum_{\chi(d)}^* \chi(a) = \prod_{p^k || d} S(a; p^k).$$

Thus, it suffices to show that

$$|S(a; p^k)| \leq (a-1, p^k).$$

This follows from

$$\begin{aligned} S(a; p^k) &= \sum_{\chi(p^k)} \chi(a) - \sum_{\chi(p^{k-1})} \chi(a) \\ &= \varphi(p^k) \mathbf{1}_{a \equiv 1 \pmod{p^k}} - \varphi(p^{k-1}) \mathbf{1}_{a \equiv 1 \pmod{p^{k-1}}}. \end{aligned}$$

□

We will need the smoothed Polyá–Vinogradov bound on Gaussian integers, which is a consequence of Poisson summation (Lemma 2.8) and the Gauss sum bound on  $\mathbb{Z}[i]$ .

LEMMA 2.18. *Let  $F: \mathbb{R}^2 \rightarrow \mathbb{C}$  be as in Lemma 2.8 and for  $z = x + iy \in \mathbb{C}$  define  $F(z) := F(x, y)$ . Let  $\chi$  be a character of modulus  $u \in \mathbb{Z}[i]$ . Then*

$$\sum_z F(z) \chi(z) \ll |u|.$$

The following lemma is required for the proof of Theorem 1.6.

LEMMA 2.19. *Let  $k \geq 1$  and let  $p$  be a prime number. Let  $\chi \in (\mathbb{Z}[i]/\widehat{p^k \mathbb{Z}[i]})^\times$  with conductor  $p^k$  or  $\pi^{k_1} \bar{\pi}^{k_2}$  with  $k = \max\{k_1, k_2\}$ . Then*

$$\sum_{r \pmod{p^k}} \chi(r + ip^\ell) \leq 2p^{k/2 + \ell/2}.$$

*Proof.* Let  $q := p^k$ . We have by  $r \mapsto r/s$

$$\sum_{r \pmod{q}} \chi(r + ip^\ell) = \frac{1}{\varphi(q)} \sum_{(s,q)=1} \sum_{r \pmod{q}} \chi(r + ip^\ell) = \frac{1}{\varphi(q)} \sum_{r,s \pmod{q}} \bar{\chi}(s) \chi(r + isp^\ell).$$

The function  $s \mapsto \bar{\chi}(s)$  defines a multiplicative character over the integers and we let  $q_1 | q$  denote its conductor. By expansion of  $\bar{\chi}(s)$  into additive characters we get

$$\left| \sum_{r \pmod{p^k}} \chi(r + ip^\ell) \right| \leq \frac{1}{q_1 \varphi(q)} \sum_{a \pmod{q_1}} \left| \sum_{t \pmod{q_1}} \bar{\chi}(t) e_{q_1}(-at) \right| \left| \sum_{r,s \pmod{q}} e_{q_1}(as) \chi(r + isp^\ell) \right|.$$

We have (writing  $s \mapsto s/p^\ell + tp^{k-\ell}, p^\ell | s$ )

$$\sum_{r,s \pmod{q}} e_{q_1}(as) \chi(r + isp^\ell) = \sum_{\substack{z \pmod{q} \\ s \equiv 0 \pmod{p^\ell}}} e_{q_1}(as/p^\ell) \chi(z) \sum_{t \pmod{p^\ell}} e_{q_1}(atp^{k-\ell}).$$

Here

$$\sum_{t \pmod{p^\ell}} e_{q_1}(atp^{k-\ell}) = p^\ell \mathbf{1}_{a \equiv 0 \pmod{q_1/(q_1, p^{k-\ell})}}.$$

Hence, denoting  $q_2 = (q_1, p^{k-\ell})$  and making the change of variables  $a \mapsto aq_1/q_2$ , we get

$$\left| \sum_{r \pmod{p^k}} \chi(r + ip^\ell) \right| \leq \frac{p^\ell}{q_1 \varphi(q)} \sum_{a \pmod{q_2}} \left| \sum_{t \pmod{q_1}} \bar{\chi}(t) e_{q_2}(-at) \right| \left| \sum_{\substack{z \pmod{q} \\ s \equiv 0 \pmod{p^\ell}}} e_{p^\ell q_2}(as) \chi(z) \right|.$$

Since  $\chi(t)$  is of conductor  $q_1$  we have the Gauss sum bound [MV07, Theorem 9.12]

$$\left| \sum_{t \pmod{q_1}} \bar{\chi}(t) e_{q_2}(-at) \right| \leq q_1^{1/2}. \quad (2.1)$$

For the second sum we expand the condition  $s \equiv 0 \pmod{p^\ell}$  with additive characters to get

$$\sum_{\substack{z \pmod{q} \\ s \equiv 0 \pmod{p^\ell}}} e_{p^\ell q_2}(as) \chi(z) = \frac{1}{p^\ell} \sum_{b \pmod{p^\ell}} \sum_{z \pmod{q}} e_{p^\ell q_2}(as) e_{p^\ell}(bs) \chi(z).$$

The function

$$z \mapsto e_{p^\ell q_2}(as) e_{p^\ell}(bs) = e\left(\frac{(a + q_2 b)s}{q_2 p^\ell}\right) =: e_q(cs)$$

is an additive character on  $\mathbb{Z}[i]/q\mathbb{Z}[i]$  since  $q_2|p^{k-\ell}$ . If  $\chi$  is of conductor  $\pi^{k_1}\bar{\pi}^{k_2}$ , then we get by the Gauss sum bound (a direct generalization of [MV07, Theorem 9.12] to  $\mathbb{Z}[i]$ , denoting  $r := q/(c, q)$ )

$$\left| \sum_{z(q)} e_q(cs)\chi(z) \right| \leq \mathbf{1}_{\pi^{k_1}\bar{\pi}^{k_2}|r} \frac{\varphi_{\mathbb{Z}[i]}(q)}{\varphi_{\mathbb{Z}[i]}(r)} q \leq q, \quad (2.2)$$

since  $\pi^{k_1}\bar{\pi}^{k_2}|r$  and  $k = \max\{k_1, k_2\}$  implies that  $r = p^k = q$ . We also get the same bound if the conductor of  $\chi$  is  $p^k$ , as then  $\chi$  is primitive. Putting the bounds (2.1) and (2.2) together and using  $q_2 = (q_1, p^{k-\ell})$  we get (noting that  $q/\varphi(q) \leq 2$  for  $q = p^k$ )

$$\begin{aligned} \left| \sum_{r(p^k)} \chi(r + ip^\ell) \right| &\leq \frac{p^\ell}{q_1\varphi(q)} q_2 q_1^{1/2} q \leq 2p^\ell q_2 q_1^{-1/2} \\ &\leq 2p^\ell q_2^{1/2} \leq 2p^{k/2+\ell/2}. \end{aligned} \quad \square$$

The previous lemma implies the following.

LEMMA 2.20. *Let  $b \in \mathbb{Z}$  and  $u \in \mathbb{Z}[i]$  and let  $Y > |u|^4$ . Let  $\chi$  be a primitive character modulo  $u$  and let  $v = (u, b)$ . Then, for any integer  $a_0$ ,*

$$\sum_{\substack{a \in (Y, Y+Y^{1-\eta}] \\ a \equiv a_0 \pmod{4} \\ (a, b) = 1}} \chi(b + ia) \ll_\varepsilon \frac{Y^{1-\eta}}{|u/v|^{1/2-\varepsilon}}.$$

*Proof.* We have

$$\sum_{\substack{a \in (Y, Y+Y^{1-\eta}] \\ a \equiv a_0 \pmod{4} \\ (a, b) = 1}} \chi(b + ia) = \sum_{\substack{c|b \\ (c, 4u) = 1}} \mu(c) \sum_{\substack{ac \in (Y, Y+Y^{1-\eta}] \\ ac \equiv a_0 \pmod{4}}} \chi(b + iac).$$

Let  $u = mw$ , where  $m$  consists of all prime factors  $p \equiv 3 \pmod{4}$ . Let  $n$  denote the smallest integer such that  $w|n$ . The contribution from  $c > Y/mn$  is trivially bounded by

$$\ll \sum_{\substack{c|b \\ c > Y/mn}} \frac{Y}{c} \ll mn\tau(b).$$

For  $c \leq Y/mn$ , we have

$$\sum_{\substack{ac \in (Y, Y+Y^{1-\eta}] \\ ac \equiv a_0 \pmod{4}}} \chi(b + iac) = \frac{Y^{1-\eta}}{16cmn} \sum_{\substack{r(4mn) \\ rc \equiv a_0 \pmod{4}}} \chi(b + irc) + O(mn) \ll \frac{Y^{1-\eta}(b, mn)^{1/2}}{(mn)^{1/2}},$$

once we show that

$$\sum_{\substack{r(4mn) \\ rc \equiv a_0 \pmod{4}}} \chi(b + irc) = \sum_{\substack{r(4mn) \\ r \equiv a_0 \pmod{4}}} \chi(b + ir) \leq (mn)^{1/2}(b, mn)^{1/2}.$$

To prove this, write

$$\chi = \prod_{p^k || m} \chi_{p^k} \prod_{\pi^k || w} \chi_{\pi^k}.$$



By the Chinese remainder theorem, we get (denoting  $k = \max\{k_1, k_2\}$ )

$$\sum_{\substack{r(4mn) \\ r \equiv a_0(4)}} \chi(b + ir) = \left( \prod_{p^k \parallel m} \sum_{r(p^k)} \chi_{p^k}(b + ir) \right) \left( \prod_{\pi^{k_1} \bar{\pi}^{k_2} \parallel w} \sum_{r(p^k)} \chi_{\pi^{k_1}} \chi_{\bar{\pi}^{k_2}}(b + ir) \right)$$

and the claim follows by Lemma 2.19.  $\square$

### 3. Set-up and statement of the quasi-explicit formula

Let  $\lambda_b$  be divisor-bounded coefficients and define

$$\omega_2(b) := 2 \prod_{\substack{p|b \\ p \neq 2}} \left( 1 - \frac{\rho(p)}{p} \right)^{-1}.$$

Define the sequences over Gaussian integers  $\mathcal{A} = (a_z)$  with

$$a_z := \lambda_{\operatorname{Re}(z)} \mathbf{1}_{(z, \bar{z})=1}, \quad a_z^\omega := \lambda_{\operatorname{Re}(z)} \omega_2(\operatorname{Re}(z)) \mathbf{1}_{(z, \bar{z})=1}.$$

Note that  $(z, \bar{z}) = 1$  implies that  $(z, (1+i)) = 1$  and for  $z = b + ia$  that  $(a, b) = 1$ .

For an ideal  $\mathfrak{n} = (z)$  we define

$$a_{\mathfrak{n}} = \sum_{u \in \{\pm 1, \pm i\}} a_{uz}, \quad a_{\mathfrak{n}}^\omega = \sum_{u \in \{\pm 1, \pm i\}} a_{uz}^\omega$$

so that for any function  $f$  on the ideals we have

$$\sum_{\mathcal{N}\mathfrak{n} \sim X} a_{\mathfrak{n}} f(\mathfrak{n}) = \sum_{|z|^2 \sim X} a_z f((z)) = \sum_{\substack{a^2 + b^2 \sim X \\ (a, b) = 1 \\ (a^2 + b^2, 2) = 1}} \lambda_b f((b + ia)).$$

In the case that there is a Siegel zero for the character  $\chi_1$ , for any finite set of integers  $B$  let  $B_1 \subseteq B$  be the largest subset such that for all  $b \in B_1$  we have

$$\sum_{\substack{a \sim (X - b^2)^{1/2} \\ (a, b) = 1 \\ (a^2 + b^2, 2) = 1}} \chi_1((b + ia)) > 0 \quad (3.1)$$

and let us denote

$$\Omega(B) := \sum_{\mathcal{N}\mathfrak{n} \sim X} a_{\mathfrak{n}}^\omega \quad \text{for } \lambda_b := \mathbf{1}_B(b).$$

We split the proof of Theorem 1.2 into two cases depending on whether there is a Siegel zero  $\beta_1 > 1 - \varepsilon_1 / \log X$  or not and according to the size of  $\Omega(B_1)$ . Theorem 1.2 is then an immediate corollary of the following two theorems.

**THEOREM 3.1** (Exceptional case). *Let  $\varepsilon_1 \in (0, 1/10)$ . Let  $B \subseteq [\eta X^{1/2}, (1 - \eta)(2X)^{1/2}] \cap \mathbb{Z}$  with  $|B| = X^{1/2 - \delta}$  and let  $\lambda_b = \mathbf{1}_B(b)$ . Suppose that for some  $\chi_1$  of modulus  $\mathcal{M}(u_1) \leq X^{\delta + \eta}$  the  $L$ -function  $L(s, \chi_1)$  has a zero  $\beta_1 > 1 - \varepsilon_1 / \log X$  and that  $\Omega(B_1) \geq \Omega(B)/2$ . Suppose that  $\delta$  is sufficiently small in terms of  $\varepsilon_1$ . Then we have, for all  $\varepsilon > 0$ ,*

$$\sum_{\mathcal{N}\mathfrak{p} \sim X} a_{\mathfrak{p}} \gg_{\varepsilon} X^{1/2 - \varepsilon} |B|.$$

**THEOREM 3.2** (Regular case). *Let  $\varepsilon_1 \in (0, 1/10)$ . Let  $B \subseteq [\eta X^{1/2}, (1 - \eta)(2X)^{1/2}] \cap \mathbb{Z}$  with  $|B| = X^{1/2-\delta}$  and let  $\lambda_b = \mathbf{1}_B(b)$ . Suppose that the  $L$ -functions  $L(s, \chi)$  have no zeros  $\beta > 1 - \varepsilon_1 / \log X$  or that  $\Omega(B_1) \leq \Omega(B)/2$  and suppose that  $\delta$  is sufficiently small in terms of  $\varepsilon_1$ . Then*

$$\sum_{\mathcal{N} \sim X} a_{\mathfrak{p}} \gg_{\varepsilon_1} \frac{1}{\log X} \sum_{\mathcal{N} \sim X} a_{\mathfrak{n}}^{\omega}.$$

In Theorem 3.2 the possible exceptional zero is not a Siegel zero, the potential zeros are only somewhat close to the line  $\operatorname{Re}(s) = 1$  and we show that these zeros can have only a small influence on the main term. This means that in all error terms it suffices to save only a power of  $\log X$  instead of a power of  $X$ . Theorem 3.1 is proved in § 5. Theorem 3.2 is proved in § 10 and is a quick consequence of the following result. Recall that  $\mathcal{M}(u)$  denotes the smallest integer  $m$  with  $u|m$ .

**THEOREM 3.3** (Quasi-explicit formula). *Let  $\eta > 0$  be small. Let  $\delta \in (0, 1/10)$  and  $X \gg 1$ . For every  $C_1 > 0$ , there is some  $C_2 > 0$  such that for some*

$$J \leq (\log X)^{C_2},$$

*there is a set of primitive Hecke characters  $\{\xi_{k_j} \chi_j\}_{j \leq J}$  with Dirichlet characters  $\chi_j$  to moduli  $u_j \in \mathbb{Z}[i]$  with  $\mathcal{M}(u_j) \leq X^{\delta+\eta}$  and  $|k_j| \leq X^{\eta}$  such that the following holds. Let  $\lambda_b$  be coefficients with  $|\lambda_b| \leq X^{o(1)}$ , supported on  $[\eta X^{1/2}, (1 - \eta)(2X)^{1/2}] \cap \mathbb{Z}$ , and satisfying*

$$\sum_b |\lambda_b| \geq X^{1/2-\delta}.$$

*Then*

$$\begin{aligned} \sum_{\mathcal{N} \sim X} a_{\mathfrak{n}} \Lambda(\mathcal{N} \mathfrak{n}) &= \frac{4}{\pi} \sum_{\mathcal{N} \sim X} a_{\mathfrak{n}}^{\omega} \left( 1 - \sum_{j \leq J} \overline{\xi_{k_j} \chi_j}(\mathfrak{n}) \sum_{\substack{\rho_j \\ L(\rho_j, \xi_{k_j} \chi_j) = 0 \\ |\operatorname{Im}(\rho_j)| \leq X^{\eta}}} (\mathcal{N} \mathfrak{n})^{\rho_j - 1} \right) \\ &\quad + O\left(\frac{1}{(\log X)^{C_1}} X^{1/2} \sum_b |\lambda_b|\right). \end{aligned}$$

The restriction  $\mathcal{M}(u_j) \leq X^{\delta+\eta}$  instead of a condition involving  $|u_j|$  may appear unusual, but, in fact, it occurs very naturally in the proof of the Type II estimate (Proposition 6.2), where we consider the distribution of Gaussian integers in arithmetic progressions to moduli  $d$  which are regular integers, so that  $u_j|d$ . We prove Theorem 3.3 in § 11. It is possible that the range  $\delta < 1/10$  may be improved a bit with further work but we seem to hit a hard barrier at  $\delta = 1/6$ , cf. Remark 8.1 for more details. It is also plausible that one could obtain a power saving in the error term by taking into account  $J \leq X^{\eta}$  bad characters. The factors of  $X^{\eta}$  in the ranges for  $u_j, k_j, \operatorname{Im}(\rho_j)$  may be replaced by  $(\log X)^{C_3}$  for some large  $C_3 > 0$  but this is inconsequential for our applications. We also note that the right-hand side may be expressed as a sum over Gaussian integers by writing

$$\begin{aligned} \sum_{\mathcal{N} \sim X} a_{\mathfrak{n}}^{\omega} \overline{\xi_{k_j} \chi_j}(\mathfrak{n}) (\mathcal{N} \mathfrak{n})^{\rho_j - 1} &= \sum_{u \in \{\pm 1, \pm i\}} \sum_{|z|^2 \sim X} a_{uz}^{\omega} \overline{\xi_{k_j} \chi_j}(z) (|z|^2)^{\rho_j - 1} \mathbf{1}_{z \equiv 1 \pmod{2(1+i)}} \\ &= \frac{1}{4} \sum_{\chi \in \mathbb{Z}[i]/2(1+i)\mathbb{Z}[i]} \sum_{u \in \{\pm 1, \pm i\}} \sum_{|z|^2 \sim X} a_{uz}^{\omega} \overline{\xi_{k_j} \chi_j}(z) (|z|^2)^{\rho_j - 1} \\ &= \sum_{\chi \in \mathbb{Z}[i]/2(1+i)\mathbb{Z}[i]} \mathbf{1}_{\xi_{k_j} \chi_j(i) = 1} \sum_{|z|^2 \sim X} a_z^{\omega} \overline{\xi_{k_j} \chi_j}(z) (|z|^2)^{\rho_j - 1}, \end{aligned}$$

where the last equation follows from the change of variables  $z \mapsto z/u$  and summing over  $u$ .

We conclude this section by making elementary reductions for the proof of Theorem 3.3, to reduce to the case when  $\lambda_b = \mathbf{1}_B(b)$  with  $B$  in a short interval.

### 3.1 Reduction to bounded $\lambda_b$

We can reduce from coefficients satisfying  $|\lambda_b| \leq X^{o(1)}$  to  $|\lambda_b| \leq 1$  as follows. Let  $\varepsilon > 0$  be small enough so that  $\delta + 2\varepsilon < 1/10$  and denote

$$B_0 := \{b : 0 < |\lambda_b| \leq 1\} \quad \text{and} \quad B_j := \{b : 2^{j-1} < |\lambda_b| \leq 2^j\}, \quad 1 \leq j \leq \log X.$$

The contribution from  $b \in B_j$  with  $|B_j| \leq X^{1/2-\delta-\varepsilon}$  is negligible, since for such  $j$  by  $|\lambda_b| \leq X^{o(1)}$

$$\sum_{b \in B_j} |\lambda_b| \ll X^{o(1)} |B_j| \ll X^{1/2-\delta-\varepsilon+o(1)}.$$

For each  $j$  such that  $|B_j| > X^{1/2-\delta-\varepsilon}$ , we can renormalize  $\lambda_b$  by  $2^{-j}$  to get bounded coefficients which satisfy

$$\sum_{b \in B_j} |2^{-j} \lambda_b| \geq \frac{1}{2} |B_j| \geq X^{1/2-\delta-2\varepsilon}.$$

Then the general case follows by applying the bounded case for each such  $j$  separately with the weights  $\lambda_b^{(j)} := 2^{-j} \lambda_b \mathbf{1}_{b \in B_j}$ . Indeed, if we denote the claim in Theorem 3.3 by  $S(\lambda) = M(\lambda) + O(E(\lambda))$ , then assuming that the theorem holds for the bounded  $\lambda_b^{(j)}$ , we get by linearity

$$S(\lambda) = \sum_j 2^j S(\lambda^{(j)}) = \sum_j 2^j (M(\lambda^{(j)}) + O(E(\lambda^{(j)}))) = M(\lambda) + O(E(\lambda)).$$

Thus, it suffices to prove Theorem 3.3 for  $|\lambda_b| \leq 1$ .

### 3.2 Reduction from $\lambda_b$ to $\mathbf{1}_B$

We can reduce the proof from general bounded weights  $\lambda_b$  to the weights of the type  $\mathbf{1}_B(b)$  by a finer-than-dyadic decomposition in terms of the values of  $\lambda_b$ . That is, for  $\nu = (\log X)^{-C}$  we write

$$I_j := \begin{cases} ((1-\nu)^{j+1}, (1-\nu')^j] & 0 \leq j < (\log X)/\nu, \\ [0, (1-\nu)^j], & j = \lfloor (\log X)/\nu \rfloor, \end{cases}$$

$$\mathbf{1}_{|z|_\infty \leq 1} = \sum_{j_1, j_2} \mathbf{1}_{z \in I_{j_2} \times I_{j_1}}.$$

Since  $|\lambda_b| \leq 1$ , we obtain a partition

$$\lambda_b = \sum_{j_1, j_2} (\mathbf{1}_{\lambda_b \in I_{j_2} \times I_{j_1}} ((1-\nu)^{j_1} + i(1-\nu)^{j_2}) + O(\nu |\lambda_b| \mathbf{1}_{\lambda_b \in I_{j_2} \times I_{j_1}})),$$

where the contribution from the error term is negligible by crude bounds. We consider  $\mathbf{1}_B$  for

$$B := B(j_1, j_2) = \{b \in [\eta X^{1/2}, (1-\eta) X^{1/2}] : \lambda_b \in I_{j_2} \times I_{j_1}\}$$

and note that for  $j_1, j_2$  with  $|B(j_1, j_2)| < X^{1/2-\delta-\eta}$  we can bound the contribution trivially. Hence, it suffices to show Theorem 3.3 for  $\lambda_b = \mathbf{1}_B(b)$ .

### 3.3 Reduction to $B$ in a short interval

Let  $B \subseteq [\eta X^{1/2}, (2-\eta) X^{1/2}]$  and split

$$B = \bigcup_j B_j, \quad B_j := B \cap [j X^{1/2-\eta}, (j+1) X^{1/2-\eta}].$$

The contribution from  $B_j$  with  $|B_j| \leq X^{-2\eta}|B|$  is negligible by a crude bound. Thus, we only need to deal with  $|B_j| \geq |B|X^{-2\eta} = X^{1/2-\delta-2\eta}$ . Therefore, it suffices to prove Theorem 3.3 for  $\lambda_b = \mathbf{1}_B(b)$  with

$$B \subseteq [Y, Y + X^{1/2-\eta}] \quad \text{for some } Y \in [\eta X^{1/2}, (2-\eta)X^{1/2}]. \quad (3.2)$$

From now on we always assume that  $\lambda_b = \mathbf{1}_B(b)$  for a set  $B$  satisfying (3.2).

#### 4. Type I information

For  $b \in B$  and for any function  $f$  on the ideals of  $\mathbb{Z}[i]$  denote

$$a_{z,f}^{(b)} := \mathbf{1}_{\operatorname{Re}(z)=b} \mathbf{1}_{(z,\bar{z}=1)} f((z)).$$

Define

$$g^{(b)}(w) := \mathbf{1}_{(w,b\bar{w})=1} \frac{1}{|w|^2}.$$

We have Type I information provided by the following proposition.

**PROPOSITION 4.1** (Type I information). *Let  $\alpha_w$  be divisor-bounded coefficients supported on  $(w, 2\bar{w}) = 1$ . Let  $\chi$  be a Dirichlet character to modulus  $u$  and let  $\xi = \xi_k$  with  $|k| \ll X^{\eta/1000}$ . Let  $q := \mathcal{M}(u) \leq X^{1/4}$ . Then*

$$\sum_{b \in B} \left| \sum_{\substack{|w|^2 \leq X^{1-\delta-\eta}/q^2 \\ (w,u)=1}} \alpha_w \left( \sum_{\substack{|z|^2 \sim X \\ z \equiv 0 \pmod{w}}} a_{z,\xi\chi}^{(b)} - g^{(b)}(w) \sum_{\substack{a \sim (X^2-b^2)^{1/2} \\ (a,b)=1 \\ (a^2+b^2,2)=1}} \xi\chi((b+ia)) \right) \right| \\ \ll X^{1/2-\eta/100} |B|.$$

The same is true if  $|z|^2 \sim X$  is replaced by  $|z|^2 \in [X', X'(1+\nu)]$  or by a smooth weight  $F(|z|^2/X')$  as in § 2.1 with  $X' \sim X$  and  $\nu \geq X^{-\eta/1000}$ , with the same change applied to the condition  $a \sim (X^2 - b^2)^{1/2}$ .

*Proof.* Let us split  $d$  dyadically into  $d \sim D \leq X^{1-\delta-\eta}/q^2$  and denote

$$S^{(b)}(\alpha) := \sum_{\substack{|w|^2 \sim D \\ (w,u)=1}} \alpha_w \left( \sum_{\substack{|z|^2 \sim X \\ z \equiv 0 \pmod{w}}} a_{z,\xi\chi}^{(b)} - g^{(b)}(w) \sum_{\substack{a \sim (X^2-b^2)^{1/2} \\ (a,b)=1 \\ (a^2+b^2,2)=1}} \xi\chi((b+ia)) \right).$$

We apply § 2.1 with  $\nu = X^{-\eta/80}$  to split  $a$  into finer than dyadic ranges to get (bounding the part  $A \leq X^{1/2-\eta}$  trivially using the divisor bound and dropping the condition  $a^2 + b^2 \sim X$ )

$$S^{(b)}(\alpha) = \frac{1}{\nu} \int_{X^{1/2-\eta}}^{2X^{1/2}} S^{(b)}(\alpha, A) \frac{dA}{A} + O(X^{1/2-\eta/100}),$$

with

$$S^{(b)}(\alpha, A) := \sum_{\substack{|w|^2 \sim D \\ (w, u)=1}} \alpha_w \left( \sum_{\substack{a \equiv -bi(w) \\ (a, b)=1 \\ (a^2+b^2, 2)=1}} F_A(a) \xi \chi((b+ia)) - g^{(b)}(w) \sum_{\substack{(a, b)=1 \\ (a^2+b^2, 2)=1}} F_A(a) \xi \chi((b+ia)) \right).$$

Since  $b+ia$  is odd, we have

$$\xi \chi((b+ia)) = \xi \chi(u) \xi \chi(b+ia)$$

for some unit  $u$  depending on  $z$  modulo  $2(1+i)$ . Since  $a$  is restricted to a short interval, for a fixed  $b$  the function  $\xi_k(b+ia)$  with  $k \ll X^{\eta/1000}$  is equal to a constant up to a negligible error term and may therefore be dropped. By splitting  $B$  into residue classes modulo 4 we may assume that  $u$  does not depend on  $b$ . Thus, we can split  $a$  into congruence classes modulo  $4q$  and  $d = |w|^2$  with  $a \equiv a_0 d \pmod{4q}$  to get, for some unit  $u_{a_0 d}$ ,

$$S^{(b)}(\alpha, A) = \sum_{\substack{|w|^2 \sim D \\ (2bu, w)=1}} \alpha_w \sum_{a_0 \pmod{4q}} \xi \chi(u_{a_0 d}) \chi(b+ia_0 d) \\ \times \sum_{\substack{\nu \in \mathbb{Z}/d\mathbb{Z} \\ \nu \equiv -i(w)}} \left( \sum_{\substack{a \equiv \nu b \pmod{d} \\ a \equiv a_0 d \pmod{4q} \\ (a, b)=1}} F_A(a) - g^{(b)}(w) \sum_{\substack{a \equiv a_0 d \pmod{4q} \\ (a, b)=1}} F_A(a) \right),$$

where the sum over  $\nu$  contains just the one element  $\nu = -r/s$  for  $w = r+is$ . Note that  $(a, b) = 1$  implies that  $(w, b) = 1$  and that  $(w, \bar{w}) = 1$ , so that  $(r, s) = 1$  and we may restrict to  $\nu \in \mathbb{Z}/d\mathbb{Z}$  in the above.

It suffices to show that

$$\sum_{b \in B} \left| \sum_{\substack{|w|^2 \sim D \\ (2bu, w)=1}} \alpha_w \sum_{a_0 \pmod{4q}} \xi \chi(u_{a_0 d}) \chi(b+ia_0 d) \right. \\ \left. \times \sum_{\substack{\nu \pmod{d} \\ \nu \equiv -i(w)}} \left( \sum_{\substack{a \equiv \nu b \pmod{d} \\ a \equiv a_0 d \pmod{4q} \\ (a, b)=1}} F_A(a) - g^{(b)}(w) \sum_{\substack{a \equiv a_0 d \pmod{4q} \\ (a, b)=1}} F_A(a) \right) \right| \\ \ll X^{1/2-\eta/40} |B|.$$

Let us first deal with the cross-condition  $\chi(b+ia_0 d)$  between  $b$  and  $d$ . If  $(b, q) = 1$ , we could just make the change of variables  $a_0 \mapsto a_0 b$  and  $\chi(b)$  would factor out. In general, we have for some characters  $\chi_{p^k}$  modulo  $p^k$ :

$$\chi(b+ia_0 d) = \prod_{p^k \parallel q} \chi_{p^k}(b+ia_0 d).$$

Denoting

$$(b, q) = q_1 = \prod_{p^\ell \parallel (b, q)} p^\ell$$

and making the change of variables  $a_0 \mapsto a_0 b/p^\ell$  modulo  $p^k$  we get

$$\chi(b + ia_0 d) = \left( \prod_{p^\ell \parallel q_1} \chi_{p^k}(b/p^\ell) \right) \prod_{\substack{p^k \parallel q \\ p^\ell \parallel q_1}} \chi_{p^k}(p^\ell + ia_0 d).$$

Note that the first factor depends only on  $b$  and the second factor no longer depends on  $b$  but on  $q_1$ . The residue classes  $b/p^\ell$  and  $p^\ell$  combine to unique residue classes  $\theta_{b,q}$  and  $\gamma_{q_1,q}$  modulo  $8q$  by the Chinese remainder theorem. Thus, we get

$$\sum_{q_1 \mid q} \sum_{\substack{b \in B \\ (b,q)=q_1}} \left| \sum_{\substack{|w|^2 \sim D \\ (2bu,w)=1}} \alpha_w \sum_{a_0 \pmod{4q}} \xi \chi(u_{a_0 d}) \chi(\gamma_{q_1,q} + ia_0 d) \right. \\ \left. \times \sum_{\substack{\nu \pmod{d} \\ \nu \equiv -i(w)}} \left( \sum_{\substack{a \equiv \nu b \pmod{d} \\ a \equiv a_0 \theta_{b,q} d \pmod{4q} \\ (a,b)=1}} F_A(a) - g^{(b)}(w) \sum_{\substack{a \equiv a_0 \theta_{b,q} d \pmod{4q} \\ (a,b)=1}} F_A(a) \right) \right|.$$

Dropping the condition  $(b, q) = q_1$ , using the divisor bound for  $\sum_{q_1 \mid q} 1 = \tau(q)$ , taking the sum over  $a_0$  to the outside, and absorbing the factor  $\chi(u_{a_0 d}) \chi(\gamma_{q_1,q} + ia_0 d)$  into the coefficient  $\alpha_w$ , it suffices to show that for any  $a_0$ ,

$$\sum_{\substack{b \in B \\ (b+ia_0, u)=1}} \left| \sum_{\substack{|w|^2 \sim D \\ (2bu,w)=1}} \alpha_w \sum_{\substack{\nu \pmod{d} \\ \nu \equiv -i(w)}} \left( \sum_{\substack{a \equiv \nu b \pmod{d} \\ a \equiv a_0 \theta_{b,q} d \pmod{4q} \\ (a,b)=1}} F_A(a) - g^{(b)}(w) \sum_{\substack{a \equiv a_0 \theta_{b,q} d \pmod{4q} \\ (a,b)=1}} F_A(a) \right) \right| \\ \ll \frac{X^{1/2-\eta/39} |B|}{q}.$$

We expand the condition  $(a, b) = 1$  using the Möbius function and use triangle inequality to get (note that  $(b + ia_0, u) = 1$  implies that  $(c, q) = 1$ )

$$\sum_{\substack{(c,q)=1}} \sum_{\substack{b \in B \\ b \equiv 0 \pmod{c}}} \left| \sum_{\substack{|w|^2 \sim D \\ (2bu,w)=1}} \alpha_w \sum_{\substack{\nu \pmod{d} \\ \nu \equiv -i(w)}} \left( \sum_{\substack{a \equiv \nu b \pmod{d} \\ a \equiv a_0 \theta_{b,q} d \pmod{4q} \\ a \equiv 0 \pmod{c}}} F_A(a) - g^{(b)}(w) \sum_{\substack{a \equiv a_0 \theta_{b,q} d \pmod{4q} \\ a \equiv 0 \pmod{c}}} F_A(a) \right) \right| \\ =: S_{\leq \eta} + S_{> \eta},$$

where we have partitioned the sum depending on whether  $c \leq X^{\eta/20}$  or  $c > X^{\eta/20}$ . The second sum can be bounded trivially by

$$\sum_{a_0 \pmod{4q}} S_{> \eta} \ll \sum_{c > X^{\eta/20}} \sum_{\substack{b \in B \\ b \equiv 0 \pmod{c}}} \sum_{\substack{d \sim D \\ (2bq, d)=1}} \left( \sum_{\substack{\nu^2+1 \equiv 0 \pmod{d}}} \sum_{\substack{a \equiv \nu b \pmod{d} \\ a \equiv 0 \pmod{c}}} F_A(a) + g^{(b)}(d) \sum_{a \equiv 0 \pmod{c}} F_A(a) \right)$$

$$\begin{aligned} &\ll \sum_{c > X^{\eta/20}} \sum_{\substack{b \in B \\ b \equiv 0 \pmod{c}}} \sum_{\substack{d \sim D \\ (2bq, d) = 1}} \left( \sum_{\substack{a^2 + b^2 \equiv 0 \pmod{d} \\ a \equiv 0 \pmod{c}}} F_A(a) + g^{(b)}(d) \sum_{a \equiv 0 \pmod{c}} F_A(a) \right) \\ &\ll \sum_{c > X^{\eta/20}} \sum_{\substack{b \in B \\ b \equiv 0 \pmod{c}}} \left( \sum_{a \equiv 0 \pmod{c}} F_A(a) \tau(a^2 + b^2) + (\log X)^{O(1)} \sum_{a \equiv 0 \pmod{c}} F_A(a) \right). \end{aligned}$$

By the divisor bound for  $\tau(a^2 + b^2)$  and  $\tau(b)$ , we obtain

$$\sum_{a_0(4q)} S_{>\eta} \ll \sum_{b \in B} \sum_{\substack{c|b \\ c > X^{\eta/20}}} X^{\eta/80} \left( \frac{X^{1/2}}{c} + 1 \right) \ll X^{1/2 - \eta/20 + \eta/80} \sum_{b \in B} \tau(b) \ll X^{1/2 - \eta/40} |B|.$$

Hence, it remains show that for any  $a_0$

$$S_{\leq \eta} \ll \frac{X^{1/2 - \eta/40} |B|}{q}.$$

By writing

$$S_{\leq \eta} = \sum_{c \leq X^{\eta/20}} S_c$$

it suffices to show that for every  $c \leq X^{\eta/20}$

$$S_c \ll \frac{X^{1/2 - \eta/10} |B|}{q}.$$

Applying Poisson summation (Lemma 2.7, note that  $d, q, c$  are all pairwise coprime) we get

$$S_c \leq T_c(\alpha, A) + X^{\eta/10} U_c(\alpha, A) + O_\varepsilon(X^{-100}),$$

where for

$$H := \frac{qDX^{\eta/10}}{X^{1/2}}, \quad (4.1)$$

we have (denoting  $\alpha'_w := \alpha_w(D/|w|^2)$ )

$$\begin{aligned} T_c(\alpha, A) &:= \hat{F}(0)A \sum_{\substack{b \in B \\ b \equiv 0 \pmod{c}}} \left| \sum_{\substack{|w|^2 \sim D \\ (2bu, w) = 1}} \alpha_w \left( \frac{1}{c dq} - \frac{g^{(b)}(w)}{cq} \right) \right|, \\ U_c(\alpha, A) &:= \frac{1}{H} \sum_{0 < |h| \leq H} \sum_{\substack{b \in B \\ b \equiv 0 \pmod{c}}} \left| \sum_{\substack{|w|^2 \sim D \\ (2bu, w) = 1}} \alpha'_w \hat{F}(hA/c dq) \sum_{\substack{\nu(d) \\ \nu \equiv -i(w)}} e_d(hb\overline{4cq\nu}) e_{4q}(a_0 d\theta_{b,q} h\overline{cd}) \right|. \end{aligned}$$

Note that the phase  $e_{4q}(a_0 d\theta_{b,q} h\overline{cd}) = e_{4q}(a_0 \theta_{b,q} h\overline{c})$  does not depend on  $d$  or  $\nu$  and thus it may be replaced by 1.

By definition,  $g^{(b)}(w)$  matches with  $1/d$  when  $(b, w) = 1$ , so that, in fact,  $T(\alpha, A) = 0$ . Thus, it suffices to show that

$$U_c(\alpha, A) \ll \frac{X^{1/2 - \eta/5} |B|}{q}. \quad (4.2)$$

4.0.1 *Bounding  $U(\alpha, A)$ .* The smooth cross-condition  $\widehat{F}(hA/c\,dq)$  may be removed by using the Mellin inversion formula (Lemma 2.9) at a cost of a factor  $\ll X^{\eta/40}$ . We expand the condition  $(b, d) = 1$  using the Möbius function to get

$$\mathbf{1}_{(b,d)=1} = \sum_e \mu(e) \mathbf{1}_{e|b} \mathbf{1}_{e|d}.$$

For each  $c$  and  $e$ , let

$$\begin{aligned} V_{c,e}(\alpha, A) &:= \sum_{\substack{|w|^2 \sim D/e \\ (d, 2cq)=1}} \sum_{\substack{\nu(d) \\ \nu \equiv -i(w)}} \left| \sum_{n \ll HX^{1/2}/(ce)} \gamma_n^{(ce)} e_d(\overline{4cq} h b \nu) \right| \\ &= \sum_{\substack{d \sim D/e \\ (d, 2cq)=1}} \sum_{\nu^2+1 \equiv 0(d)} \left| \sum_{n \ll HX^{1/2}/(ce)} \gamma_n^{(ce)} e_d(\overline{4cq} h b \nu) \right|, \end{aligned}$$

for some coefficients

$$|\gamma_n^{(ce)}| \leq \frac{1}{H} \sum_{n=hb} \mathbf{1}_B(bce).$$

Then, by rearranging the sums, we have

$$U_c(\alpha, A) \ll X^{\eta/40} \sum_e V_{c,e}(\alpha, A)$$

and for (4.2) it suffices to show that

$$V_{c,e}(\alpha, A) \ll \frac{X^{1/2-\eta/4} |B|}{eq}.$$

By a divisor bound

$$\sum_{n \ll HX^{1/2}/(ce)} |\gamma_n^{(ce)}|^2 \ll X^{\eta/100} \frac{1}{H} \sum_{n \ll HX^{1/2}/(ce)} |\gamma_n^{(ce)}| \ll X^{\eta/100} \frac{1}{H} |B|.$$

By Cauchy–Schwartz and Lemma 2.12, using (4.1), we have

$$\begin{aligned} V_{c,e}(\alpha, A) &\ll \left( \frac{D}{e} \right)^{1/2} \left( \sum_{\substack{d \sim D/e \\ (d, 2cq)=1}} \sum_{\nu^2+1 \equiv 0(d)} \left| \sum_{n \ll HX^{1/2}/(ce)} \gamma_n^{(ce)} e_d(hb\overline{4cq}\nu) \right|^2 \right)^{1/2} \\ &\ll X^{\eta/100} \frac{D^{1/2}}{eH^{1/2}} (qD + HX^{1/2})^{1/2} |B|^{1/2} \\ &\ll e^{-1} D^{1/2} X^{1/4+\eta/100} |B|^{1/2} \\ &\ll \frac{X^{1-\delta-\eta/4}}{eq}, \end{aligned}$$

by using

$$D \leq X^{1-\delta-\eta}/q^2.$$

□

## 5. Proof of Theorem 3.1

In this section, we prove Theorem 3.1 by following a similar strategy as in [FI10, Chapter 24.2].

We first note that by restricting to  $b \in B_1$  (recall (3.1)) we may assume that for all  $b$  we have



$$\sum_{\substack{a^2+b^2 \sim X \\ (a,b)=1 \\ (a^2+b^2,2)=1}} \chi_1((b+ia)) \geq 0. \quad (5.1)$$

We define the Dirichlet convolution on ideals of the Gaussian integers as

$$(f * g)(\mathfrak{a}) := \sum_{\mathfrak{c}\mathfrak{d}=\mathfrak{a}} f(\mathfrak{c})g(\mathfrak{d}).$$

We define the auxiliary function

$$\lambda_1(\mathfrak{a}) := (1 * \chi_1)(\mathfrak{a}),$$

which assuming the existence of a Siegel zero for  $L(1, \chi_1)$  is sparsely supported, precisely, for square-free  $\mathfrak{a}$  it is supported on  $\mathfrak{a}$  such that for all  $\mathfrak{p}|\mathfrak{a}$  we have  $\chi_1(\mathfrak{p}) = +1$ .

We set  $\mathcal{A}' = (a'_n)$  with  $a'_n = a_n \lambda_1(\mathfrak{n}) \mathbf{1}_{(n, u_1)=1}$  and define the multiplicative functions

$$g_1(\mathfrak{d}) := \prod_{\mathfrak{p}|\mathfrak{d}} g_1(\mathfrak{p}), \quad g_1(\mathfrak{p}) := \frac{\mathbf{1}_{\mathfrak{p} \neq \bar{\mathfrak{p}}}}{\mathcal{N}\mathfrak{p}} \left( 1 + \chi_1(\mathfrak{p}) - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}} \right) \left( 1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}} \right) \left( 1 - \frac{1}{\mathcal{N}\mathfrak{p}^2} \right)^{-1},$$

$$g'(d) := \sum_{\mathcal{N}\mathfrak{d}=d} g_1(\mathfrak{d}), \quad g'(p) = \frac{\rho(p)}{p} \left( 1 + \chi_1(\mathfrak{p}) - \frac{\chi_1(\mathfrak{p})}{p} \right) \left( 1 - \frac{\chi_1(\mathfrak{p})}{p} \right) \left( 1 - \frac{1}{p^2} \right)^{-1},$$

denoting  $p = \mathfrak{p}\bar{\mathfrak{p}}$ , which is well defined since  $\rho(p) \neq 0$  precisely if  $p$  splits in  $\mathbb{Z}[i]$  and  $\chi_1(\mathfrak{p}) = \chi_1(\bar{\mathfrak{p}})$  since  $\chi_1$  is real. We also set

$$V'(Z) := \prod_{2 < p < Z} (1 - g'(p)),$$

$$\mathbb{Y} := L(1, \chi_1) \sum_{\substack{a^2+b^2 \sim X \\ b \in B \\ (a,b)=1 \\ (b+ia, 2u_1)=1}} (1 + \chi_1((b+ia)) \prod_{\mathfrak{p}|b} \left( 1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}} \right) \prod_{\substack{p|b \\ p \neq 2}} (1 - g'(p))^{-1} \left( 1 - \frac{1}{p^2} \right).$$

Note that by (5.1) the contribution from the terms with  $\chi_1((b+ia))$  is positive and, therefore, we may drop it to conclude

$$\mathbb{Y} \gg L(1, \chi_1) \sum_{\substack{a^2+b^2 \sim X \\ b \in B \\ (a,b)=1 \\ (b+ia, 2u_1)=1}} \prod_{\mathfrak{p}|b} \left( 1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}} \right) \prod_{\substack{p|b \\ p \neq 2}} (1 - g'(p))^{-1} \left( 1 - \frac{1}{p^2} \right).$$

We define

$$S(\mathcal{A}', Z) := \sum_{\substack{\mathcal{N}\mathfrak{n} \sim X \\ (\mathcal{N}\mathfrak{n}, P(Z))=1}} a'_n.$$

We note for large primes  $p$  the function  $g'(p)$  is essentially  $p^{-1}\rho(p)(1 + \chi_1(\mathfrak{p}))$ .

Theorem 3.1 is then a direct corollary of the following and the lower bound  $L(1, \chi_1) \gg_\varepsilon |u_1|^{-\varepsilon}$ . Note that from the assumption  $\Omega(B_1) \geq \Omega(B)/2$  it follows that  $|B_1| \gg_\varepsilon |B|X^{-\varepsilon}$  since  $\omega(b) = X^{\pm o(1)}$ .

**PROPOSITION 5.1.** *Suppose that the assumptions of Theorem 3.1 hold. Then for  $Z = |u_1|^4$  we have*

$$S(\mathcal{A}', 2X^{1/2}) = V'(Z) \mathbb{Y} \left( 1 + O \left( \frac{\log Z}{\log X} + \delta(Z, X) \right) \right),$$

where

$$\delta(Z, X) = \sum_{Z \leq \mathcal{N}\mathfrak{p} < 2X^{1/2}} \frac{\lambda_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}} < (1 - \beta)(\log X + O(\log Z)). \quad (5.2)$$

*Remark 5.2.* Note that by introducing the weight  $\lambda_1(\mathfrak{n})$  we have already removed all prime factors of  $\mathfrak{n}$  with  $\chi_1(\mathfrak{p}) = -1$ , which by the Siegel zero assumption is most of the primes if  $\varepsilon_1$  is small. Thus, we are in a situation of a low-dimensional sieve as we only need to shift out prime divisors with  $\chi_1(\mathfrak{p}) = +1$ .

We first gather Type I information. Note that  $(a, b) = 1$  implies that  $\mathfrak{d}$  is not divisible by a rational prime.

**LEMMA 5.3.** Denote  $q = \mathcal{M}(u_1)$ . Let  $\alpha_{\mathfrak{d}}$  be divisor-bounded coefficients supported on square-free  $\mathfrak{d}$  with  $(\mathfrak{d}, \bar{\mathfrak{d}}) = 1$ . Let  $g_1^{(b)}(\mathfrak{d}), \mathbb{Y}^{(b)}$  be defined similarly as in § 4. Then

$$\sum_{\mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4} \alpha_{\mathfrak{d}} \left( \sum_{\substack{\mathcal{N}\mathfrak{n} \sim X \\ \mathfrak{n} \equiv 0 \pmod{\mathfrak{d}}}} a'_{\mathfrak{n}} - \sum_{b \in B} g_1^{(b)}(\mathfrak{d}) \mathbb{Y}^{(b)} \right) \ll X^{1-\delta-\eta/100}.$$

*Proof.* The function  $\lambda_1(\mathfrak{n})$  is multiplicative and we have similarly to [F10, (24.5)] (denoting  $\mathfrak{n} = \mathfrak{d}\mathfrak{m}$ )

$$\lambda_1(\mathfrak{d}\mathfrak{m}) = \sum_{\mathfrak{c} | (\mathfrak{d}, \mathfrak{m})} \mu(\mathfrak{c}) \chi_1(\mathfrak{c}) \lambda_1(\mathfrak{d}/\mathfrak{c}) \lambda_1(\mathfrak{m}/\mathfrak{c}).$$

We have

$$\lambda_1(\mathfrak{m}/\mathfrak{c}) = \sum_{\mathfrak{a}\mathfrak{b} = \mathfrak{m}/\mathfrak{c}} \chi_1(\mathfrak{a}) = (1 + \chi_1(\mathfrak{m}/\mathfrak{c})) \sum_{\substack{\mathfrak{a}\mathfrak{b} = \mathfrak{m}/\mathfrak{c} \\ \mathcal{N}\mathfrak{a} < \mathcal{N}\mathfrak{b}}} \chi_1(\mathfrak{a}) + O(\mathbf{1}_{\mathcal{N}\mathfrak{m}/\mathfrak{c} = \square}).$$

The contribution from the error term where  $\mathcal{N}\mathfrak{m}/\mathfrak{c} = \square$  may be bounded by crude bounds. We have

$$\chi_1(\mathfrak{m}/\mathfrak{c}) = \chi_1(\mathfrak{c}\mathfrak{d}) \chi_1(\mathfrak{n}).$$

The contribution from  $\mathcal{N}\mathfrak{c} > X^\eta$  can be bounded by crude bounds. We obtain

$$\begin{aligned} \sum_{\mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4} \alpha_{\mathfrak{d}} \sum_{\substack{\mathcal{N}\mathfrak{n} \sim X \\ \mathfrak{n} \equiv 0 \pmod{\mathfrak{d}}}} a'_{\mathfrak{n}} + O(X^{1-\delta-\eta/100}) &= \sum_{\mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4} \alpha_{\mathfrak{d}} \sum_{\substack{\mathfrak{c} | \mathfrak{d} \\ \mathcal{N}\mathfrak{c} \leq X^\eta}} \mu(\mathfrak{c}) \chi_1(\mathfrak{c}) \sum_{\mathcal{N}\mathfrak{c}\mathfrak{d}\mathfrak{a}^2 \leq 2X} \chi_1(\mathfrak{a}) \\ &\quad \times \sum_{\substack{\mathcal{N}\mathfrak{n} \sim X \\ \mathcal{N}\mathfrak{n} > \mathcal{N}\mathfrak{c}\mathfrak{d}\mathfrak{a}^2 \\ \mathfrak{n} \equiv 0 \pmod{\mathfrak{c}\mathfrak{d}\mathfrak{a}}}} a_{\mathfrak{n}} \mathbf{1}_{(\mathfrak{n}, u_1) = 1} (1 + \chi_1(\mathfrak{c}\mathfrak{d}) \chi_1(\mathfrak{n})). \end{aligned}$$

We now relax the cross-conditions  $\mathcal{N}\mathfrak{n} > \mathcal{N}\mathfrak{c}\mathfrak{d}\mathfrak{a}^2$  and  $\mathcal{N}\mathfrak{n} \sim X$  by introducing a finer-than-dyadic decomposition for  $\mathcal{N}\mathfrak{n}$  (using § 2.1 with  $\nu = X^{-\eta/1000}$ ) to get for  $X' \sim X$  sums of the type

$$\sum_{\mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4} \alpha_{\mathfrak{d}} \sum_{\substack{\mathfrak{c}|\mathfrak{d} \\ \mathcal{N}\mathfrak{c} \leq X^\eta}} \mu(\mathfrak{c}) \chi_1(\mathfrak{c}) \sum_{\mathcal{N}\mathfrak{c}\mathfrak{d}\mathfrak{a}^2 \leq X'} \chi_1(\mathfrak{a}) \\ \times \sum_{\substack{\mathfrak{n} \equiv 0 \pmod{\mathfrak{c}\mathfrak{d}\mathfrak{a}}}} a_{\mathfrak{n}} F_{X'}(\mathcal{N}\mathfrak{n}) \mathbf{1}_{(\mathfrak{n}, u_1)=1} (1 + \chi_1(\mathfrak{c}\mathfrak{d}) \chi_1(\mathfrak{n})) =: S_1 + S_2,$$

where

$$S_1 := \sum_{\substack{\mathcal{N}\mathfrak{f} \leq X^{1-\delta-\eta}/q^2 \\ (\mathfrak{f}, \bar{\mathfrak{f}})=1}} \alpha_1(\mathfrak{f}) \sum_{\mathfrak{n} \equiv 0 \pmod{\mathfrak{f}}} a_{\mathfrak{n}} F_{X'}(\mathcal{N}\mathfrak{n}) \mathbf{1}_{(\mathfrak{n}, u_1)=1}, \\ S_2 := \sum_{\substack{\mathcal{N}\mathfrak{f} \leq X^{1-\delta-\eta}/q^2 \\ (\mathfrak{f}, \bar{\mathfrak{f}})=1}} \alpha_2(\mathfrak{f}) \sum_{\mathfrak{n} \equiv 0 \pmod{\mathfrak{f}}} a_{\mathfrak{n}} F_{X'}(\mathcal{N}\mathfrak{n}) \chi_1(\mathfrak{n}),$$

with

$$\alpha_1(\mathfrak{f}) := \sum_{\substack{\mathfrak{f} = \mathfrak{c}\mathfrak{d}\mathfrak{a} \\ \mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4 \\ \substack{\mathfrak{c}|\mathfrak{d} \\ \mathcal{N}\mathfrak{c} \leq X^\eta \\ \mathcal{N}\mathfrak{c}\mathfrak{d}\mathfrak{a}^2 \leq X'}}} \alpha_{\mathfrak{d}} \lambda_1(\mathfrak{d}/\mathfrak{c}) \mu(\mathfrak{c}) \chi_1(\mathfrak{c}) \chi_1(\mathfrak{a}), \\ \alpha_2(\mathfrak{f}) := \sum_{\substack{\mathfrak{f} = \mathfrak{c}\mathfrak{d}\mathfrak{a} \\ \mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4 \\ \substack{\mathfrak{c}|\mathfrak{d} \\ \mathcal{N}\mathfrak{c} \leq X^\eta \\ \mathcal{N}\mathfrak{c}\mathfrak{d}\mathfrak{a}^2 \leq X'}}} \alpha_{\mathfrak{d}} \chi_1(\mathfrak{d}) \lambda_1(\mathfrak{d}/\mathfrak{c}) \mu(\mathfrak{c}) \chi_1(\mathfrak{a}).$$

Note that  $\mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4$ ,  $\mathcal{N}\mathfrak{c} \leq X^\eta$ , and  $\mathcal{N}\mathfrak{c}\mathfrak{d}\mathfrak{a}^2 \leq X'$  imply that  $\mathcal{N}\mathfrak{c}\mathfrak{d}\mathfrak{a} \leq X^{1-\delta-\eta}/q^2$ . If  $\mathfrak{n} = (z)$ , then in  $S_2$  the character value

$$\chi_1(\mathfrak{n}) = \chi_1((z))$$

depends only on the residue class of  $z$  modulo  $4(1+i)u_1$  (note that  $(a, b) = 1$  implies  $2 \nmid z$ ). Thus, by Proposition 4.1, we get

$$S_1 = \sum_{b \in B} \sum_{\substack{\mathcal{N}\mathfrak{f} \leq X^{1-\delta-\eta}/q^2 \\ (\mathfrak{f}, u_1 \bar{\mathfrak{f}})=1}} \alpha_1(\mathfrak{f}) g^{(b)}(\mathfrak{f}) \mathbb{X}_1^{(b)} + O(X^{1/2-\eta/100}|B|), \\ S_2 = \sum_{b \in B} \sum_{\substack{\mathcal{N}\mathfrak{f} \leq X^{1-\delta-\eta}/q^2 \\ (\mathfrak{f}, u_1 \bar{\mathfrak{f}})=1}} \alpha_2(\mathfrak{f}) g^{(b)}(\mathfrak{f}) \mathbb{X}_2^{(b)} + O(X^{1/2-\eta/100}|B|),$$

with

$$\mathbb{X}_1^{(b)} := \sum_{(a,b)=1} F_{X'}(a^2 + b^2) \mathbf{1}_{(b+ia, 2u_1)=1}, \\ \mathbb{X}_2^{(b)} := \sum_{(a,b)=1} F_{X'}(a^2 + b^2) \chi_1((b+ia)) \mathbf{1}_{(b+ia, 2)=1}.$$

Note that we have picked up the condition  $(\mathfrak{f}, \bar{\mathfrak{f}})$  from  $(a, b) = 1$  implicit in  $a_n$ . We have

$$\begin{aligned} \sum_{\substack{\mathcal{N}\mathfrak{f} \leq X^{1-\delta-2\eta}/q^2 \\ (\mathfrak{f}, u_1 \bar{\mathfrak{f}}) = 1}} \alpha_1(\mathfrak{f}) g^{(b)}(\mathfrak{f}) &= \sum_{\mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4} \alpha_{\mathfrak{d}} \sum_{\substack{\mathfrak{c}|\mathfrak{d} \\ \mathcal{N}\mathfrak{c} \leq X^\eta}} \lambda_1(\mathfrak{d}/\mathfrak{c}) \mu(\mathfrak{c}) \chi_1(\mathfrak{c}) \\ &\times \sum_{\substack{\mathcal{N}\mathfrak{a} \leq (X'/\mathcal{N}\mathfrak{c}\mathfrak{d})^{1/2} \\ (\mathfrak{a}, \mathfrak{a}\mathfrak{d}) = 1}} \chi_1(\mathfrak{a}) g^{(b)}(\mathfrak{a}) \end{aligned}$$

and

$$\begin{aligned} \sum_{\substack{\mathcal{N}\mathfrak{f} \leq X^{1-\delta-2\eta}/q^2 \\ (\mathfrak{f}, u_1 \bar{\mathfrak{f}}) = 1}} \alpha_2(\mathfrak{f}) g^{(b)}(\mathfrak{f}) &= \sum_{\mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4} \alpha_{\mathfrak{d}} \chi_1(\mathfrak{d}) \sum_{\substack{\mathfrak{c}|\mathfrak{d} \\ \mathcal{N}\mathfrak{c} \leq X^\eta}} \lambda_1(\mathfrak{d}/\mathfrak{c}) \mu(\mathfrak{c}) \\ &\times \sum_{\substack{\mathcal{N}\mathfrak{a} \leq (X'/\mathcal{N}\mathfrak{c}\mathfrak{d})^{1/2} \\ (\mathfrak{a}, \mathfrak{a}\mathfrak{d}) = 1}} \chi_1(\mathfrak{a}) g^{(b)}(\mathfrak{a}). \end{aligned}$$

To evaluate the sum over  $\mathfrak{a}$ , we have (by applying a Möbius expansion to  $(\mathfrak{a}, \bar{\mathfrak{a}}) = 1$ )

$$\begin{aligned} &\sum_{\substack{\mathcal{N}\mathfrak{a} \leq (X'/\mathcal{N}\mathfrak{c}\mathfrak{d})^{1/2} \\ (\mathfrak{a}, \mathfrak{a}\mathfrak{d}) = 1}} \chi_1(\mathfrak{a}) g^{(b)}(\mathfrak{a}) \\ &= \sum_{\substack{\mathcal{N}\mathfrak{a} \leq (X'/\mathcal{N}\mathfrak{c}\mathfrak{d})^{1/2} \\ (\mathfrak{a}, \mathfrak{a}\mathfrak{d}) = 1}} \chi_1(\mathfrak{a}) \frac{1}{\mathcal{N}\mathfrak{a}} = \sum_{(k, \mathfrak{d}\mathfrak{b}) = 1} \frac{\mu(k)}{k^2} \sum_{\substack{k^2 \mathcal{N}\mathfrak{a} \leq (X'/\mathcal{N}\mathfrak{c}\mathfrak{d})^{1/2} \\ (\mathfrak{a}, \mathfrak{d}\mathfrak{b}) = 1}} \chi_1(\mathfrak{a}) \frac{1}{\mathcal{N}\mathfrak{a}} \\ &= \prod_{(p, \mathfrak{d}\mathfrak{b}) = 1} \left(1 - \frac{1}{p^2}\right) \prod_{(p, \mathfrak{d}\mathfrak{b}) = 1} \left(1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right)^{-1} + O(X^{-\eta}) \\ &= \frac{L(1, \chi_1)}{\zeta(2)} \prod_{p|b} \left(1 - \frac{1}{p^2}\right)^{-1} \prod_{p|b} \left(1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right) \prod_{p|\mathfrak{d}} \left(1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right) \left(1 - \frac{1}{\mathcal{N}\mathfrak{p}^2}\right)^{-1} + O(X^{-\eta}), \end{aligned}$$

since  $\chi_1(\bar{\mathfrak{p}}) = \chi_1(\mathfrak{p})$  and  $(\mathfrak{d}, \bar{\mathfrak{d}}) = 1$ . Let us denote (noting that the contribution from  $\mathcal{N}\mathfrak{c} > X^\eta$  may now be added back in at a negligible cost)

$$\begin{aligned} g_1(\mathfrak{d}) &:= \frac{1}{\mathcal{N}\mathfrak{d}} \prod_{p|\mathfrak{d}} \left(1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right) \left(1 - \frac{1}{\mathcal{N}\mathfrak{p}^2}\right)^{-1} \sum_{\mathfrak{c}|\mathfrak{d}} \lambda_1(\mathfrak{d}/\mathfrak{c}) \mu(\mathfrak{c}) \chi_1(\mathfrak{c}) \frac{1}{\mathcal{N}\mathfrak{c}}, \\ g_2(\mathfrak{d}) &:= \frac{\chi_1(\mathfrak{d})}{\mathcal{N}\mathfrak{d}} \prod_{p|\mathfrak{d}} \left(1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right) \left(1 - \frac{1}{\mathcal{N}\mathfrak{p}^2}\right)^{-1} \sum_{\mathfrak{c}|\mathfrak{d}} \lambda_1(\mathfrak{d}/\mathfrak{c}) \mu(\mathfrak{c}) \frac{1}{\mathcal{N}\mathfrak{c}}. \end{aligned}$$

The functions  $g_j$  are multiplicative with

$$\begin{aligned} g_1(\mathfrak{p}) &= \frac{1}{\mathcal{N}\mathfrak{p}} \left(1 + \chi_1(\mathfrak{p}) - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right) \left(1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right) \left(1 - \frac{1}{\mathcal{N}\mathfrak{p}^2}\right)^{-1}, \\ g_2(\mathfrak{p}) &= \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}} \left(1 + \chi_1(\mathfrak{p}) - \frac{1}{\mathcal{N}\mathfrak{p}}\right) \left(1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right) \left(1 - \frac{1}{\mathcal{N}\mathfrak{p}^2}\right)^{-1} = g_1(\mathfrak{p}), \end{aligned}$$

where the last equality holds since  $\chi_1(\mathfrak{p})^2 = 1$ . Hence, we have

$$S_1 + S_2 = \sum_{b \in B} \sum_{\mathcal{N}\mathfrak{d} \leq X^{1-2\delta-4\eta}/q^4} \alpha_{\mathfrak{d}} g_1^{(b)}(\mathfrak{d}) \mathbb{Y}^{(b)} + O(X^{-\eta})$$

with

$$\begin{aligned} \mathbb{Y}^{(b)} &= \frac{L(1, \chi_1)}{\zeta(2)} \sum_{\substack{a \\ (a,b)=1 \\ (b+ia, 2u_1)=1}} F_{X'}(a^2 + b^2)(1 + \chi_1((b + ia))) \prod_{\substack{p|b \\ p \neq 2}} \left(1 - \frac{1}{p^2}\right)^{-1} \\ &\quad \times \prod_{\substack{\mathfrak{p}|b \\ p \neq 2}} \left(1 - \frac{\chi_1(\mathfrak{p})}{\mathcal{N}\mathfrak{p}}\right). \end{aligned} \quad \square$$

### 5.1 Proof of Proposition 5.1

We apply a sieve argument to the sequence  $\mathcal{A}'' = (a_n'')$  over integers defined by

$$a_n'' := \sum_{\mathcal{N}\mathfrak{n}=n} a_{\mathfrak{n}}'.$$

Note that then

$$S(\mathcal{A}', 2X^{1/2}) = S(\mathcal{A}'', 2X^{1/2}).$$

The proof is essentially the same as in [FI10, Proof of Proposition 24.1], but we give it for completeness as it is short. Let  $D := X^{1-2\delta-4\eta}/q^4$  and denote  $s = \log D / \log Z$ . By Buchstab's identity we have

$$S(\mathcal{A}'', 2X^{1/2}) = S(\mathcal{A}'', Z) - \sum_{Z \leq p < 2X^{1/2}} S(\mathcal{A}_p'', p).$$

By the fundamental lemma of the sieve (Lemma 2.5, see also Remark 2.6) and Lemma 5.3 we have

$$S(\mathcal{A}'', Z) = V'(Z)\mathbb{Y}(1 + O(e^{-s})) + O(X^{1-\delta-\eta/100}),$$

where

$$e^{-s} \ll s^{-1} \ll \frac{\log Z}{\log X}.$$

By an upper bound sieve (e.g. using Lemma 2.5) with level  $D/p > X^{1/4}$  and Lemma 5.3 we have

$$\sum_{Z \leq p < 2X^{1/2}} S(\mathcal{A}_p'', p) \leq \sum_{Z \leq \mathcal{N}p < 2X^{1/2}} S(\mathcal{A}_p'', Z) \ll \delta(Z, X)V'(Z)\mathbb{Y} + O(X^{1-\delta-\eta/100}).$$

Combining the two estimates we get Proposition 5.1, noting that the bound (5.2) can be proved by a similar argument as in [FI10, (24.20)].

*Remark 5.4.* Instead of taking a small  $Z$  and using the fundamental lemma of the sieve, we can take a larger  $Z$  and use the linear sieve lower and upper bounds [FI10, Theorem 11.12]. Taking  $Z = X^{1/4}$  we get that the lower bound for  $S(\mathcal{A}'', 2X^{1/2})$  is proportional to (denoting the linear sieve functions by  $f, F$ )

$$4f(4(1 - O(\delta))) - \varepsilon_1 4 \int_{1/4}^{1/2} \frac{F(4(1 - \alpha - O(\delta)))}{\alpha} d\alpha + O(\delta) = (1 - \varepsilon_1)2 \log(3) + O(\delta).$$

To make the right-hand side positive we can certainly take any  $\varepsilon_1 < 1$  and  $\delta = (1 - \varepsilon_1)/100$ , for instance.

## 6. Type II information: preliminaries

We now fix small parameters  $\eta, \eta_1, \eta_2, \eta_3 > 0$  such that  $\eta_1$  is small compared with  $\eta$ ,  $\eta_2$  is small compared with  $\eta_1$ , and  $\eta_3$  is small compared with  $\eta_2$ , that is,

$$\eta_3 \ll \eta_2 \ll \eta_1 \ll \eta.$$

For instance, for our purposes it would suffice to take any small  $\eta > 0$  and let

$$\eta_1 = \eta/100, \quad \eta_2 = \eta_1/2, \quad \eta_3 = \eta_2/100.$$

We denote

$$\nu_j = X^{-\eta_j}, \quad j \in \{1, 2, 3\}.$$

We often refer to these parameters in the course of the following sections.

Let  $G: \mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{C}$  be a non-negative  $C^\infty$ -smooth function supported on  $[-\nu_1, \nu_1]$  as in Lemma 2.10, satisfying

$$|G^{(j)}| \ll_j \nu_1^{-j}, \quad j \geq 0 \quad \text{and} \quad \int G(\theta) d\theta = \nu_1.$$

and recall that

$$G(\arg z) = \sum_k \check{G}(k) \xi_k(z).$$

For a set of Hecke characters  $\Psi = \{\xi\chi\}$  and  $u \in \mathbb{Z}[i] \setminus \{0\}$  we let  $\Psi_u$  denote the set of characters induced to modulus  $u$ . We say that  $\xi\chi$  is primitive if  $\chi$  is a primitive Dirichlet character. The following definition depends on the choice of  $\eta, \eta_j$  but since these are fixed throughout the argument we omit this dependency in the notation.

**DEFINITION 1** ( $Q$ -regularity). Let  $C_1, C_2 > 0$  and  $Q, N, X \geq 1$ . Let  $\beta_{\mathfrak{n}}$  be complex coefficients. We say that  $\beta_{\mathfrak{n}}$  is  $(Q, N, X, C_1, C_2)$ -regular if there exists a set of primitive characters  $\Psi = \{\xi\chi\}$  with  $|\Psi| \leq (\log X)^{C_2}$  such that the following holds. For any  $\xi\chi \in \Psi$ , we have

$$\mathcal{M}(\text{cond}(\chi)) \leq Q \quad \text{and} \quad \xi = \xi_k, \quad |k| \leq \nu_1^{-2}$$

and for any  $u \in \mathbb{Z}[i]$  with  $\mathcal{M}(u) \leq Q$  and  $N' \sim N$  we have

$$\frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_{\mathcal{N}_{\mathfrak{n}} \in [N', N'(1+\nu_2)]} \beta_{\mathfrak{n}} \xi_k \psi(\mathfrak{n}) \right) \right|^2 \leq \frac{\nu_1^2 \nu_2^2 N^2}{\varphi_{\mathbb{Z}[i]}(d)(\log X)^{C_1}}.$$

Given a function  $Q = Q(N, X)$ , we say that  $\beta_{\mathfrak{n}}$  is  $Q$ -regular if for any  $C_1 > 0$  there is some  $C_2 > 0$  and some  $X_0 > 0$  such that for all  $X \geq X_0$  and for all  $N \geq X^\eta$  the coefficient  $\beta_{\mathfrak{n}}$  is  $(Q, N, X, C_1, C_2)$ -regular.

Informally speaking, coefficient  $\beta_{\mathfrak{n}}$  is  $Q$ -regular if it is equidistributed in residue classes and polar boxes apart from a set of Hecke characters of size  $\leq (\log X)^{O(1)}$ , uniformly in the size of the modulus of the  $\mathcal{M}(u) \leq Q$ . We need the fact that the Möbius function restricted to rough numbers is  $Q$ -regular for  $Q$  close to  $N^{1/3}$ , which we prove in § 6.4 using the zero density estimate (Lemma 2.14).

**LEMMA 6.1.** Let  $W := X^{1/(\log \log X)^2}$ . The coefficient

$$\beta_{\mathfrak{n}} := \mu(\mathcal{N}_{\mathfrak{n}}) \mathbf{1}_{(\mathfrak{n}, P(W))=1}$$

is  $Q$ -regular for any  $Q = Q(N, X)$  with  $N \geq X^\eta Q^3$ . Furthermore, for fixed  $Q, X, C_1 > 0$ , a fixed set of characters  $\Psi$  with  $C_2 \ll_{C_1} 1$  works for all ranges of  $N > X^\eta Q^3$  in the definition of  $(Q, N, X, C_1, C_2)$ -regularity.

The exponent 3 in  $N > X^\eta Q^3$  is not the best that can be obtained but it suffices for our purposes. This could be improved by using results on large values of Dirichlet polynomials (analogous to [Hux74]). We have Type II information given by the following proposition. We will apply it with  $\beta_n$  as above but it applies equally well to, e.g., products of  $k$  primes.

**PROPOSITION 6.2** (Type II information). *Let  $W := X^{1/(\log \log X)^2}$  and let  $\nu = (\log X)^{-C}$  for some  $C > 0$ . For every  $C_1 > 0$ , there is some  $C_2 \ll_{C_1} 1$  such that the following holds. Let  $MN = X' \sim X$  with*

$$X^{3\delta+4\eta} < N < X^{1/2-\delta-\eta}.$$

Let  $\alpha_m, \beta_n$  be bounded coefficients, supported on  $(mn, P(W)) = 1$  and

$$\mathcal{N}m \in [M, M(1+\nu)], \quad \mathcal{N}n \in [N, N(1+\nu)].$$

Let  $F$  be a smooth function as in § 2.1 with the parameter  $\nu$ . Suppose that for  $Q = X^{\delta+\eta}$  the coefficient  $\beta_n$  is  $Q$ -regular and let  $\xi_{k_j} \chi_j$  denote the corresponding Hecke characters with  $j \leq J \leq (\log X)^{C_2}$ . Then

$$\sum_{m,n} \alpha_m \beta_n a_{mn} = \sum_{j \leq J} \sum_{m,n} \frac{\alpha_m \beta_n \xi_{k_j} \chi_j(mn)}{\mathcal{N}mn} \sum_{\mathfrak{a}} \frac{F(\mathcal{N}mn/\mathcal{N}\mathfrak{a})}{\widehat{F}(0)} a_{\mathfrak{a}}^{\omega} \overline{\xi_{k_j} \chi_j(\mathfrak{a})} + O\left(\frac{X^{1/2}|B|}{(\log X)^{C_1}}\right).$$

*Remark 6.3.* Note that on the right-hand side of Proposition 6.2 we have the sequence  $a_{\mathfrak{a}}^{\omega}$  which is twisted by  $\omega$ , which arises from the assumption  $(mn, P(W)) = 1$ . In particular, Proposition 6.2 as stated would be false without restricting to  $(mn, P(W)) = 1$ .

We set

$$\alpha_w := \alpha(\overline{w}) \quad \text{and} \quad \beta_z := \beta_{(z)} \mathbf{1}_{z \text{ primary}},$$

so that in the latter we may swap freely between Gaussian integers  $z$  and ideals  $\mathfrak{n}$  and that

$$\sum_{m,n} \alpha_m \beta_n a_{mn} = \sum_{w,z} \alpha_w \beta_z a_{\overline{w}z}.$$

We have included the complex conjugate in  $\overline{w}$  to make our notation match with those of [FI98b].

## 6.1 Sketch of the argument

As the proof of Proposition 6.2 is quite technical, we include here a simplified non-rigorous sketch. In § 6.2 we construct an approximation  $\beta_n^{\#}$  for  $\beta$  such that the difference

$$\beta_n^b := \beta_n - \beta_n^{\#}$$

is balanced along arithmetic progressions. Write

$$S(\alpha, \beta) = S(\alpha, \beta^{\#}) + S(\alpha, \beta^b).$$

The approximation will be simple enough that  $S(\alpha, \beta^{\#})$  can be evaluated by Type I information, so let us consider

$$S(\alpha, \beta^b) = \sum_w \sum_z \alpha_w \beta_z^b \mathbf{1}_B(\operatorname{Re}(\overline{w}z)),$$

where the aim is to capture the oscillations from  $\beta_z^b$ . By applying Cauchy–Schwarz we get

$$S(\alpha, \beta^b) \ll M^{1/2} U^{1/2},$$

with

$$U := \sum_{z_1, z_2} \beta_{z_1}^b \overline{\beta_{z_2}^b} \sum_w \mathbf{1}_B(\operatorname{Re}(\overline{w}z_1)) \mathbf{1}_B(\operatorname{Re}(\overline{w}z_2)) F_M(|w|^2),$$

for some smooth majorant  $F_M$  of the interval  $[M, 2M]$ . The goal is to evaluate the sum over  $w$  with a main term  $M(z_1, z_2)$  and then show that

$$\sum_{z_1, z_2} \beta_{z_1}^b \overline{\beta_{z_2}^b} M(z_1, z_2) \ll \frac{N|B|^2}{(\log X)^C}$$

is small due to cancellations from the coefficients  $\beta_z^b$ .

Similarly as in [FI98b], we note that by denoting  $b_j = \operatorname{Re}(\overline{w}z_j)$  and

$$\Delta = \operatorname{Im}(\overline{z_1}z_2), \quad a \equiv z_2/z_1 \pmod{\Delta},$$

we have

$$i\Delta w = z_2 b_1 - z_1 b_2. \quad (6.1)$$

Note that typically  $|\Delta| \approx N$ . The parts where  $\Delta = 0$  or  $|\Delta| < N/(\log X)^C$  correspond to diagonal contributions and may be bounded by crude estimates. Thus, we assume for simplicity that  $|\Delta| \asymp N$ .

Let  $b_0 := (b_1, b_2)$  and write  $b_j = b_0 b'_j$ . Note that in the situation that  $B \subseteq q_1 \mathbb{Z}$  we have  $q_1 | b_0$ , so that  $b_0$  can be quite large for a large subset of  $B \times B$ . As usual, in most places dealing with greatest common divisors does not cause serious problems but the dependency on  $b_0$  will be crucial for our argument.

We have  $b_0 | \Delta w$  by (6.1) and, for simplicity, let us assume that  $b_0 | \Delta$ . Then  $w = (z_2 b_1 - z_1 b_2)/(i\Delta)$  is fixed once we fix  $z_j, b_j$ , so that (ignoring the smooth weight  $F_M$ ) we have to bound

$$V = \sum_{b_0} \sum_{\substack{z_1, z_2 \\ |\Delta| \asymp N \\ b_0 | \Delta}} \beta_{z_1}^b \overline{\beta_{z_2}^b} \sum_{b'_2 \equiv ab'_1 \pmod{\Delta/b_0}} \mathbf{1}_B(b_0 b'_1) \mathbf{1}_B(b_0 b'_2).$$

Note that

$$a \equiv z_2/z_1 \equiv \frac{\operatorname{Re}(\overline{z_1}z_2)}{|z_1|^2} \pmod{\Delta}$$

is congruent to an integer, so that the congruence  $b'_2 \equiv ab'_1 \pmod{\Delta/b_0}$  lives in  $\mathbb{Z}/(\Delta/b_0)\mathbb{Z}$ .

By expansion with Dirichlet characters and sorting into primitive characters we get (ignoring issues with greatest common divisors)

$$V = \sum_{b_0} \sum_{\substack{z_1, z_2 \\ |\Delta| \asymp N \\ b_0 | \Delta}} \beta_{z_1}^b \overline{\beta_{z_2}^b} \frac{1}{\varphi(\Delta/b_0)} \sum_{d | \Delta/b_0} \sum_{\chi(d)}^* \overline{\chi}(a) \left| \sum_{b_0 b \in B} \chi(b) \right|^2,$$

where, morally,

$$\frac{1}{\varphi(\Delta/b_0)} \approx \frac{b_0}{N}.$$

We split this into two parts,  $db_0 > X^{\delta+\eta}$  and  $db_0 \leq X^{\delta+\eta}$ . The contribution from the small  $d$  is our main term  $M(z_1, z_2)$  referred to in the above.



For the large  $db_0$ , we get

$$\begin{aligned} & \sum_{b_0} b_0 \sum_{db_0 > X^{\delta+\eta}} \sum_{\substack{D \leq N \\ D \equiv 0 \pmod{db_0}}} \left( \frac{1}{N} \sum_{\substack{z_1, z_2 \\ |\Delta|=D}} 1 \right) \sum_{\chi(d)}^* \left| \sum_{b_0 b \in B} \chi(b) \right|^2 \\ & \ll N \sum_{b_0} \sum_{X^{\delta+\eta} < db_0 \leq N} \frac{1}{d} \sum_{\chi(d)}^* \left| \sum_{b_0 b \in B} \chi(b) \right|^2, \end{aligned}$$

and applying the large sieve for multiplicative characters (Lemma 2.11) we get

$$\ll N(N + X^{1/2-\delta-\eta})|B| \ll X^{-\eta} N|B|^2$$

by using  $N \ll X^{-\eta}|B|$ . Note that we are applying the large sieve to a very sparse set  $B/b_0 \cap \mathbb{Z}$ , which causes a loss in the diagonal terms and we are forced to take  $db_0$  at least a bit bigger than  $X^\delta$ .

For the small  $db_0$ , we can rewrite the conditions  $b_0 d | \Delta$ ,  $a \equiv z_2/z_1 \pmod{\Delta}$  as  $z_2 \equiv az_1 \pmod{b_0 d}$  to get

$$\frac{1}{N} \sum_{b_0} b_0 \sum_{db_0 \leq X^{\delta+\eta}} \sum_{a \pmod{b_0 d}} \sum_{\chi(d)}^* \bar{\chi}(a) \left| \sum_{b_0 b \in B} \chi(b) \right|^2 \sum_{z_2 \equiv az_1 \pmod{b_0 d}} \beta_{z_1}^b \overline{\beta_{z_2}^b}. \quad (6.2)$$

We now see what precisely is required of the balanced function  $\beta_z^b$ , we need

$$\sum_{z_2 \equiv az_1 \pmod{b_0 d}} \beta_{z_1}^b \overline{\beta_{z_2}^b} \ll \frac{N^2}{\varphi_{\mathbb{Z}[i]}(b_0 d)(\log X)^C},$$

where the modulus may be as large as  $X^{\delta+\eta}$ . This would follow if there were no zeros of  $L(s, \chi)$  with a real part  $> 1 - C' \log \log X / \log X$ . Since this is not known, we need to construct the approximation  $\beta_z^\#$  in a way that takes into account these possible bad characters. By the zero-density estimate (Lemma 2.14) this means that the approximation needs to see  $\ll (\log X)^{O(1)}$  of the characters. For technical reasons (due to the smooth weight  $F_M$ ) the approximation also needs to see the distribution of  $\beta_z$  with respect to  $\arg z$  and  $|z|^2$ . Note that, in the case, that  $B \subseteq q_1 \mathbb{Z}$  we have always  $q_1 | b_0$ , where  $q_1$  can be as large as  $X^\delta$ .

## 6.2 An approximation for $\beta_n$

For the approximation, it is convenient to use a rough finer-than-dyadic partition of unity instead of § 2.1, so that the different parts do not overlap. Let  $\nu_2 = X^{-\eta_2}$ , and let

$$H_{N'}(\mathbf{n}) := \mathbf{1}_{(N', N'(1+\nu_2)]}(\mathcal{N}\mathbf{n}),$$

so that

$$\mathbf{1}_{(N, 2N]}(\mathcal{N}\mathbf{n}) = \sum_{\substack{N' = N(1+\nu_2)^n \\ n \in [N, 2N]}} H_{N'}(\mathbf{n}).$$

We can, of course, choose  $\nu_2$  so that  $2 = (1 + \nu_2)^k$  for some  $k \asymp \nu_2^{-2}$ .

Let  $\beta_n$  be  $Q$ -regular and let  $\Psi = \{\xi_{k_j} \chi_j\}$  denote the set of  $J \leq (\log X)^{C_2}$  characters, and denote the moduli of the characters by  $u_1, \dots, u_J$  and the primitive characters by  $\chi_1, \dots, \chi_J$ . For any two coefficients  $\alpha, \beta$ , we define their normalized  $W$ -rough correlation as

$$\mathcal{C}_W(\beta, \alpha) := \left( \sum_{\substack{(\mathbf{n}, P(W))=1 \\ (\mathbf{n}, \bar{\mathbf{n}})=1}} \beta_{\mathbf{n}} \overline{\alpha_{\bar{\mathbf{n}}}} \right) \left( \sum_{\substack{(\mathbf{n}, P(W))=1 \\ (\mathbf{n}, \bar{\mathbf{n}})=1}} |\alpha_{\mathbf{n}}| \right)^{-1}$$

if the denominator is non-zero. We then define the approximation  $\beta_{\mathbf{n}}^{\#}$  for  $\beta_{\mathbf{n}}$ ,

$$\beta_{\mathbf{n}}^{\#} = \beta_{\mathbf{n}}^{\#}(\Psi) := \sum_{N'=N(1+\nu_2)^j \in [N, 2N]} H_{N'}(\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{n}P(W))=1} \left( \sum_{j \leq J} \overline{\xi_{k_j} \chi_j}(\mathbf{n}) \mathcal{C}_W(\beta, \overline{\xi_{k_j} \chi_j} H_{N'}) \right),$$

and the balanced function,

$$\beta_{\mathbf{n}}^b := \beta_{\mathbf{n}} - \beta_{\mathbf{n}}^{\#},$$

so that we have a decomposition

$$\beta_{\mathbf{n}} = \beta_{\mathbf{n}}^{\#} + \beta_{\mathbf{n}}^b.$$

Morally, the approximation  $\beta^{\#}$  can be viewed as a kind of ‘expansion’ with respect to a ‘basis’, which is justified since the functions  $\xi_{\chi} H_{N'}$  are approximately orthogonal over  $W$ -rough number, as the following lemma shows. For the lemma, recall that all functions of odd Gaussian integers  $z$  are extended to  $\mathbf{n}$  by considering the primary generator, for instance, we write  $G(\arg \mathbf{n}) = G(\arg z)$  if  $z$  is the primary generator of  $\mathbf{n}$ .

LEMMA 6.4. *Let  $\psi, \chi$  be characters to coprime moduli  $u, u_1$  and let  $\xi = \xi_k$  with  $|k| \ll (\nu_1)^{-2}$ . Let  $N > X^{\eta} |u|^2 |u_1|$ . Then, for any  $C > 0$ , we have*

$$\sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{n}P(W))=1} (\xi \psi \chi(\mathbf{n}) - \mathbf{1}_{\xi \psi \chi=1}) \ll_C \frac{\nu_2 N}{(\log N)^C}$$

and, for any  $(\mathbf{m}, u) = 1$ ,

$$\sum_{(\mathbf{n}, u)=1} H_{N'}(\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{n}P(W))=1} \left( \mathbf{1}_{\mathbf{n} \equiv \mathbf{m} (u)} G(\arg \mathbf{n}) \xi \chi(\mathbf{n}) - \frac{\mathbf{1}_{\chi=1} \check{G}(-k)}{\varphi_{\mathbb{Z}[i]}(u)} \right) \ll_C \frac{\nu_1 \nu_2 N}{\varphi_{\mathbb{Z}[i]}(u) (\log N)^C}.$$

*Proof.* We prove the second claim, the first is similar but easier. By applying § 2.1 with a smooth function  $F$  with the parameter  $\nu_1$  we split  $\mathcal{N}\mathbf{n}$  smoothly into finer-than-dyadic intervals. The contribution from the edges of the support of  $H_{N'}$  gives a negligible contribution by trivial bounds. It then suffices to show that for any  $N_1 \sim N$

$$\sum_{(\mathbf{n}, u)=1} F_{N_1}(\mathcal{N}\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{n}P(W))=1} \left( \mathbf{1}_{\mathbf{n} \equiv \mathbf{m} (u)} G(\arg \mathbf{n}) \xi \chi(\mathbf{n}) - \frac{\mathbf{1}_{\chi=1} \check{G}(-k)}{\varphi_{\mathbb{Z}[i]}(u)} \right) \ll_C \frac{\nu_1^2 N}{\varphi_{\mathbb{Z}[i]}(u) (\log N)^C}.$$

We let  $z$  denote a primary generator of  $\mathbf{n}$ . The condition  $(z, \bar{z}) = 1$  may be dropped with a negligible error term since  $z$  is supported on  $(z, P(W)) = 1$ . We write

$$\mathbf{1}_{(z, P(W))=1} = \sum_{\substack{v|(z, P(W)) \\ |v|^2 \leq N^{\eta_1}}} \mu(v) + \sum_{\substack{v|(z, P(W)) \\ |v|^2 > N^{\eta_1}}} \mu(v)$$

and write

$$\begin{aligned} & \sum_{(\mathbf{n}, u)=1} F_{N_1}(\mathcal{N}\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{n}P(W))=1} \left( \mathbf{1}_{\mathbf{n} \equiv \mathbf{m} (u)} G(\arg \mathbf{n}) \xi \chi(\mathbf{n}) - \frac{\mathbf{1}_{\chi=1} \check{G}(-k)}{\varphi_{\mathbb{Z}[i]}(u)} \right), \\ &= \sum_{z \equiv 1 \pmod{2(1+i)}} F_{N_1}(|z|^2) \mathbf{1}_{(z, P(W))=1} \left( \mathbf{1}_{z \equiv w (u)} G(\arg z) \xi \chi(z) - \frac{\mathbf{1}_{\chi=1} \check{G}(-k)}{\varphi_{\mathbb{Z}[i]}(u)} \right). \\ &= S_{\leq} + S_{>}. \end{aligned}$$

For the large  $v$  we note that by  $v|P(W)$  there is some factor  $v_0|v$  such that  $|v_0|^2 \in (X^{\eta_1}, X^{\eta_1}W]$ . Thus, by  $\check{G}(-k) \ll \nu_1$

$$\begin{aligned} S_{>} &\ll \sum_{\substack{|v_0|^2 \in (X^{\eta_1}, X^{\eta_1}W] \\ v_0|P(W) \\ (v_0, u)=1}} \sum_{z \equiv 0(v_0)} \tau(z)^{O(1)} F_{N_1}(|z|^2) \mathbf{1}_{z \equiv w(u)} G(\arg z) \\ &+ \frac{\nu_1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\substack{|v_0|^2 \in (X^{\eta_1}, X^{\eta_1}W] \\ v_0|P(W) \\ (v_0, u)=1}} \sum_{z \equiv 0(v_0)} \tau(z)^{O(1)} F_{N_1}(|z|^2). \end{aligned}$$

Recall that  $\eta$  is large compared with  $\eta_1$ . Hence, by counting the sum over  $z \equiv 0(v_0)$  (using Lemma 2.4 to handle  $\tau(z)^{O(1)}$ ) and applying Lemma 2.2, we get

$$S_{>} \ll \frac{\nu_1^2 N}{\varphi_{\mathbb{Z}[i]}(u)(\log N)^C}.$$

For small  $v$  we split into two cases depending on  $\chi \neq 1$  and  $\chi = 1$ . For  $\chi \neq 1$ , by writing  $z = uvz' + \alpha$  we have

$$\begin{aligned} S_{\leq} &= \sum_{\substack{v|P(W) \\ |v|^2 \leq N^{\eta_1} \\ (v, uu_1)=1}} \mu(v) \sum_{z \equiv 0(v)} F_{N_1}(|z|^2) \mathbf{1}_{z \equiv w(u)} G(\arg z) \xi \chi(z) \\ &= \sum_{\substack{v|P(W) \\ |v|^2 \leq N^{\eta_1}}} \mu(v) \sum_{z'} F_{N_1}(|uvz' + \alpha|^2) G(\arg uvz + \alpha) \xi(uvz' + \alpha) \chi(uvz' + \alpha) \\ &= \sum_{\substack{v|P(W) \\ |v|^2 \leq N^{\eta_1}}} \mu(v) \chi(uv) \sum_{z'} F_{N_1}(|uvz' + \alpha|^2) G(\arg uvz + \alpha) \xi(uvz' + \alpha) \chi(z' + \alpha(uv)^{-1}), \end{aligned}$$

since  $(uv, u_1) = 1$ . Treating the weight

$$z' \mapsto F_{N_1}(|uvz' + \alpha|^2) G(\arg uvz + \alpha) \xi(uvz' + \alpha)$$

as a smooth weight, we get by the Polyá-Vinogradov bound (Lemma 2.18)

$$S_{\leq} \ll N^{O(\eta_1)} |u_1| \ll N^{-\eta/2} \frac{\nu_1^2 N}{\varphi_{\mathbb{Z}[i]}(u)}$$

by  $N^{1-\eta} > |u|^2 |u_1|$  since  $\eta_1$  is small compared with  $\eta$ .

For  $\chi = 1$ , we have by Lemma 2.10 and  $(\check{G}\xi_k)(\ell) = \check{G}(\ell - k)$

$$\begin{aligned} S_{\leq} &= \sum_{\substack{v|P(W) \\ |v|^2 \leq N^{\eta_1} \\ (v, uu_1)=1}} \mu(v) \sum_{z \equiv 0(v)} F_{N_1}(|z|^2) \left( \mathbf{1}_{z \equiv w(u)} G(\arg z) \xi(z) - \frac{\check{G}(-k)}{\varphi_{\mathbb{Z}[i]}(u)} \right) \\ &= \sum_{\substack{v|P(W) \\ |v|^2 \leq N^{\eta_1} \\ (v, uu_1)=1}} \mu(v) \sum_{\ell \neq 0} \check{G}(\ell - k) \sum_{z \equiv 0(v)} F_{N_1}(|z|^2) \mathbf{1}_{z \equiv w(u)} \xi_{\ell}(z) \\ &+ \sum_{\substack{v|P(W) \\ |v|^2 \leq N^{\eta_1} \\ (v, uu_1)=1}} \mu(v) \check{G}(-k) \sum_{z \equiv 0(v)} F_{N_1}(|z|^2) \left( \mathbf{1}_{z \equiv w(u)} - \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \right). \end{aligned}$$

Estimating the contribution from  $|\ell| > (\nu_1)^{-3}$  trivially (by Lemma 2.10) and for  $|\ell| \leq (\nu_1)^{-3}$  applying the Poisson summation formula on  $\mathbb{Z}[i]$  we get

$$S_{\leq} \ll N^{-\eta/2} \frac{\nu_1^2 N}{\varphi_{\mathbb{Z}[i]}(u)}. \quad \square$$

Our main lemma about the approximation is the following, which says that  $\beta_{\mathfrak{n}}^b$  is balanced over arithmetic progressions restricting to small polar boxes.

LEMMA 6.5. *Let  $N > X^\eta Q^3$ . For any  $C_1 > 0$ , there is some  $C_2 \ll_{C_1} 1$  such that the following holds. Let  $\xi = \xi_k$  with  $k = (\log N)^{O(1)}$ . Let  $\beta_{\mathfrak{n}}$  be  $Q$ -regular, and let  $\beta_{\mathfrak{n}}^b$  be as above with  $|\Psi| = J \leq (\log X)^{C_2}$ . Suppose that  $\beta_{\mathfrak{n}}$  is supported on  $(\mathfrak{n}, P(W)) = 1$ . Let  $N' = N(1 + \nu_2)^j \in [N, 2N]$  and  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ . Then, for any  $u \in \mathbb{Z}[i]$  with  $\mathcal{M}(u) \leq Q$ , we have*

$$S = \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\mathfrak{n}} \beta_{\mathfrak{n}}^b \psi(\mathfrak{n}) H_{N'}(\mathfrak{n}) G(\arg \mathfrak{n} - \theta) \right|^2 \ll \frac{\nu_1^2 \nu_2^2 N^2}{\varphi_{\mathbb{Z}[i]}(u) (\log X)^{C_1}}.$$

*Proof.* Let us first show that in the approximation  $\beta_{\mathfrak{n}}^\#$  we can replace the characters with conductor dividing  $u$  by characters with modulus  $u$ . Suppose that  $\psi$  is induced by a primitive character  $\psi'$  of modulus  $u' < u$ . Then

$$\psi(\mathfrak{n}) = \psi'(\mathfrak{n}) \mathbf{1}_{(z, u/u')} = \psi'(\mathfrak{n}) - \psi'(\mathfrak{n}) \mathbf{1}_{(\mathfrak{n}, u/u') > 1}.$$

Since we have defined the correlation by sums over  $(\mathfrak{n}, P(W)) = 1$ , the characters agree unless  $\mathcal{N}(\mathfrak{n}, (u/u')) > W$ , so that we have

$$\mathcal{C}_W(\beta, \xi_{k_j} \psi' H_{N'}) = \mathcal{C}_W(\beta, \xi_{k_j} \psi H_{N'}) + O(W^{-1} (\log X)^{O(1)}).$$

Thus, if the approximation includes a character whose conductor is a proper divisor of  $u$ , we may replace it by the character with  $(\mathfrak{n}, u) = 1$  at a negligible cost (by using orthogonality of characters). Let us assume that this has been done, so that the moduli of the characters  $\chi_j$  satisfy either  $u_j = u$  or  $(u_j, u) = 1$ . Let  $\Psi_u$  denote the characters  $\xi\psi$  modulo  $u$  that are equal to  $\xi_{k_j} \chi_j$  for some  $j$ . By expanding  $G$  with Lemma 2.10 we get

$$S = \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_k \check{G}(k) \left( \sum_{\mathfrak{n}} \beta_{\mathfrak{n}}^b \xi_k \psi(\mathfrak{n}) H_{N'}(\mathfrak{n}) \right) \right|^2.$$

Denote

$$S(\Psi_u) := \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\substack{k \\ \xi_k \psi \in \Psi_u}} \check{G}(k) \left( \sum_{\mathfrak{n}} \beta_{\mathfrak{n}}^b \xi_k \psi(\mathfrak{n}) H_{N'}(\mathfrak{n}) \right) \right|^2,$$

$$S(\Psi_u^c) := \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_{\mathfrak{n}} \beta_{\mathfrak{n}}^b \xi_k \psi(\mathfrak{n}) H_{N'}(\mathfrak{n}) \right) \right|^2.$$

Then by Cauchy–Schwarz  $((A + B)^2 \leq 2(A^2 + B^2))$  we have

$$S \ll S(\Psi_u) + S(\Psi_u^c).$$

6.2.1 *Bounding  $S(\Psi_u)$ .* By the definition of  $\beta_n^b$ , we see that for  $\xi_k \psi = \xi_{k_{j_0}} \chi_{j_0}$  we have

$$\sum_n \beta_n^b \xi_k \psi(n) H_{N'}(n) = \sum_{j \neq j_0} \mathcal{C}_W(\beta, \overline{\xi_{k_j} \chi_j} H_{N'}) \sum_n H_{N'}(n) \mathbf{1}_{(n, \bar{n}P(W))=1} \xi_k \psi \overline{\xi_{k_j} \chi_j}(n).$$

Thus, by Cauchy-Schwarz on  $j$  and  $k$  (recall that for  $\psi \xi_k \in \Psi$  we have  $|k| \leq (\nu_1)^{-2}$  by  $Q$ -regularity) and using  $|\check{G}(k)| \ll \nu_1$  we have

$$\begin{aligned} S(\Psi_u) &\ll \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{j \leq J} \mathcal{C}_W(\beta, \overline{\xi_{k_j} \chi_j} H_{N'}) \right. \\ &\quad \times \sum_{\substack{k \\ \xi_k \psi \in \Psi_u \\ \xi_k \psi \neq \xi_{k_j} \chi_j}} \check{G}(k) \sum_n H_{N'}(n) \mathbf{1}_{(n, \bar{n}P(W))=1} \xi_k \psi \overline{\xi_{k_j} \chi_j}(n) \left. \right|^2 \\ &\ll \sup_{j \leq J} \frac{(\log X)^{3C'}}{\varphi_{\mathbb{Z}[i]}(u)} \sum_k |\check{G}(k)|^2 \sum_{\substack{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times \\ \xi_k \psi \in \Psi_u \\ \xi_k \psi \neq \xi_{k_j} \chi_j}} \left| \sum_n H_{N'}(n) \mathbf{1}_{(n, \bar{n}P(W))=1} \xi_k \psi \overline{\xi_{k_j} \chi_j}(n) \right|^2 \\ &\ll \sup_{\substack{j \leq J \\ |k| \leq \nu_1^{-2} \\ \xi_k \psi \neq \xi_{k_j} \chi_j}} \frac{\nu_1^2 (\log X)^{4C'}}{\varphi_{\mathbb{Z}[i]}(u)} \left| \sum_n H_{N'}(n) \mathbf{1}_{(n, \bar{n}P(W))=1} \xi_k \psi \overline{\xi_{k_j} \chi_j}(n) \right|^2. \end{aligned}$$

By Lemma 6.4 we get

$$S(\Psi_u) \ll \frac{\nu_1^2 \nu_2^2 N^2}{\varphi_{\mathbb{Z}[i]}(u) (\log X)^C}.$$

6.2.2 *Bounding  $S(\Psi_u^{\mathbb{C}})$ .* By Cauchy-Schwarz  $((A+B)^2 \leq 2(A^2+B^2))$  we have

$$\begin{aligned} S(\Psi_u^{\mathbb{C}}) &= \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_n \beta_n^b \xi_k \psi(n) H_{N'}(n) \right) \right|^2 \\ &\ll S_1(\Psi_u^{\mathbb{C}}) + S_2(\Psi_u^{\mathbb{C}}), \end{aligned}$$

with

$$\begin{aligned} S_1(\Psi_u^{\mathbb{C}}) &:= \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_n \beta_n \xi_k \psi(n) H_{N'}(n) \right) \right|^2, \\ S_2(\Psi_u^{\mathbb{C}}) &:= \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{j \leq J} \mathcal{C}_W(\beta, \overline{\xi_{k_j} \chi_j} H_{N'}) \right. \\ &\quad \times \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_n H_{N'}(n) \mathbf{1}_{(n, \bar{n}P(W))=1} \xi_k \psi \overline{\xi_{k_j} \chi_j}(n) \right) \left. \right|^2. \end{aligned}$$

By the assumption that  $\beta_{\mathbf{n}}$  is  $Q$ -regular we have

$$S_1(\Psi_u^{\mathbf{C}}) \ll \frac{\nu_1^2 \nu_2^2 N^2}{\varphi_{\mathbb{Z}[i]}(u)(\log X)^{C_1}}.$$

For  $S_2(\Psi_u^{\mathbf{C}})$  we have by Cauchy–Schwarz on  $j$

$$S_2(\Psi_u^{\mathbf{C}}) := \sup_{j \leq J} \frac{(\log X)^{C'}}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{\mathbf{n}}P(W))=1} \xi_k \psi \overline{\xi_{k_j} \chi_j}(\mathbf{n}) \right) \right|^2,$$

we write

$$\sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) = \sum_k \check{G}(k) - \check{G}(k_j) \mathbf{1}_{\psi=\chi_j} - \sum_{\substack{k \\ \xi_k \psi \in \Psi_u \\ \xi_k \psi \neq \xi_{k_j} \chi_j}} \check{G}(k).$$

The contribution from the third sum may be extracted from  $S_2(\Psi_u^{\mathbf{C}})$  by Cauchy–Schwarz and bounded by the same argument as with  $S(\Psi_u)$ . Thus, we are left with bounding

$$\begin{aligned} S_3(\Psi_u^{\mathbf{C}}) &:= \sup_{j \leq J} \frac{(\log X)^{C'}}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_k \check{G}(k) \left( \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{\mathbf{n}}P(W))=1} \xi_k \psi \overline{\xi_{k_j} \chi_j}(\mathbf{n}) \right) \right. \\ &\quad \left. - \check{G}(k_j) \mathbf{1}_{\psi=\chi_j} \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{\mathbf{n}}P(W))=1} \right|^2 \\ &= \sup_{j \leq J} \frac{(\log X)^{C'}}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) G(\arg \mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{\mathbf{n}}P(W))=1} \psi \overline{\xi_{k_j} \chi_j}(\mathbf{n}) \right. \\ &\quad \left. - \check{G}(k_j) \mathbf{1}_{\psi=\chi_j} \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \mathbf{1}_{(\mathbf{n}, \bar{\mathbf{n}}P(W))=1} \right|^2. \end{aligned}$$

The claim now follows by orthogonality of characters and Lemma 6.4 since  $N^{1-\eta} > Q^3$  and  $|u|, |u_j| \leq Q$ .  $\square$

*Remark 6.6.* When constructing the approximation  $\beta_{\mathbf{n}}^{\#}$  we have a choice of using either the physical space or the Fourier space. To approximate  $\beta_{\mathbf{n}}$  with respect to arithmetic progressions and sectors of  $\mathbb{Z}[i]$  we use the Fourier space (i.e. characters  $\chi, \xi$ ), whereas to approximate  $\beta_{\mathbf{n}}$  with respect to the size of  $\mathcal{N}\mathbf{n}$  we use the physical space (i.e. smooth partition  $H_{N'}$ ). These are the most convenient choice for using existing zero-density estimates and information about exceptional characters.

### 6.3 Partitioning the Type II sum

With the approximation for  $\beta_{\mathbf{n}}$  defined as in § 6.2, we can extract the main term from the Type II sum by writing

$$S(\alpha, \beta) = S(\alpha, \beta^{\#}) + S(\alpha, \beta^b).$$

By  $(a, b) = 1$ , we may restrict to  $(w, \bar{w}) = 1$ . Since we are working with rough numbers the condition  $(a, b) = 1$  may be dropped with a negligible error term. By definition,  $a_{\mathbf{n}} = \sum_{u \in \{\pm 1, \pm i\}} a_{uz}$

so that we have

$$S(\alpha, \beta^b) = \sum_{\substack{|w|^2 \sim M \\ (w, \bar{w})=1}} \sum_{|z|^2 \sim N} \alpha_w \beta_z^b \mathbf{1}_B(\operatorname{Re}(\bar{w}z)) + O_C\left(\frac{X^{1/2}|B|}{(\log X)^C}\right).$$

The two contributions are bounded by the following two propositions, which together imply Proposition 6.2.

**PROPOSITION 6.7.** *Suppose that the assumptions of Proposition 6.2 hold and let  $\beta_n^b$  be as in § 6.2. Then, for every  $C_1 > 0$ , there is some  $C_2 > 0$*

$$\sum_{\substack{|w|^2 \sim M \\ (w, \bar{w})=1}} \sum_{|z|^2 \sim N} \alpha_w \beta_z^b \mathbf{1}_B(\operatorname{Re}(\bar{w}z)) \ll \frac{X^{1/2}|B|}{(\log X)^{C_1}}.$$

**PROPOSITION 6.8.** *Suppose that the assumptions of Proposition 6.2 hold and let  $\beta_n^\#$  be as in § 6.2. Then, for any  $C > 0$ ,*

$$\sum_{\mathfrak{m}, \mathfrak{n}} \alpha_{\mathfrak{m}} \beta_{\mathfrak{n}}^\# a_{\mathfrak{m}\mathfrak{n}} = \sum_{j \leq J} \sum_{\mathfrak{m}, \mathfrak{n}} \frac{\alpha_{\mathfrak{m}} \beta_{\mathfrak{n}} \xi_{k_j} \chi_j(\mathfrak{m}\mathfrak{n})}{\mathcal{N}\mathfrak{m}\mathfrak{n}} \sum_{\mathfrak{a}} \frac{F(\mathcal{N}\mathfrak{m}\mathfrak{n}/\mathcal{N}\mathfrak{a})}{\widehat{F}(0)} a_{\mathfrak{a}}^{\omega \xi_{k_j} \chi_j(\mathfrak{a})} + O\left(\frac{X^{1/2}|B|}{(\log X)^C}\right).$$

*Remark 6.9.* We want to carry the condition  $(w, \bar{w}) = 1$  through the application of Cauchy–Schwarz. To see why, consider a situation where  $B \subseteq d_0\mathbb{Z}$  with  $d_0 > X^\eta$  being very smooth so that  $\tau(d_0)$  is larger than any fixed power of  $\log X$ . Let us split the Type II sum according to the g.c.d. of  $w$  and  $d_0$ , which gives us

$$\sum_{e|d_0} \sum_{w \equiv 0(e)} \sum_z \mathbf{1}_B(\operatorname{Re}(\bar{w}z)).$$

Now in the inner sum  $e|w$  means that  $e|b$  is automatic, so that the density on the inside is bumped up, that is, we morally have

$$\sum_{w \equiv 0(e)} 1 \approx \frac{M}{e^2}, \quad \sum_z \mathbf{1}_B(\operatorname{Re}(\bar{w}z)) \approx e \frac{X^{1/2}|B|}{M}.$$

Prior to Cauchy–Schwarz this is not an issue since we still get converging sum  $\sum_{e|d_0} e^{-1}$ . However, after applying Cauchy–Schwarz to  $w$ , we get

$$\sum_{e|d_0} \sum_{w \equiv 0(e)} \left| \sum_z \mathbf{1}_B(\operatorname{Re}(\bar{w}z)) \right|^2 \approx N|B|^2 \sum_{e|d_0} 1,$$

which means that we have picked up a large divisor function  $\tau(d_0)$ . This would be problematic since we can only save a fixed power of  $\log X$  from Lemma 6.5. We resolve this issue by keeping the condition  $(w, \bar{w}) = 1$  so that  $w$  has no non-trivial integer divisors, but there are also other ways to deal with this.

#### 6.4 Proof of Lemma 6.1

Note that  $\mathcal{M}(u) \leq Q$  implies that  $|u| \leq Q$ , so that by  $N > X^\eta Q^3$  we have  $N > X^\eta |u|^3$ . By Lemma 2.14 we can take for  $\sigma_Q := 1 - \frac{C'_2 \log \log Q}{\log Q}$  with some large  $C'_2 > 0$  to get

$$J \leq N^*(\sigma_Q, X^\eta, X^\eta, Q^2) \leq (\log X)^{C_2}$$

and let  $\Psi$  be the set of primitive characters  $\xi\psi$  such that  $L(s, \xi\psi)$  has a zero counted in the above with  $\mathcal{M}(u) \leq Q$ . Recall that we now specify

$$\beta_{\mathbf{n}} := \mathbf{1}_{\mathcal{N}\mathbf{n} \sim N} \mu(\mathcal{N}\mathbf{n}) \mathbf{1}_{(\mathbf{n}, P(W))=1}.$$

Then, for Lemma 6.1, we need to show that if  $\Psi_u$  denotes the set of characters modulo  $u$  which are induced by  $\Psi$ , then for any  $N' \in [N, 2N]$  for any  $C_1 > 0$  there is some  $C_2 > 0$  such that

$$\frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_{\mathcal{N}\mathbf{n} \in (N', N'(1+\nu_2)]} \beta_{\mathbf{n}} \xi_k \psi(\mathbf{n}) \right) \right|^2 \ll \frac{\nu_1^2 \nu_2^2 N^2}{\varphi_{\mathbb{Z}[i]}(d) (\log X)^{C_1}}. \quad (6.3)$$

The contribution from the trivial character  $\psi = \psi_0$  is bounded by a similar but easier argument as below, using Heath-Brown's identity and the Vinogradov strength zero-free region of Coleman [Col90]. We then restrict to  $\psi \neq \psi_0$ . We apply § 2.1 to  $\mathcal{N}\mathbf{n}$  with  $\nu_1 = X^{-\eta_1}$ , using the assumption that  $\eta_2$  is small compared with  $\eta_1$  to replace  $\mathcal{N}\mathbf{n} \in (N', N'(1+\nu_2)]$  with a smooth weight. It then suffices to show that for any  $N_1 \sim N$ , we have

$$\frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\substack{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times \\ \psi \neq \psi_0}} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_{\mathbf{n}} \beta_{\mathbf{n}} \xi_k \psi(\mathbf{n}) F_{N_1}(\mathcal{N}\mathbf{n}) \right) \right|^2 \ll \frac{\nu_1^4 N^2}{\varphi_{\mathbb{Z}[i]}(d) (\log X)^{C_1}}.$$

The proof strategy is classical so we will be brief. We use the Heath-Brown identity [IK04, (13.58)] with  $K = 3$

$$\mu(n) = \sum_{k=1}^3 (-1)^{k+1} \binom{3}{k} \sum_{\substack{n=m_1 m_2 m_3 \\ m_j \leq 2N^{1/3}}} \mu(m_1) \mu(m_2) \mu(m_3).$$

Let  $F_1$  be as in § 2.1 with  $\nu = 1/2$ . Using the Heath-Brown identity and a dyadic decomposition (smooth for the free variable  $\mathfrak{a}$ ), we get sums of Type I

$$S_I := \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\substack{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times \\ \psi \neq \psi_0}} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_{\substack{\mathcal{N}\mathfrak{m} \sim M \\ (\mathfrak{m}\mathfrak{a}, P(W))=1}} F_{N_1}(\mathcal{N}\mathfrak{m}\mathfrak{a}) \alpha(\mathfrak{m}) F_{1,A}(\mathcal{N}\mathfrak{a}) \xi_k \psi(\mathfrak{m}\mathfrak{a}) \right) \right|^2,$$

with  $M \ll 2N^{2/3}$  and  $AM \sim N$  and sums of Type II

$$S_{II} := \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\substack{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times \\ \psi \neq \psi_0}} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_{\substack{\mathcal{N}\mathfrak{m}_j \sim M_j \\ (\mathfrak{m}_j, P(W))=1}} F_{N_1}(\mathcal{N}\mathfrak{m}_1 \mathfrak{m}_2 \mathfrak{m}_3) \right. \right. \\ \left. \left. \times \alpha_1(\mathfrak{m}_1) \alpha_2(\mathfrak{m}_2) \alpha_3(\mathfrak{m}_3) \xi_k \psi(\mathfrak{m}_1 \mathfrak{m}_2 \mathfrak{m}_3) \right) \right|^2,$$

with  $\alpha_1(\mathfrak{m}) = 1$  or  $\alpha_1(\mathfrak{m}) = \mu(\mathcal{N}\mathfrak{m})$ ,  $M_1 M_2 M_3 \sim M$  and

$$N^{1/6} \ll M_1 \ll N^{1/3}.$$



To see this note that we are in the Type I case unless all of the variables  $n_j$  are  $\ll N^{1/3}$ , and in that case we can take the  $M_1$  for the Type II sum to be the range of largest variable  $m_j, n_j$ , which must be  $\gg N^{1/6}$ .

To show (6.3) it then suffices to show that

$$S_I, S_{II} \ll \frac{\nu_1^4 N^2}{\varphi_{\mathbb{Z}[i]}(d)(\log X)^{C_1}}.$$

For  $S_I$ , we let  $D = X^{\eta_1}$  and write

$$\mathbf{1}_{(\mathfrak{a}, P(W))=1} = \sum_{\mathfrak{d} | (\mathfrak{a}, P(W))} \mu(\mathfrak{d}) = \sum_{\substack{\mathfrak{d} | (\mathfrak{a}, P(W)) \\ \mathcal{N}\mathfrak{d} \leq D}} \mu(\mathcal{N}\mathfrak{d}) + \sum_{\substack{\mathfrak{d} | (\mathfrak{a}, P(W)) \\ \mathcal{N}\mathfrak{d} > D}} \mu(\mathcal{N}\mathfrak{d}).$$

The contribution from  $\mathcal{N}\mathfrak{d} > D$  is bounded by using Lemma 2.2, after using orthogonality of characters and Lemma 2.10. For  $S_I$ , we then need to bound

$$S'_I := \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\substack{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times \\ \psi \neq \psi_0}} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} \check{G}(k) \left( \sum_{\substack{\mathfrak{a} \\ (\mathfrak{m}, P(W))=1}} \sum_{\substack{\mathfrak{d} | (\mathfrak{n}, P(W)) \\ \mathcal{N}\mathfrak{d} \sim D}} \mu(\mathfrak{d}) F_{1,A}(\mathcal{N}\mathfrak{d}\mathfrak{a}) \alpha(\mathfrak{m}) \xi_k \psi(\mathfrak{d}\mathfrak{m}\mathfrak{a}) \right) \right|^2,$$

with  $D \ll X^{\eta_1}$  and  $M \ll N^{2/3}$ .

Denote

$$\begin{aligned} M(t, \xi\psi) &:= \sum_{\mathcal{N}\mathfrak{m} \sim M} \mathbf{1}_{(\mathfrak{m}, P(W))=1} \alpha(\mathfrak{m}) (\mathcal{N}\mathfrak{m})^{-it} \xi\psi(\mathfrak{m}), \\ M_j(t, \xi\psi) &:= \sum_{\mathcal{N}\mathfrak{m} \sim M_j} \mathbf{1}_{(\mathfrak{m}, P(W))=1} \alpha_j(\mathfrak{m}) (\mathcal{N}\mathfrak{m})^{-it} \xi\psi(\mathfrak{m}), \\ A(t, \xi\psi) &:= \sum_{\mathfrak{a}} \sum_{\substack{\mathfrak{d} | (\mathfrak{a}, P(W)) \\ \mathcal{N}\mathfrak{d} \leq D}} \mu(\mathfrak{d}) F_{1,A}(\mathcal{N}\mathfrak{d}\mathfrak{a}) (\mathcal{N}\mathfrak{d}\mathfrak{a})^{-it} \xi\psi(\mathfrak{a}). \end{aligned}$$

Then, for  $\psi \neq \psi_0$ , we have the standard point-wise bounds

$$A(t, \xi_k \psi) \ll D(1 + |t|)(1 + |k|)|u| \quad (6.4)$$

and for  $\alpha_1 = 1$  or  $\alpha_1 = \mu$  with  $M_1 \gg X^{1/6}$  once  $C_2$  is large compared with  $C_1$ , with  $\xi\psi \notin \Psi_u$ ,

$$M_1(t, \xi\psi) \ll \frac{M_1}{(\log X)^{C_1}}. \quad (6.5)$$

The bound (6.4) follows by the Polyá–Vinogradov bound (i.e. the convexity bound for  $L(s, \xi\psi)$  in the  $u$  aspect, Lemma 2.18). The bound (6.5) follows by the truncated Perron's formula and shifting the contour to  $(1 + \sigma_Q)/2$  (justified by  $\xi\psi \notin \Psi_u$ ), using the bound Lemma 2.16 for  $1/L(s, \xi\psi)$ , and taking  $C'_2 > 0$  in the definition of  $\sigma_Q$  sufficiently large.

By Mellin inversion (Lemma 2.9), we get

$$F_{N_1}(x) = \frac{1}{2\pi i} \int \dot{F}(s) N_1^s x^{-s} ds,$$

with

$$|\dot{F}(s)| \ll_C \nu_1(1 + \nu_1|t|)^{-C}.$$

Hence, we have

$$S'_I \ll J_I \quad \text{and} \quad S_{II} \ll J_{II},$$

with

$$J_I := \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\substack{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times \\ \psi \neq \psi_0}} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} |\check{G}(k)| \int |\dot{F}(it)| |MA(t, \xi_k \psi)| dt \right|^2,$$

$$J_{II} := \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\substack{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times \\ \psi \neq \psi_0}} \left| \sum_{\substack{k \\ \xi_k \psi \notin \Psi_u}} |\check{G}(k)| \int |\dot{F}(it)| |M_1 M_2 M_3(t, \xi_k \psi)| dt \right|^2.$$

To bound  $J_I$  we apply Cauchy–Schwarz on  $t, k$ , (6.4), and orthogonality of characters to get, for some coefficients  $\gamma_n$  and for some  $t, \xi$ ,

$$\begin{aligned} J_I &\ll \nu_1^{O(1)} \frac{D^2 |u|^2}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} \left| M(t, \xi \psi) \right|^2 \\ &\ll \nu_1^{O(1)} D^2 |u|^2 \sum_{\substack{\mathbf{n}_1 \equiv \mathbf{n}_2 (u) \\ \mathcal{N}_{\mathbf{n}_1}, \mathcal{N}_{\mathbf{n}_2} \ll M}} |\gamma_{\mathbf{n}_1} \overline{\gamma_{\mathbf{n}_2}}| \\ &\ll \nu_1^{O(1)} D^2 M^2 + \nu_1^{O(1)} |u| D^2 M. \end{aligned}$$

By  $N > N^\eta |u|^3$ ,  $M \ll N^{2/3}$ , and  $D = X^{\eta_1}$  we get

$$J_I \ll N^{4/3+O(\eta_1)} + |u| N^{2/3+O(\eta_1)} \ll N^{-\eta} \frac{N^2}{\varphi_{\mathbb{Z}[i]}(u)}$$

since  $\eta_1$  is small compared with  $\eta$ , which is sufficient for bounding  $J_I$ .

For  $J_{II}$  we apply the bound (6.5) for  $M_1$  and Cauchy–Schwarz in the  $t, k$  variables to get

$$\begin{aligned} J_{II} &\ll \frac{M_1^2}{(\log X)^{C_1}} \iint |\dot{F}(it_1) \dot{F}(it_2)| \sum_{k_1, k_2} |\check{G}(k_1) \check{G}(k_2)| \\ &\quad \times \frac{1}{\varphi_{\mathbb{Z}[i]}(u)} \sum_{\psi \in (\mathbb{Z}[i]/u\mathbb{Z}[i])^\times} |M_2(t_1, \xi_{k_1} \psi) M_3(t_2, \xi_{k_2} \psi)|^2 dt_1 dt_2. \end{aligned}$$

By orthogonality of characters and Lemmas 2.10 and 2.9 we get

$$J_{II} \ll \frac{M_1^2}{(\log X)^{C_1}} T^2 \sum_{\substack{\mathbf{m}_{21}, \mathbf{m}_{22}, \mathbf{m}_{31}, \mathbf{m}_{32} \\ \mathcal{N}_{\mathbf{m}_{jk}} \sim M_j \\ \mathbf{m}_{21} \mathbf{m}_{31} \equiv \mathbf{m}_{22} \mathbf{m}_{32} (u) \\ |\mathcal{N}_{\mathbf{m}_{21}} - \mathcal{N}_{\mathbf{m}_{22}}| \ll \nu_1 M_2 \\ |\arg \mathbf{m}_{21} - \arg \mathbf{m}_{22}| \ll \nu_1 \\ |\mathcal{N}_{\mathbf{m}_{31}} - \mathcal{N}_{\mathbf{m}_{32}}| \ll \nu_1 M_3 \\ |\arg \mathbf{m}_{31} - \arg \mathbf{m}_{32}| \ll \nu_1}} 1.$$

Let us denote  $\mathbf{m}_{2j} = (w_j)$  and  $\mathbf{m}_{3j} = (z_j)$  so that

$$J_{II} \ll \frac{M_1^2}{(\log X)^{C_1}} T^2 \sum_{\substack{w_1, w_2, z_1, z_2 \\ |w_j|^2 \sim M_2 \\ |z_j|^2 \sim M_3 \\ w_1 z_1 \equiv w_2 z_2 (u) \\ |w_1 - w_2|^2 \ll \nu_1 M_2 \\ |z_1 - z_2|^2 \ll \nu_1 M_3}} 1.$$

Writing  $w_2 = w_1 + u$ ,  $z_1 = z_2 + v$  we get (using Lemma 2.4 to handle  $\tau(w)\tau(z)$ )

$$\begin{aligned} J_{II} &\ll \frac{M_1^2}{(\log X)^{C_1}} T^2 \sum_{\substack{w_1, u, v, z_2 \\ |w_1|^2 \sim M_2 \\ |u|^2 \ll \nu_1 M_2 \\ |z_2|^2 \sim M_3 \\ v \ll \nu_1 M_3 \\ w_1 v \equiv u z_2 \pmod{u}}} 1 \\ &\ll \frac{M_1^2}{(\log X)^{C_1}} T^2 \sum_{\substack{|w|^2, |z|^2 \ll \nu_1 M_2 M_3 \\ w \equiv z \pmod{u}}} \tau(w)\tau(z) \\ &\ll \frac{\nu^4 N^2}{\varphi_{\mathbb{Z}[i]}(u)(\log N)^{C_1 - O(1)}}, \end{aligned}$$

using  $M_1 \ll N^{1/3}$ ,  $M_1 M_2 M_3 \asymp N$  to get  $M_2 M_3 \gg N^{2/3} \gg X^\eta |u|^2$ .

## 7. Type II information: proof of Proposition 6.7

### 7.1 Cauchy–Schwarz

Let  $F_M(m) = F(m/M)$  with a fixed smooth majorant  $F$  for the interval  $[1, 2]$ , supported on  $[1/2, 3]$ . By applying Cauchy–Schwarz we get

$$\sum_{\substack{|w|^2 \sim M \\ (w, \bar{w})=1}} \sum_{|z|^2 \sim N} \alpha_w \beta_z^b \mathbf{1}_B(\operatorname{Re}(\bar{w}z)) \ll M^{1/2} U(\beta)^{1/2},$$

where

$$U(\beta) := \sum_{|z_1|^2, |z_2|^2 \sim N} \beta_{z_1}^b \overline{\beta_{z_2}^b} \sum_{(w, \bar{w})=1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_1)).$$

It then suffices to show that

$$U(\beta) \ll \frac{N|B|^2}{(\log X)^{C_1}}. \quad (7.1)$$

Define

$$\Delta = \Delta(z_1, z_2) := \operatorname{Im}(\overline{z_1} z_2) = |z_1 z_2| \sin(\arg z_2 - \arg z_1).$$

Note that typically  $|\Delta| \asymp N$ . We partition the sum into a main term and diagonal terms by writing

$$U(\beta) = V(\beta) + O(U_0(\beta) + U_1(\beta)),$$

where

$$\begin{aligned} V(\beta) &:= \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ \Delta \neq 0 \\ (z_1, z_2)=1}} \beta_{z_1}^b \overline{\beta_{z_2}^b} \sum_{(w, \bar{w})=1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_1)) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_2)), \\ U_0(\beta) &:= \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ \Delta=0}} |\beta_{z_1}^b \beta_{z_2}^b| \sum_{(w, \bar{w})=1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_1)) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_2)), \\ U_1(\beta) &:= \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ \Delta \neq 0 \\ (z_1, z_2) > 1}} |\beta_{z_1}^b \beta_{z_2}^b| \sum_{(w, \bar{w})=1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_1)) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_2)). \end{aligned}$$

For  $V(\beta)$  we apply § 2.1 with  $G: \mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{R}$  being a non-negative smooth function with the parameter  $\nu_1 = X^{-\eta_1}$  to the variables  $\arg z_1$  and  $\arg z_2$  to get

$$V(\beta) = \nu_1^{-2} \int_{(\mathbb{R}/2\pi\mathbb{Z})^2} V(\beta, \theta) d\theta_1 d\theta_2,$$

with

$$V(\beta, \theta) = \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ \Delta \neq 0 \\ (z_1, z_2) = 1}} \beta_{z_1, \theta_1}^b \overline{\beta_{z_2, \theta_2}^b} \sum_{(w, \bar{w}) = 1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_1)) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_2))$$

and

$$\beta_{z, \theta}^b = \beta_z^b G(\arg z - \theta).$$

We now further partition  $V(\beta)$  according to the size of  $|\Delta|$ . Let  $\nu_3 = X^{-\eta_3}$  and write

$$V(\beta) = V_{>\nu_3}(\beta) + V_{\leq \nu_3}(\beta),$$

where  $V_{>\nu_3}(\beta)$  is the part where  $|\sin(\theta_1 - \theta_2)| > \nu_3$ , which implies (since  $\eta_3$  is small compared with  $\eta_1$ )

$$|\Delta| = |z_1 z_2| |\sin(\arg z_2 - \arg z_1)| \gg \nu_3 N.$$

We then have the following lemmas, which together imply Proposition 6.7.

LEMMA 7.1 (Off-diagonal contribution). For  $|\sin(\theta_1 - \theta_2)| > \nu_3$ ,

$$V(\beta, \theta) \ll \frac{\nu_1^2 N |B|^2}{|\sin(\theta_1 - \theta_2)| (\log X)^{C_1}}.$$

LEMMA 7.2 (Diagonal contribution). We have

$$U_0(\beta) \ll_\varepsilon X^{1/2+\varepsilon} |B|$$

LEMMA 7.3 (Pseudo-diagonal contribution I). We have

$$U_1(\beta) \ll \frac{N |B|^2}{W^{1-\eta}}$$

LEMMA 7.4 (Pseudo-diagonal contribution II). We have

$$V_{\leq \nu_3}(\beta) \ll_\varepsilon X^\varepsilon \nu_3 N |B|^2$$

## 7.2 Proof of Lemma 7.1

It suffices to show that for  $|\theta_1 - \theta_2 \pmod{\pi}| > \nu_3$  we have

$$\begin{aligned} V(\beta, \theta) &= \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2) = 1}} \beta_{z_1, \theta_1}^b \overline{\beta_{z_2, \theta_2}^b} \\ &\quad \times \sum_{(w, \bar{w}) = 1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_1)) \mathbf{1}_B(\operatorname{Re}(\bar{w}z_2)) \\ &\ll \frac{\nu_1^2 N |B|^2}{|\sin(\theta_1 - \theta_2)| (\log X)^C}. \end{aligned} \tag{7.2}$$

We note that  $(z_1, z_2) = (z_1, \bar{z}_1) = (z_2, \bar{z}_2) = 1$  implies  $(\Delta, |z_1|^2 |z_2|^2) = 1$ . By symmetry, this can be seen from

$$(\overline{\Delta}, \overline{z_1}) = (\Delta, z_1) = (\operatorname{Im}(z_2 \bar{z}_1), z_1) = (z_2 \bar{z}_1, z_1) = 1.$$

Denoting  $\operatorname{Re}(\overline{w}z_1) = b_j$ , we have

$$i\Delta w = z_2b_1 - z_1b_2.$$

Let  $b_0 := (b_1, b_2)$ . Since  $(w, \overline{w}) = 1$ , we know that  $b_0|\Delta$ . Thus,

$$i(\Delta/b_0)w = z_2b_1/b_0 - z_1b_2/b_0. \quad (7.3)$$

Denoting  $b_j = b_0b'_j$ , we have

$$V(\beta, \theta) = \sum_{b_0 \geq 1} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{b_0}}} \beta_{z_1, \theta_1}^b \overline{\beta_{z_2, \theta_2}^b} \sum_{\substack{b'_2 \equiv ab'_1 \pmod{\Delta/b_0} \\ (b'_1b'_2, \Delta/b_0)=1 \\ (b'_1, b'_2)=1 \\ (w, \overline{w})=1}} \mathbf{1}_B(b_0b'_1) \mathbf{1}_B(b_0b'_2) F_M \left( \left| b_0 \frac{z_2b'_1 - z_1b'_2}{\Delta} \right|^2 \right).$$

We now wish to remove the smooth weight  $F_M$ . Recall that already  $b_0b'_j \in [Y, Y + X^{1/2-\eta}]$  by (3.2). We introduce a rough finer-than-dyadic partition for  $|z_1|^2, |z_2|^2$  by using  $H_{N'}$  as in § 6.2 with  $\nu_2 = X^{-\eta_2}$ . Let  $N_1, N_1 \sim N$ , and denote

$$\beta_{z,i}^b := \beta_z^b H_{N_i}(z) G(\arg z - \theta_i).$$

To prove (7.2) it then suffices to prove that for  $N_1, N_1 \sim N$  and for  $|\sin(\theta_1 - \theta_2)| \geq \nu_3$  we have

$$\begin{aligned} V'(\beta, N, \theta) &:= \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \\ &\quad \times \sum_{(w, \overline{w})=1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\overline{w}z_1)) \mathbf{1}_B(\operatorname{Re}(\overline{w}z_2)) \\ &\ll \frac{\nu_1^2 \nu_2^2 N |B|^2}{|\sin(\theta_1 - \theta_2)| (\log X)^C}. \end{aligned} \quad (7.4)$$

Using  $|\sin(\theta_1 - \theta_2)| > \nu_3$  we have  $\Delta = |z_1z_2| \sin(\arg z_2 - \arg z_1) \gg \nu_3 N$ , so that for some constant  $F_M(N, \theta)$

$$F_M \left( \left| b_0 \frac{z_2b'_1 - z_1b'_2}{\Delta} \right|^2 \right) = F_M(N, \theta) + O(\nu_2 \nu_3^{-2}). \quad (7.5)$$

Therefore, we have

$$V'(\beta, N, \theta) = F_M(N, \theta) V(\beta, N, \theta) + \nu_2 \nu_3^{-2} O(W(\beta, N, \theta)),$$

where

$$\begin{aligned} V(\beta, N, \theta) &:= \sum_{b_0} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{b_0}}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \sum_{\substack{b'_2 \equiv ab'_1 \pmod{\Delta/b_0} \\ (b'_1b'_2, \Delta/b_0)=1 \\ (b'_1, b'_2)=1 \\ (w, \overline{w})=1}} \mathbf{1}_B(b_0b'_1) \mathbf{1}_B(b_0b'_2) \\ W(\beta, N, \theta) &:= \sum_{b_0} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{b_0}}} |\beta_{z_1, 1}^b| |\beta_{z_2, 2}^b| \sum_{\substack{b'_2 \equiv ab'_1 \pmod{\Delta/b_0} \\ (b'_1b'_2, \Delta/b_0)=1 \\ (b'_1, b'_2)=1 \\ (w, \overline{w})=1}} \mathbf{1}_B(b_0b'_1) \mathbf{1}_B(b_0b'_2). \end{aligned}$$

Then (7.4) follows from the following two lemmas once  $\eta_3$  is small compared with  $\eta_2$  so that  $\nu_2 \nu_3^{-3} = X^{-\eta_2+3\eta_3} < X^{-\eta_2/2}$ , say.

LEMMA 7.5. For any  $C_1 > 0$ , there is some  $C_2 > 0$  for  $\beta_n^b$  as in § 6.2 such that for  $|\sin(\theta_1 - \theta_2)| > \nu_3$  we have

$$V(\beta, \mathbf{N}, \boldsymbol{\theta}) \ll \nu_1^2 \nu_2^2 \frac{N|B|^2}{|\sin(\theta_1 - \theta_2)|(\log X)^{C_1}}.$$

LEMMA 7.6. For  $|\sin(\theta_1 - \theta_2)| > \nu_3$ , we have

$$W(\beta, \mathbf{N}, \boldsymbol{\theta}) \ll \frac{\nu_1^2 \nu_2^2 N|B|^2 (\log X)^{O(1)}}{|\sin(\theta_1 - \theta_2)|}.$$

7.2.1 *Proof of Lemma 7.5.* The condition  $(w, \bar{w}) = 1$  has served its purpose so we remove it by expanding

$$\mathbf{1}_{(w, \bar{w})=1} = \sum_{\ell|w} \mu(\ell),$$

where  $\ell$  runs over integers. Writing  $w = \ell w'$ , we get from (7.3)

$$i(\Delta/b_0)\ell w' = z_2 b'_1 - z_1 b'_2. \quad (7.6)$$

We have

$$\left(\frac{\ell\Delta}{b_0}, b'_j\right) = 1. \quad (7.7)$$

To see this, by using  $(b'_1, b'_2) = 1$  and (7.6), we get that for  $j \in \{1, 2\}$ ,

$$\left(\frac{\ell\Delta}{b_0}, b'_j\right) \mid \left(\frac{\ell\Delta}{b_0}, z_j\right),$$

and (7.7) follows since the left-hand side is an integer and by  $(z_j, \bar{z}_j) = 1$  the only integer dividing  $z_j$  is 1. Similarly, we also get  $(z_1 z_2, \ell) = 1$  by (7.6). Thus,

$$a := z_2/z_1 \equiv b'_2/b'_1 \pmod{\ell}$$

is congruent to a rational integer, which is equivalent to saying that  $\Delta \equiv 0 \pmod{\ell}$ . Thus, we get

$$V(\beta, \mathbf{N}, \boldsymbol{\theta}) = \sum_{b_0, \ell} \mu(\ell) \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2) = 1 \\ \Delta \equiv 0 \pmod{[b_0, \ell]}}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \sum_{\substack{b'_2 \equiv ab'_1 \pmod{\ell\Delta/b_0} \\ (b'_1 b'_2, \ell\Delta/b_0) = 1 \\ (b'_1, b'_2) = 1}} \mathbf{1}_B(b_0 b'_1) \mathbf{1}_B(b_0 b'_2).$$

By expanding the congruence  $b'_2 \equiv ab'_1 \pmod{\ell\Delta/b_0}$  into Dirichlet characters and splitting into primitive characters we get

$$\begin{aligned} V(\beta, \mathbf{N}, \boldsymbol{\theta}) &= \sum_{b_0, \ell} \mu(\ell) \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2) = 1 \\ \Delta \equiv 0 \pmod{[b_0, \ell]}}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \\ &\quad \times \frac{1}{\varphi(\ell\Delta/b_0)} \sum_{d|\ell\Delta/b_0} \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2) = 1 \\ (b'_1 b'_2, \ell\Delta/(db_0)) = 1}} \chi(b'_1) \overline{\chi(b'_2)} \right). \end{aligned}$$

We separate  $db_0 > X^{\delta+\eta/2}$  and  $db_0 \leq X^{\delta+\eta/2}$ , that is, write

$$V(\beta, \mathbf{N}, \boldsymbol{\theta}) = V_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) + V_{>}(\beta, \mathbf{N}, \boldsymbol{\theta}).$$

For the large  $db_0$ , we use the expansion

$$\mathbf{1}_{(b'_1, b'_2)=1} = \sum_{c|(b'_1, b'_2)} \mu(c)$$

to get

$$V_{>}(\beta, \mathbf{N}, \boldsymbol{\theta}) = \sum_{b_0, \ell} \mu(\ell) \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{[b_0, \ell]}}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \\ \times \frac{1}{\varphi(\ell\Delta/b_0)} \sum_{\substack{d|\ell\Delta/b_0 \\ db_0 > X^{\delta+\eta/2}}} \sum_{(c, d)=1} \mu(c) \sum_{\chi(d)}^* \chi(a) \left| \sum_{\substack{cb_0 b \in B \\ (cb, \ell\Delta/(db_0))=1}} \chi(b) \right|^2.$$

We split the sum according to  $a_0 = (b_0, \ell)$  which gives us

$$V_{>}(\beta, \mathbf{N}, \boldsymbol{\theta}) = \sum_{\substack{b_0 \\ a_0|b_0}} \sum_{(\ell, b_0)=a_0} \mu(\ell) \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{(b_0\ell/a)}}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \\ \times \frac{1}{\varphi(\ell\Delta/b_0)} \sum_{\substack{d|\ell\Delta/b_0 \\ db_0 > X^{\delta+\eta/2}}} \sum_{(c, d)=1} \mu(c) \sum_{\chi(d)}^* \chi(a) \left| \sum_{\substack{cb_0 b \in B \\ (cb, \ell\Delta/(db_0))=1}} \chi(b) \right|^2.$$

Recall that by  $|\sin(\theta_1 - \theta_2)| \gg \nu_3$  we have  $|\Delta| \gg \nu_3 N$ . For any fixed  $D$  with  $\nu_3 N \ll D \ll N$  we have (combining the variables  $z_2 \bar{z}_1 = z$ , using a divisor bound, and recalling that  $\Delta = \text{Im}(z_2 \bar{z}_1)$ )

$$\sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ |\Delta| = D \gg \nu_3 N}} \frac{1}{\varphi(\ell\Delta/b_0)} \ll X^{\eta/200} \sum_{\substack{|z|^2 \sim N^2 \\ |\text{Im}(z)| = D \gg \nu_3 N}} \frac{b_0}{\ell D} \ll \frac{X^{\eta/200} b_0}{\nu_3 \ell},$$

which gives us (since  $\eta_3$  is small compared with  $\eta$ )

$$V_{>}(\beta, \mathbf{N}, \boldsymbol{\theta}) \ll \sup_{L \ll X} \frac{X^{\eta/100}}{L} \sum_{\substack{b_0 \\ a_0|b_0}} \sum_{\substack{\ell \sim L \\ \ell \equiv 0 \pmod{a_0}}} b_0 \sum_{\substack{D \ll N \\ D \equiv 0 \pmod{(b_0\ell/a_0)}}} \sum_{\substack{d|\ell D/b_0 \\ db_0 > X^{\delta+\eta/2}}} \sum_{(c, d)=1} \\ \times \sum_{\chi(d)}^* \left| \sum_{\substack{cb_0 b \in B \\ (cb, \ell D/(db_0))=1}} \chi(b) \right|^2.$$

By a dyadic partition, we get (denoting  $D' = \ell D/(db_0)$ )

$$V_{>}(\beta, \mathbf{N}, \boldsymbol{\theta}) \ll \sup_{\substack{B_0, D_0, L \\ X^{\delta+\eta/2} \ll D_0 B_0 \ll LN}} \frac{X^{\eta/50} B_0}{L} \sum_{\substack{b_0 \sim B_0 \\ a_0|b_0}} \sum_{\substack{\ell \sim L \\ \ell \equiv 0 \pmod{a_0}}} \sum_{\substack{D' \sim LN/B_0 D_0 \\ D' \equiv 0 \pmod{\ell^2/a_0}}} \\ \times \sum_{d \sim D_0} \sum_{(c, d)=1} \sum_{\chi(d)}^* \left| \sum_{\substack{cb_0 b \in B \\ (cb, D')=1}} \chi(b) \right|^2$$

$$\begin{aligned}
 &\ll \sup_{\substack{B_0, D_0, L \\ X^{\delta+\eta/2} \ll D_0 B_0 \ll LN}} \frac{X^{\eta/50} B_0}{L} \sum_{\substack{b_0 \sim B_0 \\ a_0 | b_0}} \sum_{\substack{\ell \sim L \\ \ell \equiv 0 \pmod{a_0}}} \sum_{\substack{D' \sim LN/B_0 D_0 \\ D' \equiv 0 \pmod{\ell^2/a_0}}} \\
 &\times \sup_{D''} \sum_{d \sim D_0} \sum_{(c,d)=1} \sum_{\chi(d)}^* \left| \sum_{\substack{cb_0 b \in B \\ (cb, D'')=1}} \chi(b) \right|^2 \\
 &\ll X^{\eta/40} N \sup_{\substack{B_0, D_0, L \\ X^{\delta+\eta/2} \ll D_0 B_0 \ll LN}} \frac{1}{D_0 L} \sum_{b_0 \sim B_0} \sum_c \sup_{D''} \sum_{d \sim D_0} \sum_{\chi(d)}^* \left| \sum_{\substack{cb_0 b \in B \\ (cb, D'')=1}} \chi(b) \right|^2.
 \end{aligned}$$

By the large sieve for multiplicative characters (Lemma 2.11) we obtain

$$V_{>}(\beta, \mathbf{N}, \boldsymbol{\theta}) \ll X^{\eta/30} N(N + X^{1/2-\delta-\eta/2})|B| \ll X^{-\eta/4} N|B|^2,$$

since  $N \leq X^{-\eta}|B|$ . This is sufficient for Lemma 7.5 since  $\eta_1, \eta_2$  are small compared with  $\eta$ .

To bound the contribution from the small  $db_0$  recall that

$$\begin{aligned}
 V_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) &= \sum_{b_0, \ell} \mu(\ell) \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{[b_0, \ell]}}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \\
 &\times \frac{1}{\varphi(\ell \Delta / b_0)} \sum_{\substack{d | \ell \Delta / b_0 \\ db_0 \leq X^{\delta+\eta/2}}} \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ (b'_1 b'_2, \ell \Delta / (db_0))=1}} \chi(b'_1) \overline{\chi(b'_2)} \right).
 \end{aligned}$$

We expand the conditions  $(b'_1 b'_2, \ell \Delta / (db_0)) = 1$  by using the Möbius function to get

$$\begin{aligned}
 V_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) &= \sum_{(e_1, e_2)=1} \mu(e_1) \mu(e_2) \sum_{b_0, \ell} \mu(\ell) \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{[b_0, \ell]}}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \\
 &\times \frac{1}{\varphi(\ell \Delta / b_0)} \sum_{\substack{de_1 e_2 | \ell \Delta / b_0 \\ db_0 \leq X^{\delta+\eta/2}}} \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 e_1 b'_1, b_0 e_2 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(e_1 b'_1) \overline{\chi(e_2 b'_2)} \right).
 \end{aligned}$$

We have

$$\frac{1}{\varphi(\ell \Delta / b_0)} = \frac{b_0}{\ell |\Delta|} \prod_{p | \ell \Delta / b_0} \left( 1 + \frac{1}{p-1} \right) = \frac{b_0}{\ell |\Delta|} \sum_{d_1 | \ell \Delta / b_0} \frac{|\mu(d_1)|}{\varphi(d_1)}.$$

Since  $z_1, z_2$  are restricted to polar boxes, we have

$$|\Delta| = N_1^{1/2} N_2^{1/2} |\sin(\theta_2 - \theta_2)| (1 + O(\nu_2)) := D(1 + O(\nu_2)). \quad (7.8)$$

Plugging this in, we get

$$V_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) = V'_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) + O(W_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta})), \quad (7.9)$$



with

$$\begin{aligned}
 V'_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) &:= \frac{1}{D} \sum_{d_1} \frac{|\mu(d_1)|}{\varphi(d_1)} \sum_{(e_1, e_2)=1} \mu(e_1)\mu(e_2) \sum_{b_0, \ell} \frac{b_0 \mu(\ell)}{\ell} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{[b_0, \ell]}}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \\
 &\quad \times \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ [d_1, de_1e_2] | \ell \Delta / b_0}} \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0e_1b'_1, b_0e_2b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(e_1b''_1) \overline{\chi(e_2b''_2)} \right), \\
 W_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) &:= \frac{\nu_2}{D} \sum_{d_1} \frac{1}{\varphi(d_1)} \sum_{(e_1, e_2)=1} \sum_{b_0, \ell} \frac{b_0}{\ell} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ \Delta \equiv 0 \pmod{[b_0, \ell]}}} |\beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b}| \\
 &\quad \times \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ [d_1, de_1e_2] | \ell \Delta / b_0}} \left| \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0e_1b'_1, b_0e_2b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(e_1b''_1) \overline{\chi(e_2b''_2)} \right) \right|.
 \end{aligned}$$

Let us first consider  $V'_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta})$ . We note that the sums over  $d_1, e_2, e_2, \ell$  converge quickly, so that we expect to be able to bound the contribution from large ranges of  $d_1, e_2, e_2$  by crude estimates. We have

$$\frac{\ell}{(b_0, \ell)} \left| \Delta / b_0 \right| \quad \text{and} \quad \frac{[d_1, de_1e_2]}{([d_1, de_1e_2], \ell)} \left| \Delta / b_0 \right|.$$

Denoting

$$f := f(\ell, d_1, d, e_1, e_2) := b_0 \cdot \left[ \frac{[d_1, de_1e_2]}{([d_1, de_1e_2], \ell)}, \frac{\ell}{(b_0, \ell)} \right],$$

we get  $f|\Delta$ , which is equivalent to saying that

$$z_2 \equiv az_1(f)$$

for some  $a(f)$ . Note that  $|f| \ll N$ . Thus,

$$\begin{aligned}
 V'_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) &= \frac{1}{D} \sum_{d_1} \frac{|\mu(d_1)|}{\varphi(d_1)} \sum_{(e_1, e_2)=1} \mu(e_1)\mu(e_2) \sum_{b_0, \ell} \frac{b_0 \mu(\ell)}{\ell} \\
 &\quad \times \sum_{db_0 \leq X^{\delta+\eta/2}} \sum_{a(f)} \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0b'_1, b_0b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \overline{\chi(b'_2)} \right) \sum_{\substack{z_2 \equiv az_1(f) \\ (z_1, z_2)=1}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b}.
 \end{aligned}$$

We write

$$V'_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) = V''_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) + E_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}), \tag{7.10}$$

where  $E_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta})$  is the part where  $f > X^{\delta+\eta}$ .

For  $f \leq X^{\delta+\eta}$ , we note by  $(z_1, z_2) = 1$  we have  $(z_1 z_2, f) = 1$ . By dropping the condition  $(z_1, z_2) = 1$ , we get

$$V_{\leq}''(\beta, \mathbf{N}, \boldsymbol{\theta}) = V_{\leq}'''(\beta, \mathbf{N}, \boldsymbol{\theta}) + O(Z_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta})) \quad (7.11)$$

with

$$V_{\leq}'''(\beta, \mathbf{N}, \boldsymbol{\theta}) = \frac{1}{D} \sum_{d_1} \frac{|\mu(d_1)|}{\varphi(d_1)} \sum_{(e_1, e_2)=1} \mu(e_1) \mu(e_2) \sum_{b_0, \ell} \frac{b_0 \mu(\ell)}{\ell} \\ \times \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ f \leq X^{\delta+\eta}}} \sum_{a(f)} \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \overline{\chi}(b'_2) \right) \sum_{\substack{z_2 \equiv a z_1(f) \\ (z_1 z_2, f)=1}} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b}$$

and

$$Z_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) := \frac{1}{D} \sum_{d_1} \frac{1}{\varphi(d_1)} \sum_{(e_1, e_2)=1} \sum_{b_0, \ell} \frac{b_0}{\ell} \\ \times \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ f \leq X^{\delta+\eta}}} \sum_{a(f)} \left| \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \overline{\chi}(b'_2) \right) \right| \sum_{\substack{z_2 \equiv a z_1(f) \\ (z_1 z_2, f)=1 \\ |(z_1, z_2)|^2 \geq W}} |\beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b}|.$$

By expanding  $z_2 \equiv a z_1(f)$  with Dirichlet characters we have

$$V_{\leq}'''(\beta, \mathbf{N}) \leq \frac{1}{D} \sum_{d_1} \frac{1}{d_1} \sum_{(e_1, e_2)=1} \sum_{b_0, \ell} \frac{b_0}{\ell} \sum_{db_0 \leq X^{\delta+\eta/2}} \sum_{a(f)} \left| \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \overline{\chi}(b'_2) \right) \right| \\ \times \frac{\mathbf{1}_{f \leq X^{\delta+\eta}}}{\varphi_{\mathbb{Z}[i]}(f)} \sum_{\psi \in (\mathbb{Z}[i]/f\mathbb{Z}[i])^\times} \left| \sum_{z_1, z_2} \beta_{z_1, 1}^b \psi(z_1) \overline{\beta_{z_2, 2}^b \psi(z_2)} \right|.$$

By applying Cauchy–Schwarz on  $\psi$  and using Lemma 6.5 we get

$$|V_{\leq}'''(\beta, \mathbf{N})| \ll \frac{\nu_1^2 \nu_2^2 N^2}{D(\log X)^{C_1}} \sum_{d_1, e_1, e_2} \sum_{b_0, \ell} \sum_{db_0 \leq X^{\delta+\eta/2}} \frac{b_0}{d_1 \ell f^2} \\ \times \sum_{a(f)} \left| \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \overline{\chi}(b'_2) \right) \right|.$$

By Lemma 2.17 (with the argument  $ab'_1/b'_2$ ) and Lemma 2.1 this is bounded by

$$\begin{aligned} &\ll \frac{\nu_1^2 \nu_2^2 N^2}{D(\log X)^{C_1}} \sum_{d_1, b_0, \ell} \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1}} \sum_{\substack{e_1 | b'_1 \\ e_2 | b'_2}} \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ (b'_1 b'_2, d)=1}} \frac{b_0}{d_1 \ell f^2} \sum_{\substack{a(f) \\ (a, d)=1}} (d, b'_2 - ab'_1) \\ &\ll \frac{\nu_1^2 \nu_2^2 N^2}{D(\log X)^{C_1}} \sum_{d_1, b_0, \ell} \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1}} \sum_{\substack{e_1 | b'_1 \\ e_2 | b'_2}} \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ (b_1 b_2, d)=1}} \frac{\tau(d) b_0}{d_1 \ell f}. \end{aligned}$$

Using

$$f = b_0 \cdot \left[ \frac{[d_1, de_1 e_2]}{([d_1, de_1 e_2], \ell)}, \frac{\ell}{(b_0, \ell)} \right] \geq \frac{b_0 d}{(d, \ell)},$$

by Lemma 2.1 and (7.8), we get

$$\begin{aligned} |V_{\leq}'''(\beta, \mathbf{N})| &\ll \frac{\nu_1^2 \nu_2^2 N^2}{D(\log X)^{C_1}} \sum_{b_1, b_2 \in B} \sum_{d, \ell \leq X} \frac{\tau(d)(\ell, d)}{d\ell} \\ &\ll \frac{\nu_1^2 \nu_2^2 N |B|^2}{|\sin(\theta_2 - \theta_1)| (\log X)^{C_1}}, \end{aligned}$$

which is sufficient.

The argument for bounding  $Z_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta})$  from (7.11) is the same except that instead of expanding into Hecke characters and Lemma 6.5 we use the trivial bound

$$\begin{aligned} \sum_{\substack{z_2 \equiv az_1(f) \\ (z_1 z_2, f)=1 \\ |(z_1, z_2)|^2 \geq W}} |\beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b}| &= \sum_{|z_0|^2 \geq W} \sum_{\substack{z'_2 \equiv az'_1(f) \\ (z_1 z_2, f)=1}} |\beta_{z_0 z_1, 1}^b \overline{\beta_{z_0 z_2, 2}^b}| \\ &\ll (\log X)^{O(1)} \sum_{W < |z_0|^2 \ll N} \left( \frac{\nu_1 \nu_2 N}{|z_0|^2} + 1 \right) \left( \frac{\nu_1 \nu_2 N}{|z_0|^2 f^2} + 1 \right) \\ &\ll (\log X)^{O(1)} \left( \frac{\nu_1^2 \nu_2^2 N^2}{W f^2} + \nu_1 \nu_2 N (\log X) + N \right) \\ &\ll (\log X)^{O(1)} \frac{\nu_1^2 \nu_2^2 N^2}{W f^2}, \end{aligned}$$

where the last bound holds since  $N > X^{3\delta+3\eta}$  and  $f^2 \leq X^{\delta+2\eta}$ . This gives us

$$Z_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) \ll (\log X)^{O(1)} \frac{\nu_1^2 \nu_2^2 N |B|^2}{|\sin \theta_2 - \theta_1| W},$$

which suffices for Lemma 7.5.

For  $E_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta})$  from (7.10) with large  $f$  we will need to use a slightly different argument since the modulus  $f$  can be as large as  $N$  which would make  $f^2$  much bigger than  $N$ . We write

$$\begin{aligned} E_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) &= \frac{1}{D} \sum_{d_1} \frac{|\mu(d_1)|}{\varphi(d_1)} \sum_{(e_1, e_2)=1} \mu(e_1) \mu(e_2) \sum_{b_0, \ell} \frac{b_0 \mu(\ell)}{\ell} \\ &\quad \times \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ f > X^{\delta+\eta}}} \sum_{a(f)} \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \overline{\chi(b'_2)} \right) \sum_{z_2 \equiv az_1(f)} \beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b} \end{aligned}$$

$$\begin{aligned}
 &\ll \frac{1}{D} \sum_{d_1} \frac{1}{\varphi(d_1)} \sum_{e_1, e_2} \sum_{b_0, \ell} \frac{b_0}{\ell} \\
 &\quad \times \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ f > X^{\delta+\eta}}} \sum_{\chi(d)}^* \left| \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \bar{\chi}(b'_2) \right| \sum_{\substack{z_1, z_2 \\ f | \Delta}} |\beta_{z_1, 1}^b \overline{\beta_{z_2, 2}^b}| \\
 &\ll \frac{N^2}{D} \sum_{d_1} \frac{1}{\varphi(d_1)} \sum_{e_1, e_2} \sum_{b_0, \ell} \frac{b_0}{\ell} \\
 &\quad \times \sum_{\substack{db_0 \leq X^{\delta+\eta/2} \\ f > X^{\delta+\eta}}} \frac{1}{|f|} \sum_{\chi(d)}^* \left| \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \bar{\chi}(b'_2) \right|.
 \end{aligned}$$

By using  $f > X^{\delta+\eta}$  and  $db_0 \leq X^{\delta+\eta/2}$  we get  $f^{-1} \leq X^{-\eta/2} (b_0 d)^{-1}$ , which gives us

$$\begin{aligned}
 E_{\leq}(\beta, N, \theta) &\ll X^{-\eta/2} \frac{N^2}{D} \sum_{d_1 \ll X} \frac{1}{\varphi(d_1)} \sum_{e_1, e_2} \sum_{b_0, \ell \ll X} \frac{1}{\ell} \\
 &\quad \times \sum_{db_0 \leq X^{\delta+\eta}} \frac{1}{d} \sum_{\chi(d)}^* \left| \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \bar{\chi}(b'_2) \right|.
 \end{aligned}$$

By using the expansion

$$\mathbf{1}_{(b'_1, b'_2)=1} = \sum_{c | (b'_1, b'_2)} \mu(c),$$

Cauchy–Schwarz, orthogonality of characters, and a divisor bound, we have

$$\begin{aligned}
 &\sum_{\chi(d)}^* \sum_{e_1, e_2} \left| \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ e_j | b'_j}} \chi(b'_1) \bar{\chi}(b'_2) \right| \leq \sum_c \sum_{\chi(d)} \sum_{e_1, e_2} \left| \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ e_j | b'_j \\ c | b'_j}} \chi(b'_1) \bar{\chi}(b'_2) \right| \\
 &\ll_{\varepsilon} X^{\varepsilon} \left( \sum_{e_1, e'_1} \sum_{\substack{b_0 b_1, b_0 b'_1 \in B \\ e_1 | b_1 \\ e'_1 | b'_1}} d \mathbf{1}_{b_1 \equiv b'_1}(d) \right)^{1/2} \left( \sum_{e_2, e'_2} \sum_{\substack{b_0 b_2, b_0 b'_2 \in B \\ e_2 | b_2 \\ e'_2 | b'_2}} d \mathbf{1}_{b_2 \equiv b'_2}(d) \right)^{1/2} \\
 &\ll X^{\eta/4} \sum_{e_1, e'_1} \sum_{\substack{b_0 b_1, b_0 b'_1 \in B \\ e_1 | b_1 \\ e'_1 | b'_1}} d \mathbf{1}_{b_1 \equiv b'_1}(d),
 \end{aligned}$$

where the last bound follows by symmetry. Hence, we have

$$E_{\leq}(\beta, N, \theta) \ll X^{-\eta/4} \frac{N^2}{D} \sum_{\substack{b_0, d, e_1, e'_1 \\ db_0 \leq X^{\delta+\eta/2}}} \sum_{\substack{b_0 b_1, b_0 b'_1 \in B \\ e_1 | b_1 \\ e'_1 | b'_1}} \mathbf{1}_{b_1 \equiv b'_1}(d).$$

The contribution from  $b'_1 = b_1$  is bounded by

$$\ll X^{-\eta/8} \frac{N^2}{D} |B| X^{\delta+\eta/2} \ll X^{-\eta/8} \frac{N^2}{D} |B|^2.$$

The contribution from  $b'_1 \neq b_1$  is bounded by

$$\ll X^{-\eta/4} \frac{N^2}{D} \sum_{b_0, e_1, e'_1} \sum_{\substack{b_0 b_1, b_0 b'_1 \in B \\ e_1 | b_1 \\ e'_1 | b'_1}} \mathbf{1}_{b_1 \neq b'_1} \tau(b'_1 - b_1) \ll X^{-\eta/8} \frac{N^2}{D} |B|^2$$

by applying the divisor bound  $\tau(n) \ll_\varepsilon n^\varepsilon$ . Therefore, by (7.8) we get (recall that  $\nu_3$  is small compared with  $\nu$ )

$$E_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta}) \ll X^{-\eta/10} N |B|^2.$$

Finally, the error term  $W_{\leq}(\beta, \mathbf{N}, \boldsymbol{\theta})$  from (7.9) is bounded by exactly the same argument as above except that in the part  $f \leq X^{\delta+\eta}$  we use the trivial estimate

$$\sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ z_2 \equiv a z_1 (f)}} |\beta_{z_1,1}^b| |\beta_{z_2,2}^b| \ll (\log X)^{O(1)} \frac{\nu_1^4 N^2}{|f|^2}$$

instead of expanding into Hecke characters and Lemma 6.5.

**7.2.2 Proof of Lemma 7.6.** The argument is exactly the same as in § 7.2.1, except that in the part  $f \leq X^{\delta+\eta}$  we use the trivial estimate

$$\sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ z_2 \equiv a z_1 (f)}} |\beta_{z_1,1}^b| |\beta_{z_2,2}^b| \ll (\log X)^{O(1)} \frac{\nu_1^4 N^2}{|f|^2}$$

instead of expanding into Hecke characters and Lemma 6.5.

### 7.3 Proof of Lemma 7.2

Since  $\Delta = 0$ , we have

$$z_1 b_2 = z_2 b_1.$$

Multiplying both sides by  $\overline{w}$  and taking the imaginary parts we find

$$a_1 b_2 = a_2 b_1.$$

Hence, we get

$$U_0(\gamma) \ll \sum_{\substack{a_1, a_2 \ll X^{1/2} \\ b_1, b_2 \ll X^{1/2} \\ a_1 b_2 = a_2 b_1}} \mathbf{1}_B(b_1) \mathbf{1}_B(b_2) \ll_\varepsilon X^{1/2+\varepsilon} |B|,$$

by using the divisor bound  $\tau(\ell) \ll_\varepsilon \ell^\varepsilon$ .

### 7.4 Proof of Lemma 7.3

Since  $(z_1 z_2, P(W)) = 1$ , having  $(z_1, z_2) > 1$  implies  $|(z_1, z_2)|^2 \geq W$ . Let  $z_0 = (z_1, z_2)$  and  $z_j = z_0 z'_j$ .

Denoting  $w_0 = \overline{z_0} w$  and  $\Delta' = \text{Im}(\overline{z'_1} z'_2)$ , we have

$$b_j = \text{Re}(\overline{w_0} z'_j),$$

which implies that

$$i\Delta'w_0 = z'_2b_1 - z'_1b_2, \quad (7.12)$$

so that  $w_0$  is fixed once we fix  $z'_j, b_j$ . Furthermore, we have

$$\sum_{(z_0, P(W))=1} \mathbf{1}_{\overline{z_0}|w_0} \ll_{\varepsilon} W^{\varepsilon}.$$

Note also that  $(w_0, \overline{w_0}) = 1$  implies that  $(b_1, b_2) = b_0|\Delta'|$ . Then denoting  $b_j = b_0b'_j$  and  $a' = z'_2/z'_1$  we have

$$\begin{aligned} U_1(\beta) &= \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ \Delta \neq 0 \\ (z_1, z_2) > 1 \\ (z'_j, z'_j) = 1}} \beta_{z_1}^b \overline{\beta_{z_2}^b} \sum_{(w, \overline{w})=1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\overline{w}z_1)) \mathbf{1}_B(\operatorname{Re}(\overline{w}z_2)) \\ &\ll \sup_{W \ll Z \ll N} \sum_{\substack{|z'_1|^2, |z'_2|^2 \sim N/Z \\ \Delta \neq 0 \\ (z'_1, z'_2) = 1 \\ (z'_j, z'_j P(W)) = 1}} \sum_{\substack{z_0, w_0 \\ (w_0, \overline{w_0}) = 1 \\ (z_0, P(W)) = 1 \\ \overline{z_0}|w_0 \\ |z_0|^2 \sim Z}} |\beta_{z_0 z'_1}^b \overline{\beta_{z_0 z'_2}^b}| \mathbf{1}_B(\operatorname{Re}(\overline{w_0}z'_1)) \mathbf{1}_B(\operatorname{Re}(\overline{w_0}z'_2)) \\ &\ll_{\varepsilon} W^{\varepsilon} \sup_{W \ll Z \ll N} \sum_{b_0} \sum_{\substack{|z'_1|^2, |z'_2|^2 \sim N/Z \\ \Delta' \neq 0 \\ (z'_1, z'_2) = 1 \\ (z'_j, z'_j P(W)) = 1 \\ \Delta' \equiv 0 (b_0)}} \sum_{\substack{b'_2 \equiv ab'_1 (\Delta'/b_0) \\ (b'_1 b'_2, \Delta'/b_0) = 1 \\ (b'_1, b'_2) = 1}} \mathbf{1}_B(b_0 b'_1) \mathbf{1}_B(b_0 b'_2). \end{aligned}$$

Thus, denoting

$$V_Z(\beta) := \sum_{b_0} \sum_{\substack{|z'_1|^2, |z'_2|^2 \sim N/Z \\ \Delta' \neq 0 \\ (z'_1, z'_2) = 1 \\ (z'_j, z'_j P(W)) = 1 \\ \Delta' \equiv 0 (b_0)}} \sum_{\substack{b'_2 \equiv ab'_1 (\Delta'/b_0) \\ (b'_1 b'_2, \Delta'/b_0) = 1 \\ (b'_1, b'_2) = 1}} \mathbf{1}_B(b_0 b'_1) \mathbf{1}_B(b_0 b'_2),$$

it suffices to show that

$$V_Z(\beta) \ll_{\varepsilon} W^{\varepsilon} \frac{N|B|^2}{W}. \quad (7.13)$$

We apply a similar argument as in § 7.5 except that certain parts will be easier by positivity. By expanding into Dirichlet characters, we get

$$V_Z(\beta) \ll (\log X) \sum_{b_0} \sum_{\substack{|z'_1|^2, |z'_2|^2 \sim N/Z \\ \Delta' \neq 0 \\ (z'_1, z'_2) = 1 \\ (z'_j, z'_j P(W)) = 1 \\ \Delta' \equiv 0 (b_0)}} \frac{b_0}{|\Delta|} \sum_{d|\Delta'/b_0} \left| \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2) = 1 \\ (b'_1 b'_2, \Delta'/(db_0)) = 1}} \chi(b'_1) \overline{\chi}(b'_2) \right) \right|.$$

We split into the parts  $db_0 \leq X^{\delta+\eta/2}$  and  $db_0 > X^{\delta+\eta/2}$

$$V_Z(\beta) \ll (\log X)(V_{Z, \leq}(\beta) + V_{Z, >}(\beta)).$$

For large  $db_0$  we expand the condition  $(b'_1, b'_2)$  to get

$$V_{Z,>}(\beta) \ll \sum_c \sum_{b_0, \substack{|z'_1|^2, |z'_2|^2 \sim N/Z \\ \Delta' \neq 0 \\ (z'_1, z'_2)=1 \\ (z'_j, z'_j P(W))=1 \\ \Delta' \equiv 0 (b_0)}} \frac{b_0}{|\Delta|} \sum_{\substack{d|\Delta'/b_0 \\ db_0 > X^{\delta+\eta}}} \sum_{\chi(d)}^* \left| \sum_{\substack{cb_0 b \in B \\ (b, \Delta'/(db_0))=1}} \chi(b) \overline{\chi}(b'_2) \right|^2.$$

By using the estimate

$$\sum_{\substack{|z'_1|^2, |z'_2|^2 \sim N/Z \\ (z'_1, z'_2)=1 \\ (z'_j, z'_j)=1 \\ \Delta'=D}} 1 \ll_{\varepsilon} X^{\varepsilon} \frac{N}{Z}, \quad (7.14)$$

we get (denoting  $D = D' db_0$ )

$$\begin{aligned} V_{Z,>}(\beta) &\ll_{\varepsilon} X^{\varepsilon} \frac{N}{Z} \sum_c \sum_{b_0, \substack{db_0 > X^{\delta+\eta} \\ D \ll N/Z \\ db_0 | D}} b_0 \sum_{\substack{D \ll N/Z \\ db_0 | D}} \frac{1}{D} \sum_{\chi(d)}^* \left| \sum_{\substack{cb_0 b \in B \\ (b, D/(db_0))=1}} \chi(b) \overline{\chi}(b'_2) \right|^2 \\ &\ll_{\varepsilon} X^{\varepsilon} \frac{N}{Z} \sum_c \sum_{b_0, \substack{db_0 > X^{\delta+\eta} \\ D' \ll N/(Zdb_0)}} b_0 \sum_{\substack{D' \ll N/(Zdb_0)}} \frac{1}{D' db_0} \sum_{\chi(d)}^* \left| \sum_{\substack{cb_0 b \in B \\ (b, D')=1}} \chi(b) \overline{\chi}(b'_2) \right|^2 \\ &\ll_{\varepsilon} X^{\varepsilon} \frac{N}{Z} \sum_c \sum_{b_0, \substack{\sup_{D'} \\ X^{\delta+\eta} < db_0 \ll N/Z}} \sum_{\substack{X^{\delta+\eta} < db_0 \ll N/Z}} \frac{1}{d} \sum_{\chi(d)}^* \left| \sum_{\substack{cb_0 b \in B \\ (b, D')=1}} \chi(b) \overline{\chi}(b'_2) \right|^2. \end{aligned}$$

By applying the multiplicative large sieve (Lemma 2.11) similarly as in § 7.2.1, we get

$$V_{Z,>}(\beta) \ll (\log X)^{O(1)} \frac{N|B|^2}{W},$$

which suffices for (7.13).

For small  $db_0$ , using Lemma 2.17 we write

$$\begin{aligned} V_{Z,\leq}(\beta) &\ll \sum_{db_0 \leq X^{\delta+\eta}} b_0 \sum_{a(db_0)} \sum_{\substack{|z'_1|^2, |z'_2|^2 \sim N/Z \\ \Delta' \neq 0 \\ (z'_1, z'_2)=1 \\ (z'_j, z'_j P(W))=1 \\ z'_2 \equiv az'_1 (db_0)}} \frac{1}{|\Delta'|} \left| \sum_{\chi(d)}^* \chi(a) \left( \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ (b'_1 b'_2, \Delta'/(db_0))=1}} \chi(b'_1) \overline{\chi}(b'_2) \right) \right| \\ &\ll \sum_{db_0 \leq X^{\delta+\eta}} b_0 \sum_{a(db_0)} \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ (b'_1 b'_2, d)=1}} (d, b'_2 - ab'_1) \sum_{\substack{|z'_1|^2, |z'_2|^2 \sim N/Z \\ \Delta' \neq 0 \\ (z'_1, z'_2)=1 \\ (z'_j, z'_j P(W))=1 \\ z'_2 \equiv az'_1 (db_0)}} \frac{1}{|\Delta'|}. \end{aligned}$$

We note that  $z'_2 \equiv az'_1 (db_0)$  implies that for  $z = z_2 \overline{z_1}$ , we have  $z \equiv \overline{z} (db_0)$ , that is, denoting  $z = r + is$ , we get  $s \equiv 0 (db_0)$ . Thus, using the divisor bound

$$\sum_{\substack{z=z_2\overline{z_1} \\ (z_j, P(W))=1}} 1 \ll_{\varepsilon} W^{\varepsilon},$$

we get

$$\begin{aligned} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N/Z \\ \Delta' \neq 0 \\ (z'_1, z'_2)=1 \\ (z'_j, \overline{z'_j} P(W))=1 \\ z'_2 \equiv a z'_1 \pmod{db_0}}} \frac{1}{|\Delta'|} &\ll_{\varepsilon} W^{\varepsilon} \sum_{\substack{|r_1|, |r_2| \ll (N/Z)^{1/2} \\ r_2 \equiv ar_1 \pmod{db_0}}} \sum_{\substack{0 < |s| \ll N/Z \\ s \equiv 0 \pmod{db_0}}} \frac{1}{s} \\ &\ll_{\varepsilon} W^{\varepsilon} \left( \frac{N}{Z d^2 b_0^2} + \frac{N^{1/2}}{Z^{1/2} db_0} \right) \\ &\ll_{\varepsilon} W^{\varepsilon} \frac{N}{W d^2 b_0^2}, \end{aligned} \quad (7.15)$$

where the last step follows from  $N > X^{3\delta+3\eta}$  and  $db_0 \leq X^{\delta+\eta}$ . Therefore, we get by Lemma 2.1,

$$\begin{aligned} V_{Z, \leq}(\beta) &\ll W^{\varepsilon} \frac{N}{W} \sum_{db_0 \leq X^{\delta+\eta}} \sum_{\substack{b_0 b'_1, b_0 b'_2 \in B \\ (b'_1, b'_2)=1 \\ (b'_1 b'_2, d)=1}} \frac{1}{d^2 b_0} \sum_{a \pmod{db_0}} (d, b'_2 - ab'_1) \\ &\ll_{\varepsilon} W^{\varepsilon} \frac{N|B|}{W}, \end{aligned}$$

which completes the proof of (7.13).

## 7.5 Proof of Lemma 7.4

By recombining the finer-than-dyadic partitions for  $\theta_1, \theta_2$ , we get

$$V_{\leq \nu_3}(\beta) \leq U_{\ll \nu_3}(\beta),$$

with

$$U_{\ll \nu_3}(\beta) := \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ 0 < |\Delta| \ll \nu_3 N \\ (z_1, z_2)=1}} |\beta_{z_1}^b \overline{\beta_{z_2}^b}| \sum_{(w, \overline{w})=1} F_M(|w|^2) \mathbf{1}_B(\operatorname{Re}(\overline{w} z_1)) \mathbf{1}_B(\operatorname{Re}(\overline{w} z_2)).$$

Similarly as in § 7.2, we write

$$U_{\ll \nu_3}(\beta) \ll (\log X)^{O(1)} \sum_{b_0} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ 0 < |\Delta| \ll \nu_3 N \\ (z_1, z_2)=1 \\ (z'_j, \overline{z'_j})=1 \\ \Delta \equiv 0 \pmod{b_0}}} \sum_{\substack{b'_2 \equiv ab'_1 \pmod{\Delta/b_0} \\ (b'_1 b'_2, \Delta/b_0)=1 \\ (b'_1, b'_2)=1}} \mathbf{1}_B(b_0 b'_1) \mathbf{1}_B(b_0 b'_2) F_M \left( \left| b_0 \frac{z_2 b'_1 - z_1 b'_2}{\Delta} \right|^2 \right).$$

Using  $\Delta \ll \nu_2 N$  and  $MN \sim X$  we see that the smooth weight  $F_M$  restricts  $z_2 b_1 - z_1 b_2$  to a small disc, that is,

$$U_{\ll \nu_3}(\beta) \ll \sum_{b_0} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ 0 < |\Delta| \ll \nu_3 N \\ (z_1, z_2)=1 \\ (z'_j, \overline{z'_j})=1 \\ \Delta \equiv 0 \pmod{b_0}}} \sum_{\substack{b'_2 \equiv ab'_1 \pmod{\Delta/b_0} \\ (b'_1 b'_2, \Delta/b_0)=1 \\ (b'_1, b'_2)=1}} \mathbf{1}_B(b_0 b'_1) \mathbf{1}_B(b_0 b'_2) \mathbf{1}_{|z_2 b_0 b'_1 - z_1 b_0 b'_2| \ll \nu_3 N^{1/2} X^{1/2}}.$$



Now recall that by (3.2) we have  $B \subseteq [Y, Y + X^{1/2-\eta}]$  for some  $Y \asymp X^{1/2}$ . Hence, we obtain

$$U_{\ll \nu_3}(\beta) \ll \sum_{b_0} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ |\Delta| \neq 0 \\ (z_1, z_2)=1 \\ (z'_j, z'_j)=1 \\ \Delta \equiv 0 \pmod{b_0}}} \sum_{\substack{b'_2 \equiv ab'_1 \pmod{\Delta/b_0} \\ (b'_1 b'_2, \Delta/b_0)=1 \\ (b'_1, b'_2)=1}} \mathbf{1}_B(b_0 b'_1) \mathbf{1}_B(b_0 b'_2) \mathbf{1}_{|z_2 - z_1| \ll \nu_3 N^{1/2}}.$$

This can now be bounded by the same argument as in §7.4, replacing the bounds (7.14) and (7.15) by

$$\sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ (z_1, z_2)=1 \\ (z_j, \bar{z}_j)=1 \\ \Delta=D}} \mathbf{1}_{|z_2 - z_1| \ll \nu_3 N^{1/2}} \ll_\varepsilon X^\varepsilon \nu_3 N,$$

respectively, and

$$\begin{aligned} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N \\ \Delta \neq 0 \\ (z_1, z_2)=1 \\ z_2 \equiv az_1 \pmod{db_0}}} \frac{\mathbf{1}_{|z_2 - z_1| \ll \nu_3 N^{1/2}}}{|\Delta|} &\ll_\varepsilon X^\varepsilon \sum_{\substack{|r_1|, |r_2| \ll (N/Z)^{1/2} \\ |r_2 - r_1| \ll \nu_3 N^{1/2} \\ r_2 \equiv ar_1 \pmod{db_0}}} \sum_{\substack{0 < |s| \ll \nu_3 N \\ s \equiv 0 \pmod{db_0}}} \frac{1}{s} \\ &\ll_\varepsilon X^\varepsilon \left( \frac{\nu_3 N}{d^2 b_0^2} + \frac{\nu_3 N^{1/2}}{db_0} \right) \\ &\ll_\varepsilon X^\varepsilon \frac{\nu_3 N}{d^2 b_0^2}. \end{aligned}$$

We get

$$U_{\ll \nu_3}(\beta) \ll_\varepsilon X^\varepsilon \nu_3 N |B|^2,$$

which gives us Lemma 7.4.

*Remark 7.7.* Without the assumption  $B \subset [\eta_1 X^{1/2}, (2-\eta)X^{1/2}]$  we would need to take  $\nu_3 = X^{-\delta-\varepsilon}$  to get savings in this argument, since it is possible that  $Y \asymp X^{1/2-\delta}$ . This means that in the approximation for  $\beta_n$  we need to track the distribution of  $\beta_n$  in sectors with an angle  $X^{-\delta-\eta}$ . It is possible to do so by a more careful argument but we do not pursue this issue here. For this, we also note that the smooth weight

$$F_M \left( \left| \frac{b_1 z_2 - b_2 z_1}{\Delta} \right|^2 \right) = F_{M/Y^2} \left( \left| \frac{z_2 - z_1}{\Delta} \right|^2 \right) + O(X^{-\eta'})$$

could be handled more efficiently in terms of the dependency on  $\arg z_j$  since it is a function of the difference  $\arg z_2 - \arg z_1$ , so that we need an expansion to  $\xi_k$  only once instead of twice. Note that  $B$  cannot simultaneously be multiples of a large fixed  $q$  and restricted to a narrow interval. Thus, for this extension the condition  $\mathcal{M}(u) \leq X^{\delta+\eta}$  ought to be replaced by  $|k|\mathcal{M}(u) \leq X^{\delta+\eta}$ .

## 8. Type II information: proof of Proposition 6.8

Recall that we are trying to evaluate

$$\begin{aligned} \sum_{\mathcal{N} \mathbf{m} \sim M} \sum_{\mathcal{N} \mathbf{n} \sim N} \alpha_{\mathbf{m}} \beta_{\mathbf{n}}^\# a_{\mathbf{m} \mathbf{n}} &= \sum_{N'} \sum_{\mathcal{N} \mathbf{m} \sim M} \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \alpha_{\mathbf{m}} \mathbf{1}_{(\mathbf{n}, \bar{\mathbf{n}} P(W))=1} \left( \sum_{j \leq J} \overline{\xi_{k_j} \chi_j}(\mathbf{n}) \mathcal{C}_W(\beta, \overline{\xi_{k_j} \chi_j} H_{N'}) \right) a_{\mathbf{m} \mathbf{n}} \\ &=: \sum_{j \leq J} S_j. \end{aligned}$$

The condition  $(\mathbf{n}, \bar{\mathbf{n}})$  may be dropped since we have  $(z, \bar{z}) = 1$  in the definition of  $a_z$ . We have

$$\begin{aligned} S_j &= \sum_{N'} \mathcal{C}_W(\beta, \overline{\xi_{k_j} \chi_j} H_{N'}) \sum_{\mathcal{N}\mathbf{m} \sim M} \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \alpha_{\mathbf{m}} \mathbf{1}_{(\mathbf{n}, P(W))=1} \overline{\xi_{k_j} \chi_j}(\mathbf{n}) a_{\mathbf{m}\mathbf{n}} \\ &= \sum_{N'} \mathcal{C}_W(\beta, \overline{\xi_{k_j} \chi_j} H_{N'}) \sum_{\mathcal{N}\mathbf{m} \sim M} \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \alpha_{\mathbf{m}} \xi_{k_j} \chi_j(\mathbf{m}) \mathbf{1}_{(\mathbf{n}, P(W))=1} \overline{\xi_{k_j} \chi_j}(\mathbf{m}\mathbf{n}) a_{\mathbf{m}\mathbf{n}}. \end{aligned}$$

Here, for any  $C > 0$ ,

$$\begin{aligned} \mathcal{C}_W(\beta, \overline{\xi_{k_j} \chi_j} H_{N'}) &= \left( \sum_{(\mathbf{n}, \bar{\mathbf{n}})=1} H_{N'}(\mathbf{n}) \beta_{\mathbf{n}} \xi_{k_j} \chi_j(\mathbf{n}) \right) \left( \sum_{(\mathbf{n}, \bar{\mathbf{n}} P(W))=1} H_{N'}(\mathbf{n}) |\xi_{k_j} \chi_j(\mathbf{n})| \right)^{-1} \\ &= \frac{1}{\nu_2} \prod_{p \leq W} \left( 1 - \frac{\rho(p)}{p} \right)^{-1} \sum_{\mathbf{n}} H_{N'}(\mathbf{n}) \frac{\beta_{\mathbf{n}} \xi_{k_j} \chi_j(\mathbf{n})}{\mathcal{N}\mathbf{n}} + O_C((\log X)^{-C}). \end{aligned}$$

Then Proposition 4.1 follows from applying the fundamental lemma of the sieve (Lemma 2.5, see also Remark 2.6) and Proposition 4.1 to handle  $\mathbf{1}_{(\mathbf{n}, P(W))=1}$  in  $S_j$ . Note that by  $N > X^{3\delta+3\eta}$  and  $\mathcal{M}(u_j) \leq Q \leq X^{\delta+\eta}$ , we get that  $\mathcal{N}\mathbf{m} \ll X^{1-\delta-\eta}/\mathcal{M}(u_j)^2$ , which is required for Proposition 4.1. This gives us a main term of the form

$$\sum_{N'} \sum_{j \leq J} \sum_{\mathbf{m}, \mathbf{n}} \frac{\alpha_{\mathbf{m}} \beta_{\mathbf{n}} H_{N'}(\mathbf{n}) \xi_{k_j} \chi_j(\mathbf{m}\mathbf{n})}{\mathcal{N}\mathbf{m}\mathbf{n}} \sum_{\mathbf{a}} \frac{H_{N'}(\mathbf{a}/\mathbf{m})}{\nu_2} a_{\mathbf{a}}^{\omega} \overline{\xi_{k_j} \chi_j}(\mathbf{a}).$$

The weight  $\frac{H_{N'}(\mathbf{a}/\mathbf{m})}{\nu_2}$  may be replaced (with a negligible error term) by  $\frac{F(\mathcal{N}\mathbf{m}\mathbf{n}/\mathcal{N}\mathbf{a})}{\hat{F}(0)}$  by a further application of Poisson summation (Lemma 2.7) on the free variable  $a$  in  $z = b + ia$ , completing the proof.

*Remark 8.1.* There are two potential bottlenecks for improving the range of  $\delta < 1/10$  in Theorem 3.3, namely, the exponent 3 in Lemma 6.1 and the exponent 2 in  $X^{1-\delta-\eta}/q^2$  in Proposition 4.1. It is plausible that with more work these exponents may be improved to 2 and 1, respectively, which would suffice to prove Theorem 3.3 for  $\delta < 1/8$ . Both of these improvements run into quite delicate issues and we have decided not to pursue this here. It is not clear whether the boundary  $1/8$  can be improved, but we certainly hit a hard barrier at  $\delta = 1/6$  as this when even the most optimistic the Type II range  $[N^{2\delta}, N^{1/2-\delta}]$  becomes empty.

## 9. Proof of Theorem 3.3

We apply a sieve argument to the sequence  $\mathcal{A} = (a_n)$  over integers

$$a_n := F(n/X') \sum_{\mathcal{N}\mathbf{n}=n} a_{\mathbf{n}},$$

where for convenience we have split  $n$  into finer-than-dyadic intervals (as in § 2.1) with  $\nu = (\log X)^{-C}$  and  $X' \sim X$ . We also define an auxiliary sequence  $\mathcal{B}_j$  by

$$b_n^{(j)} := \sum_{\mathcal{N}\mathbf{n}=n} \sum_{\mathbf{m}} \frac{F(\mathcal{N}\mathbf{n}/\mathcal{N}\mathbf{m}) \xi_{k_j} \chi_j(\mathbf{n})}{\hat{F}(0) \mathcal{N}\mathbf{n}} F(\mathcal{N}\mathbf{m}/X') a_{\mathbf{m}}^{\omega} \overline{\xi_{k_j} \chi_j}(\mathbf{m}),$$

denoting  $\xi_{k_0} \chi_0 = 1$  for  $j = 0$ . Then Theorem 3.3 follows by using the explicit formula to evaluate the sums

$$\sum_{\mathbf{n}} \Lambda(\mathcal{N}\mathbf{n}) \frac{F(\mathcal{N}\mathbf{n}/\mathcal{N}\mathbf{m}) \xi_{k_j} \chi_j(\mathbf{n})}{\hat{F}(0) \mathcal{N}\mathbf{n}},$$

once we prove that for any  $C_1 > 0$ , there is some  $C_2 > 0$  such that

$$S(\mathcal{A}, \Lambda) = \sum_n a_n \Lambda(n) = \sum_{0 \leq j \leq J_1} S(\mathcal{B}_j, \Lambda) + O\left(\frac{X^{1/2}|B|}{(\log X)^{C_1}}\right). \quad (9.1)$$

Let  $Y = X^{3\delta+4\eta}$  and  $Z = X^{1/2+\delta+\eta}$ . Then  $YZ \leq X^{1-\delta-\eta}$  for some  $\eta > 0$  by  $\delta < 1/10$ . By Vaughan's identity [Vau75] for  $n > Y$  we have

$$\Lambda(n) = \sum_{\substack{b|n \\ b \leq Y}} \mu(b) \log \frac{n}{b} - \sum_{\substack{bc|n \\ b \leq Y \\ c \leq Z}} \mu(b) \Lambda(c) + \sum_{\substack{bc|n \\ b > Y \\ c > Z}} \mu(b) \Lambda(c). \quad (9.2)$$

Applying this with both sides multiplied by  $(n, P(W)) = 1$  we get

$$S(\mathcal{A}, \Lambda) = S_1(\mathcal{A}) + S_2(\mathcal{A}) + S_3(\mathcal{A}) + O\left(\frac{X^{1/2}|B|}{(\log X)^{C_1}}\right)$$

and, similarly, for  $j \leq J$  write

$$S(\mathcal{B}_j, \Lambda) = S_1(\mathcal{B}_j) + S_2(\mathcal{B}_j) + S_3(\mathcal{B}_j) + O\left(\frac{X^{1/2}|B|}{(\log X)^{C_1}}\right).$$

The sums  $S_1, S_2$  correspond to Type I sums and  $S_3$  is a Type II sum, with all variables coprime to  $P(W)$ .

By the fundamental lemma of the sieve (Lemma 2.5) and Type I information (Proposition 4.1), for  $k = 1, 2$  we get

$$S_k(\mathcal{A}) = S_k(\mathcal{B}_0) + O\left(\frac{X^{1/2}|B|}{(\log X)^{C_1}}\right) = \sum_{0 \leq j \leq J_1} S_k(\mathcal{B}_j) + O\left(\frac{X^{1/2}|B|}{(\log X)^{C_1}}\right),$$

since for  $j \geq 1$ ,  $k = 1, 2$  for any  $C > 0$ ,

$$S_k(\mathcal{B}_j) \ll_C X^{1/2}|B|(\log X)^{-C}.$$

By Type II information (Proposition 6.2, note that  $Y < b \ll X/Z$ ) we get

$$S_3(\mathcal{A}) = \sum_{1 \leq j \leq J_j} S_3(\mathcal{B}_j) + O\left(\frac{X^{1/2}|B|}{(\log X)^{C_1}}\right) = \sum_{0 \leq j \leq J_j} S_3(\mathcal{B}_j) + O\left(\frac{X^{1/2}|B|}{(\log X)^{C_1}}\right),$$

since for any  $C > 0$ ,

$$S_3(\mathcal{B}_0) \ll_C X^{1/2}|B|(\log X)^{-C}.$$

By recombining Vaughan's identity for  $\mathcal{B}_j$  we get (9.1).

## 10. Proof of Theorem 3.2

We have two cases, no zeros  $\beta > 1 - \varepsilon_1 / \log X$  or that in the case of a zero  $\beta > 1 - \varepsilon_1 / \log X$  we have  $\Omega(B_1) \leq \Omega(B)/2$ .

### 10.1 No zeros $\beta > 1 - \varepsilon_1 / \log X$

Let us denote by  $a_n, a_n^\omega$  the sequences corresponding to  $\lambda = \mathbf{1}_B$ . By Theorem 3.3 we have

$$\sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}} \Lambda(\mathcal{N}\mathbf{n}) = \frac{4}{\pi} \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega} \left( 1 - \sum_{j \leq J} \overline{\xi_{k_j}} \chi_j(\mathbf{n}) \sum_{\substack{\rho_j \\ L(\rho_j, \xi_{k_j} \chi_j) = 0 \\ |\operatorname{Im}(\rho_j)| \leq X^{\eta}}} (\mathcal{N}\mathbf{n})^{\rho_j - 1} \right) + O\left(\frac{1}{(\log X)^{C_1}} X^{1/2} \sum_b |\lambda_b|\right). \quad (10.1)$$

If there is a zero  $\beta_1 \geq 1 - \frac{1}{\sqrt{\delta} \log X}$  as in Lemma 2.15 corresponding to  $\chi_1$  real and  $\xi_{k_1} = 1$ , then  $\beta_1 \leq 1 - \varepsilon_1 / \log X$  and the contribution from that zero is

$$\begin{aligned} -\frac{4}{\pi} \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega} \chi_1(\mathbf{n}) (\mathcal{N}\mathbf{n})^{\beta_1 - 1} &\geq -\frac{4}{\pi} \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega} (\exp(-\varepsilon_1) + o(1)) \\ &\geq -\frac{4}{\pi} \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega} (1 - \varepsilon_1/2), \end{aligned}$$

since  $\varepsilon_1 < 1/10$ . Therefore, the contribution from the first two terms in (10.1) is

$$\frac{4}{\pi} \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega} (1 - \chi_1(\mathbf{n}) (\mathcal{N}\mathbf{n})^{\beta_1 - 1}) \gg \varepsilon_1 \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega}.$$

Denoting  $\sigma_0 = 1 - \frac{1}{\sqrt{\delta} \log X}$ , by Lemma 2.15 the remaining zeros satisfy  $\beta_j \leq \sigma_0$  and they contribute at most

$$\begin{aligned} &\ll \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega} \sum_{j \leq J} \sum_{\substack{\rho_j \\ L(\rho_j, \xi_{k_j} \chi_j) = 0 \\ |\operatorname{Im}(\rho_j)| \leq X^{\eta} \\ \beta_j \leq \sigma_0}} X^{\beta_j - 1} \\ &= \Omega(B) \sum_{j \leq J} \sum_{\substack{\rho_j \\ L(\rho_j, \xi_{k_j} \chi_j) = 0 \\ |\operatorname{Im}(\rho_j)| \leq X^{\eta} \\ \beta_j \leq \sigma_0}} \left( \int_{1/2}^{\beta_j} X^{\sigma - 1} \log X d\sigma + X^{-1/2} \right) \\ &\ll \Omega(B) \int_{1/2}^{\sigma_0} N^*(\sigma, X^{\eta}, X^{\eta}, X^{2\delta + 2\eta}) X^{\sigma - 1} \log X d\sigma + \Omega(B) X^{-1/2 + O(\delta + \eta)}. \end{aligned}$$

The last term is negligible once  $\delta, \eta$  are sufficiently small. By Lemma 2.14 the integral is bounded by

$$\begin{aligned} &\int_{1/2}^{\sigma_0} N^*(\sigma, X^{\eta}, X^{\eta}, X^{2\delta + 2\eta}) X^{\sigma - 1} \log X d\sigma \ll \int_{1/2}^{\sigma_0} X^{c_2(4\delta + 6\eta)(1 - \sigma) + \sigma - 1} \log X d\sigma \\ &\ll X^{(\sigma_0 - 1)/2} = \exp\left(-\frac{1}{2\sqrt{\delta}}\right), \end{aligned}$$

once  $\delta, \eta$  are sufficiently small compared with the constant  $c_2$  in Lemma 2.14. Combining all of the above estimates we have

$$\sum_{\mathcal{N}\mathbf{p}\sim X} a_{\mathbf{p}} \gg \frac{1}{\log X} \left( \varepsilon_1 - O\left(\exp\left(-\frac{1}{2\sqrt{\delta}}\right)\right) \right) \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega} \gg \frac{1}{\log X} \varepsilon_1 \sum_{\mathcal{N}\mathbf{n}\sim X} a_{\mathbf{n}}^{\omega}$$

once  $\delta$  is small enough compared with  $\varepsilon_1$ .

**10.2 Zero  $\beta > 1 - \varepsilon_1 / \log X$  and  $\Omega(B_1) \leq \Omega(B)/2$** 

Let us call  $B_2 = B \setminus B_1$  so that  $\Omega(B_2) \geq \Omega(B)/2$  and for all  $b \in B_2$ , we have

$$\sum_{\substack{a \sim (X-b^2)^{1/2} \\ (a,b)=1 \\ (a^2+b^2,2)=1}} \chi_1((b+ia)) \leq 0. \quad (10.2)$$

Let us denote by  $a_n, a_n^\omega$  the sequences corresponding to  $\lambda = \mathbf{1}_B$  and  $a_n^{(2)}, a_n^{(2)\omega}$  the sequences corresponding to  $\lambda = \mathbf{1}_{B_2}$ . Then, by Theorem 3.3, we have

$$\begin{aligned} (\log X) \sum_{\mathcal{N} \mathbf{p} \sim X} a_{\mathbf{p}} &\geq \log(X) \sum_{\mathcal{N} \mathbf{p} \sim X} a_{\mathbf{p}}^{(2)} \\ &\gg \sum_{\mathcal{N} \mathbf{n} \sim X} a_{\mathbf{n}}^{(2)} \Lambda(\mathcal{N} \mathbf{n}) = \frac{4}{\pi} \sum_{\mathcal{N} \mathbf{n} \sim X} a_{\mathbf{n}}^{(2)\omega} \left( 1 - \sum_{j \leq J} \overline{\xi_{k_j} \chi_j}(\mathbf{n}) \sum_{\substack{\rho_j \\ L(\rho_j, \xi_{k_j} \chi_j) = 0 \\ |\operatorname{Im}(\rho_j)| \leq X^\eta}} (\mathcal{N} \mathbf{n})^{\rho_j - 1} \right) \\ &\quad + O\left(\frac{1}{(\log X)^{C_1}} X^{1/2} \sum_b |\lambda_b|\right). \end{aligned}$$

The first term contributes

$$\sum_{\mathcal{N} \mathbf{n} \sim X} a_{\mathbf{n}}^{(2)\omega} = \Omega(B_2) \geq \frac{1}{2} \Omega(B) = \frac{1}{2} \sum_{\mathcal{N} \mathbf{n} \sim X} a_{\mathbf{n}}^\omega.$$

The contribution from  $2 \leq j \leq J$  is handled similarly as in § 10.1 and similarly for  $j = 1$  the zeros  $\beta \leq 1 - \varepsilon_1 / \log X$ . The contribution from the Siegel zero  $\beta > 1 - \varepsilon_1 / \log X$  for  $j = 1$  is essentially positive, since  $\xi_{k_1} = 1$ ,  $\chi_1$  is real, and by (10.2)

$$\begin{aligned} - \sum_{\mathcal{N} \mathbf{n} \sim X} a_{\mathbf{n}}^{(2)\omega} \chi_1(\mathbf{n}) (\mathcal{N} \mathbf{n})^{\beta-1} &= - \sum_{b \in B_2} \sum_{\substack{a \sim (X-b^2)^{1/2} \\ (a,b)=1 \\ (a^2+b^2,2)=1}} \chi_1((b+ia)) (a^2+b^2)^{\beta-1} \\ &\geq - \sum_{b \in B_2} \sum_{\substack{a \sim (X-b^2)^{1/2} \\ (a,b)=1 \\ (a^2+b^2,2)=1}} (\chi_1((b+ia)) + 2\varepsilon_1) \\ &\geq -2\varepsilon_1 \Omega(B_2) \geq -2\varepsilon_1 \sum_{\mathcal{N} \mathbf{n} \sim X} a_{\mathbf{n}}^\omega, \end{aligned}$$

since  $\varepsilon_1 < 1/10$ . Therefore, we conclude that also in the second case

$$\sum_{\mathcal{N} \mathbf{p} \sim X} a_{\mathbf{p}} \gg \frac{1}{\log X} \left( 1 - 2\varepsilon_1 O\left(\exp\left(-\frac{1}{2\sqrt{\delta}}\right)\right) \right) \sum_{\mathcal{N} \mathbf{n} \sim X} a_{\mathbf{n}}^\omega \gg \sum_{\mathcal{N} \mathbf{n} \sim X} a_{\mathbf{n}}^\omega$$

once  $\delta$  is sufficiently small.

**11. Proof of Theorem 1.6**

By similar reductions as in § 3 it suffices to consider  $\lambda_b = \mathbf{1}_B(b)$ . The goal is to show that if  $u_j$  is a modulus of one of the characters  $\xi_{k_j} \chi_j$  and  $b$  does not have a large common factor with  $u_j$ , then the sum over the free variable  $a$  of  $\xi_{k_j} \chi_j(b+ia)$  exhibits cancellation. By Theorem 3.3 and the Siegel–Walfisz bound [FI98b, Lemma 16.1] for small moduli  $u_j$ , it suffices to show that for

any  $j \leq J_1$  and  $Y \in [X^{1/2-\eta}, 2X^{1/2}]$  and any  $|u_j|^2 \gg (\log X)^{C''}$  with  $C''$  large compared with  $C'$  we have

$$S'_j := \sum_{b \in B} \left| \sum_{\substack{a \in (Y, Y(1+X^{-\eta})] \\ (a,b)=1 \\ a \equiv a_0 \pmod{4}}} \chi_j(b+ia) \right| \ll \frac{X^{-\eta} Y |B|}{(\log X)^C}.$$

Note that the weight  $\xi_j((b+ia))$  has been removed by splitting  $a$  into finer-than-dyadic intervals and using (3.2) to note that then  $b+ia$  lives in a small box.

We write

$$S'_j = \sum_{v|u_j} \sum_{\substack{b \in B \\ (b,u_1)=v}} \left| \sum_{\substack{a \in (Y, Y(1+X^{-\eta})] \\ (a,b)=1 \\ a \equiv a_0 \pmod{4}}} \chi_j(b+ia) \right|.$$

For  $|v| > |u_j|/(\log X)^{C'/2}$  we use the assumption (1.1) to get

$$\ll \frac{X^{-\eta} Y |B|}{(\log X)^{C'}} \sum_{\substack{v|u_j \\ |v| > |u_1|/(\log X)^{C'/2}}} 1 = \frac{X^{-\eta} Y |B|}{(\log X)^{C'}} \sum_{\substack{v|u_j \\ |v| \leq (\log X)^{C'/2}}} 1 \ll \frac{X^{-\eta} Y |B|}{(\log X)^{C+C_2}},$$

once  $C'$  is large compared with  $C_1$  and  $C$ . For  $|v| \leq |u_j|/(\log X)^{C'/2}$  we use Lemma 2.20 to get

$$S_{>} \ll X^{-\eta} Y \sum_{\substack{v|u_j \\ |v| \leq |u_1|/(\log X)^{C'/2}}} \frac{1}{|u/v|^{1/3}} \sum_{\substack{b \in B \\ (b,u_1)=v}} 1 \ll \frac{X^{-\eta} Y |B|}{(\log X)^{C+C_2}},$$

by taking  $C'$  large compared with  $C_1$  and  $C$ . To evaluate the main term we note that

$$\sum_{\substack{a^2+b^2 \sim X \\ (a,b)=1 \\ (b+ia,2)=1}} 2\omega_2(b) = (1 + O(X^{-\eta})) \sum_{\substack{a^2+b^2 \sim X \\ (a,b)=1}} \omega(b).$$

## 12. Proof of Theorem 1.10

This follows immediately from Theorem 3.3 with the zero-density estimate Lemma 2.14 once  $C'$  is sufficiently large, via similar arguments as in § 10.1.

### ACKNOWLEDGEMENTS

I am grateful to Akshat Mudgal for numerous discussions and to Lasse Grimmelt for suggestions on constructing an approximation for primes, as well as to James Maynard for encouragement and helpful comments. I also wish to thank the anonymous referee for comments.

### CONFLICTS OF INTEREST

None.

### FINANCIAL SUPPORT

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 851318).

## JOURNAL INFORMATION

*Compositio Mathematica* is owned by the Foundation Compositio Mathematica and published by the London Mathematical Society in partnership with Cambridge University Press. All surplus income from the publication of *Compositio Mathematica* is returned to mathematics and higher education through the charitable activities of the Foundation, the London Mathematical Society and Cambridge University Press.

## REFERENCES

- Col90 M. D. Coleman, *A zero-free region for the Hecke  $L$ -functions*, *Mathematika* **37** (1990), 287–304.
- Dra15 S. Drappeau, *Théorèmes de type Fouvry–Iwaniec pour les entiers friables*, *Compos. Math.* **151** (2015), 828–862.
- FI97 E. Fouvry and H. Iwaniec, *Gaussian primes*, *Acta Arith.* **79** (1997), 249–287.
- FI98a J. Friedlander and H. Iwaniec, *Asymptotic sieve for primes*, *Ann. of Math. (2)* **148** (1998), 1041–1065.
- FI98b J. Friedlander and H. Iwaniec, *The polynomial  $X^2 + Y^4$  captures its primes*, *Ann. of Math. (2)* **148** (1998), 945–1040.
- FI05 J. B. Friedlander and H. Iwaniec, *The illusory sieve*, *Int. J. Number Theory* **1** (2005), 459–494.
- FI10 J. Friedlander and H. Iwaniec, *Opera de cribro*, in *American Mathematical Society Colloquium Publications*, vol. 57 (American Mathematical Society, Providence, RI, 2010).
- Gal70 P. X. Gallagher, *A large sieve density estimate near  $\sigma = 1$* , *Invent. Math.* **11** (1970), 329–339.
- Har07 G. Harman, *Prime-detecting sieves*, London Mathematical Society Monographs Series, vol. 33 (Princeton University Press, Princeton, NJ, 2007).
- HB92 D. R. Heath-Brown, *Zero-free regions for Dirichlet  $L$ -functions, and the least prime in an arithmetic progression*, *Proc. Lond. Math. Soc. (3)* **s3-64** (1992), 265–338.
- HB01 D. R. Heath-Brown, *Primes represented by  $x^3 + 2y^3$* , *Acta Math.* **186** (2001), 1–84.
- HL17 D. R. Heath-Brown and X. Li, *Prime values of  $a^2 + p^4$* , *Invent. Math.* **208** (2017), 441–499.
- Hux74 M. N. Huxley, *Large values of Dirichlet polynomials. III*, *Acta Arith.* **26** (1974/75), 435–444.
- IK04 H. Iwaniec and E. Kowalski, *Analytic number theory*, in *American Mathematical Society Colloquium Publications*, vol. 53 (American Mathematical Society, Providence, RI, 2004).
- LSX20 P. C.-H. Lam, D. Schindler and S. Y. Xiao, *On prime values of binary quadratic forms with a thin variable*, *J. Lond. Math. Soc. (2)* **102** (2020), 749–772.
- Li21 X. Li, *Prime values of a sparse polynomial sequence*, *Duke Math. J.* **171** (2022), 101–208.
- Lin44 U. V. Linnik, *On the least prime in an arithmetic progression. I. The basic theorem*, *Mat. Sb.* **15** (1944), 139–178.
- May20 J. Maynard, *Primes represented by incomplete norm forms*, *Forum Math. Pi* **8** (2020), e3.
- Mer22 J. Merikoski, *The polynomials  $X^2 + (Y^2 + 1)^2$  and  $X^2 + (Y^3 + Z^3)^2$  also capture their primes*, *Proc. Lond. Math. Soc.* **127** (2023), 1057–1133.
- MV75 H. L. Montgomery and R. C. Vaughan, *The exceptional set in Goldbach’s problem*, *Acta Arith.* **27** (1975), 353–370.

- MV07 H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97 (Cambridge University Press, Cambridge, 2007).
- Pra20 K. Pratt, *Primes from sums of two squares and missing digits*, Proc. Lond. Math. Soc. (3) **120** (2020), 770–830.
- Ten15 G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Graduate Studies in Mathematics, vol. 163, third edition (American Mathematical Society, Providence, RI, 2015), translated from the 2008 French edition by P. D. F. Ion.
- Vau75 R. C. Vaughan, *Mean value theorems in prime number theory*, J. Lond. Math. Soc. (2) **s2-10** (1975), 153–162.

Jori Merikoski [jori.merikoski@maths.ox.ac.uk](mailto:jori.merikoski@maths.ox.ac.uk)

Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK