THE POLYNOMIALS $X^2 + (Y^2 + 1)^2$ AND $X^2 + (Y^3 + Z^3)^2$ ALSO CAPTURE THEIR PRIMES

JORI MERIKOSKI

Abstract. We show that there are infinitely many primes of the form $X^2 + (Y^2 + 1)^2$ and $X^2 + (Y^3 + Z^3)^2$. This extends the work of Friedlander and Iwaniec showing that there are infinitely many primes of the form $X^2 + Y^4$. More precisely, Friedlander and Iwaniec obtained an asymptotic formula for the number of primes of this form. For the sequences $X^2 + (Y^2 + 1)^2$ and $X^2 + (Y^3 + Z^3)^2$ we establish Type II information that is too narrow for an asymptotic formula, but we can use Harman’s sieve method to produce a lower bound of the correct order of magnitude for primes of form $X^2 + (Y^2 + 1)^2$ and $X^2 + (Y^3 + Z^3)^2$. Estimating the Type II sums is reduced to a counting problem which is solved by using the Weil bound, where the arithmetic input is quite different from the work of Friedlander and Iwaniec for $X^2 + Y^4$. We also show that there are infinitely many primes $p = X^2 + Y^2$ where $Y$ is represented by an incomplete norm form of degree $k$ with $k - 1$ variables. For this we require a Deligne-type bound for correlations of hyper-Kloosterman sums.

Contents

1. Introduction 1
2. The set-up for Theorems 1 and 2 6
3. Type I information 7
4. Type II information: initial reductions 14
5. Applications of the Weil bound 17
6. Proof of Proposition 13 23
7. The sieve argument for Theorems 1 and 2 38
8. The set-up for Theorem 3 45
9. A Deligne-type bound 48
10. The arithmetic information for Theorem 3 60
11. The sieve argument for Theorem 3 64
References 68

1. Introduction

Friedlander and Iwaniec [6] famously showed that there are infinitely many primes of the form $a^2 + b^4$. This result is striking as integers of this form are very sparse – the number of integers of the form $a^2 + b^4$ less than $X$ is of order $X^{3/4}$. Prior to this Fouvry and Iwaniec [3] had shown that there are infinitely many primes of the form $a^2 + p^2$ with $p$ also a prime. The result of [6] was extended to primes of the form $a^2 + p^4$ with $p$ also a prime by Heath-Brown and Li [11]. Pratt [18] has shown that there are infinitely

2020 Mathematics Subject Classification. 11N32 primary, 11N36 secondary.
many primes of the form $a^2 + b^2$ with $b$ missing any three fixed digits in its decimal expansion, which is also a very sparse sequence.

Another key result concerning primes in sparse polynomial sets is the result of Heath-Brown of infinitely many primes of the form $a^3 + 2b^3$ \[9\], which has been generalized to binary cubic forms by Heath-Brown and Moroz \[12\] and to general incomplete norm forms by Maynard \[16\]. Recently Li \[15\] showed that there are infinitely many primes of the form $a^3 + 2b^3$ with $b$ relatively small.

The work of Friedlander and Iwaniec for primes of the form $a^2 + b^4$ relies on the factorization over Gaussian integers $a^2 + b^4 = (b^2 + ia)(b^2 - ia)$ and the great regularity in the distribution of squares $b^2$. In particular, the argument fails to capture primes of the form $a^2 + f(b)^2$ for non-homogeneous quadratic polynomials $f(b)$. Our first main result resolves this, with the caveat that instead of an asymptotic formula we get lower and upper bounds of correct order of magnitude for primes of this shape.

**Theorem 1.** There are infinitely many primes of the form $a^2 + (b^2 + 1)^2$. More precisely, we have

$$\sum_{p \leq X} \sum_{p = a^2 + (b^2 + 1)^2} 1 \asymp \frac{X^{3/4}}{\log X}.$$

The methods developed in this paper are quite flexible and work for Gaussian primes with one coordinate given by a polynomial sequence that is not too sparse compared to the degree. We are also able to show the following.

**Theorem 2.** There are infinitely many primes of the form $a^2 + (c^3 + d^3)^2$ with $c, d > 0$. More precisely, we have

$$\sum_{p \leq X} \sum_{c, d > 0} 1 \asymp \frac{X^{5/6}}{\log X}.$$

**Remark 1.** The proof of Theorem 1 is easily generalized to primes of the form $a^2 + f(b)^2$, where $f(b) = b^2 + D$ for any integer $D \neq 0$, with $D = 0$ corresponding to \[6\]. The proof of Theorem 2 as given below generalizes (with minor technical nuisances) to primes of the form $a^2 + (f_1(c) + f_2(d))^2$, where $f_1$ and $f_2$ are polynomials of degree at most 3. It is possible to handle also primes of the form $a^2 + g(c, d)^2$ for a cubic polynomial $g(c, d)$ which could involve mixed terms, but this would require additional work (replacing the Weil bound by the Deligne bound for two dimensional sums to get an analogue of Lemma 21).

Finally, analogous to Theorem 2 we have the following.

**Theorem 3.** Let $k$ be sufficiently large. Let $K/\mathbb{Q}$ be a Galois extension of degree $k$ and fix a basis $\omega_1, \ldots, \omega_k$ for the ring of integers $\mathcal{O}_K$. Define the incomplete form

$$N(b_1, \ldots, b_{k-1}) := N_{K/\mathbb{Q}}(b_1\omega_1 + \cdots + b_{k-1}\omega_{k-1}).$$

Then there are infinitely many primes of the form $p = a^2 + N(b_1, \ldots, b_{k-1})^2$ with $b_i \in [0, p^{1/(2k)}]$. More precisely, we have

$$\sum_{p \leq X} \sum_{p = a^2 + N(b_1, \ldots, b_{k-1})^2} 1 \asymp \frac{X^{1-1/(2k)}}{\log X}.$$
The assumption that $K/Q$ is Galois could be removed with more work. With a lot more precise numerical work it should be possible to give a non-trivial lower bound for all $k \geq 3$. The result we prove actually approaches to an asymptotic formula as $k$ becomes large. The only reason for restricting to a norm form is that we need the fact that the number of representations

$$b = N(b_1, \ldots, b_{k-1})$$

with $b_i \ll b^{1/k}$ is bounded by $\ll \varepsilon b^\varepsilon$, which is not known for general forms. For the diagonal form $b_1^k + \cdots + b_{k-1}^k$ we can obtain a result conditional to such a hypothesis. Working with a norm form also simplifies the bounding of the relevant exponential sums of Deligne type.

1.1. Sketch of the arguments. Let us focus on the case $a^2 + (b^2 + 1)^2$ in this sketch, as the proofs of the other results are similar. Using a sieve argument counting primes is reduced to estimating Type I and Type II sums of the form

$$\sum_{d \leq D} \alpha(d) \sum_{n \sim X \atop n \equiv 0 (d)} \sum_{a^2 + (b^2 + 1)} 1 \quad \text{and} \quad \sum_{m \sim M \atop n \sim N} \alpha(m) \beta(n) \sum_{mn = a^2 + (b^2 + 1)} 1,$$

with $MN = X$ with bounded coefficients $\alpha$ and $\beta$ with $\beta$ satisfying the Siegel-Walfisz property (4.1). We can evaluate the Type I sums in the range $D \leq X^{3/4-\eta}$ by essentially the same argument as in [6, Section 3], applying the large sieve inequality for roots of quadratic congruences.

We will be able to handle the Type II sums in the range

$$X^{1/4+\eta} \ll N \ll X^{1/3-\eta}$$

for any small $\eta > 0$. This is in contrast to FI, where the range is essentially $X^{1/4+\eta} \ll N \ll X^{1/2-\eta}$. Our Type II information is not sufficient for an asymptotic formula, but using Harman’s sieve method [7] we are nonetheless able to get a lower bound of the correct order of magnitude. In contrast to the work of Friedlander and Iwaniec [5], our sieve argument is much more similar to the one in Heath-Brown’s work [9].

In the situation of Theorem 2 we are able to handle Type I sums for $D \leq X^{5/6-\eta}$ and Type II sums in the range $X^{1/6+\eta} \ll N \ll X^{2/9-\eta}$.

For the Type II sums we first use a device of Heath-Brown to reduce to the special case where $\beta(n)$ satisfies the Siegel-Walfisz property with main term 0 (see (4.3)) – this replaces the sieve argument of Friedlander and Iwaniec [5] in our work. It then suffices to show that for $\beta$ satisfying (4.3) we have

$$\sum_{m \sim M \atop n \sim N} \alpha(m) \beta(n) \sum_{mn = a^2 + (b^2 + 1)} 1 \ll C X^{3/4}/\log^C X.$$

The initial part of the argument is the same as in [6, Sections 4 and 5]. Assuming for simplicity that $(m, n) = 1$, we may write

$$m = \lvert w \rvert^2, \quad n = \lvert z \rvert^2, \quad wz = b^2 + 1 + ia$$
for Gaussian integers $w, z \in \mathbb{Z}[i]$, and denote $\alpha_w = \alpha(|w|^2)$ and $\beta_z = \beta(|z|^2)$. Then our Type II sums are essentially

$$\sum_{|w|^2 \sim M} \alpha_w \sum_{|z|^2 \sim N} \beta_z \sum_{\text{Re}(wz) = b^2 + 1} 1.$$  

By applying Cauchy-Schwarz similarly to [6, Sections 4-6], the task is reduced to showing that for all $C > 0$ (for $\psi$ a fixed smooth function)

$$\sum_{z_1, z_2} \beta_{z_1, z_2} \sum_{b_1, b_2} \psi(b_1/B_1) \psi(b_2/B_2) \ll_C X^{1/2} N \log^{-C} X,$$

for any $B_1, B_2 \ll X^{1/4}$, where for $z_j = r_j + is_j$

$$\Delta := \text{Im}(z_1 z_2) \quad \text{and} \quad a := z_2/z_1 = \frac{z_2}{|z_1|^2} = \frac{r_1 r_2 + s_1 s_2}{r_1^2 + s_1^2} (|\Delta|).$$

Essentially the only difference to [6, Section 6] here is that instead of the congruence $b_1^2 \equiv ab_2^2 (\Delta)$ we are counting solutions to

$$b_1^2 + 1 \equiv a(b_2^2 + 1) (\Delta).$$

The contribution from the diagonal $\Delta = 0$ is sufficiently small by trivial bounds provided that $N \gg X^{1/4+\eta}$.

For $\Delta \neq 0$ applying the Poisson summation formula to the sums over $b_1$ and $b_2$ we get a main term which is controlled by

$$N(a; \Delta) := \frac{B_1 B_2}{\Delta^2} |\{ x_1, x_2 (\Delta) : x_1^2 + 1 \equiv a(x_2^2 + 1) (\Delta)\}|,$$

and an error term controlled by an average over the frequencies $(h_1, h_2) \neq (0, 0)$ of the exponential sums

$$S(a, h_1, h_2; \Delta) := \sum_{x_1, x_2 (\Delta) : x_1^2 + 1 \equiv a(x_2^2 + 1) (\Delta)} e_{|\Delta|}(h_1 x_1 + h_2 x_2).$$

This sum runs over a non-singular curve (for the generic $a$), so that by the Weil bound we get (morally)

$$S(a, h_1, h_2; \Delta) \ll |\Delta|^{1/2+\varepsilon} \ll N^{1/2+\varepsilon}.$$  

This bound is sufficient provided that $N \ll X^{1/3-\eta}$. For the analogous sum in [6]

$$S_{FI}(a, h_1, h_2; \Delta) := \sum_{x_1, x_2 (\Delta) : x_1^2 + 1 \equiv a(x_2^2 + 1) (\Delta)} e_{|\Delta|}(h_1 x_1 + h_2 x_2)$$

one morally gets $S_{FI}(a, h_1, h_2; \Delta) \ll N^\varepsilon$, which is why for $a^2 + b^4$ one can handle the Type II sums up to $N \ll X^{1/2-\eta}$.

For the main term we have to evaluate $N(a; \Delta)$. Since we are counting points on a non-singular curve, this is a much easier problem than the corresponding one in [6] which is over a singular curve

$$N_{FI}(a; \Delta) := |\{ x_1, x_2 (\Delta) : x_1^2 + ax_2^2 (\Delta)\}|.$$
For simplicity assume that $\Delta$ is square-free. Writing

$$N(a; p) := p + \varepsilon_p(a), \quad \varepsilon_d := \prod_{p\mid d} \varepsilon_p,$$

we have by the Weil bound $\varepsilon_p(a) \ll p^{1/2}$ (for generic $a$), which gives $\varepsilon_d(a) \ll d^{1/2+\varepsilon}$.

By the Chinese remainder theorem

$$N(a; \Delta) = \prod_{p\mid d} N(a; p) = \prod_{p\mid d} (p + \varepsilon_p(a)) = |\Delta| \sum_{d\mid \Delta} \frac{\varepsilon_d(a)}{d},$$

so that we find (ignoring the condition $(d, \Delta/d) = 1$ for simplicity)

$$B_1 B_2 \sum_{z_1, z_2} \beta_{z_1} \beta_{z_2} \frac{B_1 B_2}{\Delta^2} N(a; \Delta) = B_1^2 B_2^2 \sum_{d\ll N} \sum_{\Delta = 0 (d)} \beta_{z_1} \beta_{z_2} \frac{1}{|\Delta|} \frac{\varepsilon_d(a)}{d}.$$

By using the bound $\varepsilon_d(a) \ll d^{1/2+\varepsilon}$ we can truncate the sum at $d \leq \log^{2C} X$, getting

$$B_1^2 B_2^2 \sum_{d\leq \log^{2C} X} \sum_{\Delta = 0 (d)} \beta_{z_1} \beta_{z_2} \frac{1}{|\Delta|} \frac{\varepsilon_d(a)}{d} + O_C(XN \log^{-C} X),$$

which can now be bounded by the same argument as in \[6, Section 16\], using the Siegel-Walfisz property \[4,5\]. In contrast, Friedlander and Iwaniec \[6\] have a singular curve with

$$N_{FI}(a; p) = p + \left(\frac{a}{p}\right) p,$$

which produces the sum

$$B_1^2 B_2^2 \sum_{d\ll N} \sum_{\Delta = 0 (d)} \beta_{z_1} \beta_{z_2} \frac{1}{|\Delta|} \left(\frac{a}{\Delta}\right).$$

This sum cannot be truncated and for large $d$ they have to show cancellation in the sums over $z_1$ and $z_2$. Thus, for the main term having $(b^2 + 1)^2$ rather than $b^4$ turns out to be a friend rather than an enemy. Similar arguments apply for the sequence $a^2 + (b^3 + c^3)^2$

The article is structured as follows. In Section 3 we state and prove a Type I estimate (Proposition 4). Using this we prove a fundamental lemma of the sieve type result (Proposition 5). In Section 4 we state our Type II information (Proposition 12) and reduce it to a simpler case (Proposition 13). In Section 5 we will state several corollaries of the Weil bound. After this preparation we give the proof of our Type II estimate in Section 6. In Section 7 we apply Harman’s sieve method with the gathered arithmetic information to prove Theorems 1 and 2. In Sections 8, 10, and 11 we give the proof of Theorem 3, where we will assume that the reader is familiar with the techniques from the previous sections. The key estimates are Lemma 32 and Lemma 36. The first gives essentially a square-root bound for the completed exponential sums unless the frequency parameters lie in some bad hyperplanes, as an application of bounds for correlations of hyper-Kloosterman sums \[17\]. The second shows that on average over the frequency parameters are not too often in these bad hyperplanes.
1.2. **Notations.** For functions \( f \) and \( g \) with \( g \geq 0 \), we write \( f \ll g \) or \( f = O(g) \) if there is a constant \( C \) such that \( |f| \leq Cg \). The notation \( f \asymp g \) means \( g \ll f \ll g \). The constant may depend on some parameter, which is indicated in the subscript (e.g. \( \ll_\varepsilon \)). We write \( f = o(g) \) if \( f/g \to 0 \) for large values of the variable. For summation variables we write \( n \sim N \) meaning \( N < n \leq 2N \), that is,

\[
\sum_{n \sim N} f_n = \sum_{N < n \leq 2N} f_n.
\]

We let \( C > 0 \) denote a large constant. In particular, we write \( f \ll C g / \log C X \) meaning that this bound holds for any fixed \( C > 0 \).

1.3. **Acknowledgements.** I am grateful to Kaisa Matomäki and James Maynard for helpful discussions and comments. I am also grateful for the anonymous referee for their careful reading of the manuscript, helpful suggestions, and spotting errors in a previous version of the manuscript. The author was supported by a grant from the Emil Aaltonen foundation. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 851318).

2. **The set-up for Theorems [1] and [2]**

Denote

\[
\Omega(c, d) := 9c^2d^2/(c^3 + d^3).
\]

Then for large \( Y \) we have

\[
\sum_{c^3 + d^3 > Y, c, d > 0} \frac{9c^2d^2}{c^3 + d^3} = (1 + O(Y^{-\eta})) \sum_{c_1 + d_1 > Y, c_1, d_1 > 0} \frac{1}{c_1 + d_1} = (1 + O(Y^{-\eta})) \sum_{n \sim Y} 1 = (1 + O(Y^{-\eta}))Y,
\]

which follows from

\[
\int_0^{Y^{1/3}} \int_0^{(Y-u^3)^{1/3}} \Omega(u, v)dvdu = Y \int_0^1 \int_0^{(1-u^3)^{1/3}} \Omega(u, v)dvdu = Y,
\]

by using (2.2) with \( Y \) and \( 2Y \) and taking the difference. Heuristically \( c_1^{-2/3}/3 \) and \( d_1^{-2/3}/3 \) are the probabilities that \( c_1 \) and \( d_1 \) are perfect cubes, and \( n = c_1 + d_1 \) is the number of representations as a sum. Also, set

\[
\rho_1(d) := |\{\nu(d) : \nu^2 + 1 \equiv 0 (d)\}|, \quad \rho_2(d) := |\{\alpha, \beta(d) : \alpha^3 + \beta^3 \equiv 0 (d)\}|
\]

and define the arithmetic factors (see Remark [2])

\[
\kappa_j := \sum_{c \geq 1} \frac{\mu(c)}{c^2} \left( \sum_{c \geq 1} \frac{\mu(c)\rho_j(c)}{c^{1+j}} \right)^{-1}.
\]
We define the normalized sequences
\[ a_n^{(1)} := \kappa_1 \sum_{n=a^2+(b^2+1)^2}^{n=b>0} 2b \quad \text{and} \quad a_n^{(2)} := \kappa_2 \sum_{n=a^2+(c^3+d^3)^2}^{n=c,d>0} \Omega(c, d). \]

To prove Theorems 1 and 2 we need to show a lower and upper bounds for the sums along primes
\[ \sum_{p \sim X} a_p^{(j)} \quad \text{for} \quad j \in \{1, 2\} \]

and heuristically we expect that both should be asymptotic to
\[ \sum_{p \sim X} b_p \quad \text{for} \quad b_n := \sum_{n=a^2+b^2}^{n=(a,b)=1} 1. \]

We will use Harman’s sieve method, where the basic idea is to try to compare the sum \( \sum_{p \sim X} a_p^{(j)} \) to \( \sum_{p \sim X} b_p \). This means that all of our arithmetic information (Type I and Type II sums) is given in terms of a comparison, see Propositions 4 and 12.

3. Type I information

We have the following two propositions, the latter of which is a corollary to the first. We will prove both of these in this section.

Proposition 4. (Type I estimate). Let \( \eta, \eta' > 0 \) and suppose that \( \eta \) is sufficiently small in terms of \( \eta' \) (\( \eta = \eta'/100 \) suffices). Let \( D_1 \leq X^{3/4-\eta'} \) and \( D_2 \leq X^{5/6-\eta'} \). Then for any bounded coefficients \( \alpha(d) \) we have
\[ \sum_{d \sim D_1, n \sim X/d} \alpha(d)(a_n^{(j)} - b_n) \ll \eta X^{1-\eta}. \]

Proposition 5. (Fundamental lemma of the sieve). Let \( \eta, \eta' > 0 \) and suppose that \( \eta \) is sufficiently small in terms of \( \eta' \). Let \( W := X^{1/\log \log x} \). Let \( D_1 \leq X^{3/4-\eta'} \) and \( D_2 \leq X^{5/6-\eta'} \). Then for any bounded coefficients \( \alpha(d) \) we have for any \( C > 0 \)
\[ \sum_{d \sim D_1, n \sim X/d} \alpha(d)1_{(n, P(W))=1}(a_n^{(j)} - b_n) \ll \eta C X \log^{-C} X. \]

where \( \eta > 0 \) depends on \( \eta' > 0 \).

Remark 2. In the definitions of \( a_n^{(k)} \), and \( b_n \) we have included the coprimality conditions such as \( (a, b^2+1) = 1 \), and \( (a, b) = 1 \) so that the Type I sums match. Without this Proposition 4 would be false. For instance, if \( d = 5 \), then we have
\[
\sum_{a^2+(b^2+1)^2=0}^{a=A} 2b = \sum_{\mu, \nu(5)}^{\mu^2+(\nu^2+1)^2=0} 2b = (1 + o(1)) \sum_{\mu, \nu(5)}^{\mu^2+(\nu^2+1)^2=0} AB \ll 25 \]
\[
= (1 + o(1)) \frac{8}{25} AB
\]
and
\[
\sum_{a \sim A \atop a^2 + b^2 \equiv 0 (5)} 1 = \sum_{\mu, \nu \equiv 0 (5)} \sum_{a \sim A \atop b \sim B} 1 = (1 + o(1)) \sum_{\mu, \nu \equiv 0 (5)} \frac{AB}{25} = (1 + o(1)) \frac{9}{25} AB.
\]

If we replace the modulus 5 by 3, then the first sum is empty while the second is \( \sim AB/9 \). The issue arises from the potential common factors of \( d \) and \( \mu \). For every odd prime divisor \( p \mid (d, \mu) \) there is exactly one \( \nu \) (\( p \)) such that \( p \mid \nu \) but there are either zero or two \( \nu \) for which \( p \mid (\nu^2 + 1) \). Inserting the coprimality conditions we get \( (p, \mu) = 1 \) which removes this problem. To see that we need to normalize by a factor \( \kappa_2 \) consider \( d = 1 \) for which
\[
\sum_{a \sim A \atop (b^2 + 1) \sim B \atop (a, b^2 + 1) = 1} 2b = \sum_e \mu(e) \sum_{a \sim A \atop b \sim B \atop e \mid a \atop e \mid (b^2 + 1)} 2b = (1 + o(1)) \sum_e \mu(e) \rho_2(e) \frac{1}{e^2} \sum_{a \sim A \atop b \sim B \atop (a, b) = 1} 1 = (1 + o(1)) \kappa_2^{-1} \sum 1.
\]

Similar arguments apply for \( k \geq 3 \).

3.1. 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 \( \delta \in (0, 1/10) \) be small (we will use \( \delta = X^{-c} \) or \( \delta = \log^{-C} X \)), and fix a non-negative \( C^\infty \)-smooth function \( \psi \) supported on \([1 - \delta, 1 + \delta]\) and satisfying
\[
|\psi^{(j)}| \ll_j \delta^{-j}, \ j \geq 0 \quad \text{and} \quad \int_{1/2}^{2} \psi(1/t) \frac{dt}{t} = \delta.
\]

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 \( t \mapsto t/n \))
\[
\sum_{n \leq N} f_n = \frac{1}{\delta} \sum_{n \leq N} \int_{1/2}^{2} \psi(1/t) \frac{dt}{t} = \frac{1}{\delta} \sum_{n \leq N} \int_{1/2}^{2} f_n \psi(n/t) \frac{dt}{t} = \frac{1}{\delta} \int_{1/2}^{2N} \sum_{n \leq N} f_n \psi(n/t) \frac{dt}{t},
\]
so that
\[
\left| \sum_{n \leq N} f_n \right| \leq \frac{1}{\delta} \int_{1/2}^{2N} \left| \sum_{n \leq N} f_n \psi(n/t) \right| \frac{dt}{t}.
\]

Hence, at the cost of a factor \( \delta^{-1} \log N \), it suffices to consider sums of the form
\[
\sum_{n \leq N} f_n \psi(n/t)
\]
for \( t \leq 2N \). Naturally, if the original sum is \( \sum_{n \sim N} f_n \) (dyadic \( n \)), then it suffices to consider \( \sum_{n \sim N} f_n \psi(n/t) \) for \( t \asymp N \) at a cost \( \delta^{-1} \). The effect is the same as with the usual smooth partition of unity except that we did not need to explicitly construct the partition functions \( \psi \). We will denote \( \psi_t(n) := \psi(n/t) \).
3.2. **Lemmas.** For the proof we need the large sieve inequality for roots of quadratic congruences (see [8, Lemma 3.2]).

**Lemma 6.** For any complex numbers $\gamma_n$ we have

$$\sum_{d \leq D} \sum_{\nu^2 + 1 \equiv 0 (d)} \left| \sum_{n \leq N} \gamma_n e_d(\nu n) \right|^2 \ll (D + N) \sum_{n \leq N} |\gamma_n|^2.$$  

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

**Lemma 7.** (Poisson summation). Let $\psi$ be as in Section 3.1 for some $\delta \in (0, 1/10)$ and denote $\psi_N(n) := \psi(n/N)$. Let $x \gg 1$ and let $q \sim Q$ be an integer. Let $\varepsilon > 0$ and denote $H := \delta^{-1}(QN)^\varepsilon Q/N$

Then for any $A > 0$

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

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

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

$$\sum_{n \equiv a (q)} \psi_N(n) = \frac{N}{q} \sum_{h} \hat{\psi} \left( \frac{hN}{q} \right) e_q(-ah)$$

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

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

which gives the result. 

We need the following lemma bounding the number of representations of an integer as a sum of two cubes.

**Lemma 8.** The number of representations of $n$ as sum of two cubes $c^3 + d^3$ with $c, d > 0$ is bounded by $\ll \sqrt{n}$. 

**Proof.** Let $n = c^3 + d^3$. By factoring out the common factor it suffices to consider representations where $(c, d) = 1$ and $(d, n) = 1$. Any such representation produces a root of the equation $\omega^3 + 1 \equiv 0 (n)$ by $\omega = cd$ where $dd \equiv 1 (n)$. If we have another representation $n = c_1^3 + d_1^3$ with $\omega = c_1d_1$, then

$$cd_1 \equiv c_1d \ (n).$$

Since $cd_1, c_1d < c^3 + d^3 = n$, this implies $cd_1 = c_1d$. Since $(c, d) = 1$ and $c, c_1, d, d_1 > 0$, we must have $c = c_1$ and $d = d_1$. The lemma now follows since the number of roots $\omega$ is by the Chinese remainder theorem at most $\ll \sqrt{n}$. The claim also follows trivially from the factorization $c^3 + d^3 = (c + d)(c^2 - cd + d^2)$ and the divisor bound. 

In the proof of Proposition 5 we use the following standard bound for exceptionally smooth numbers (see [19, Chapter III.5, Theorem 1], for instance).
Lemma 9. For any \(2 \leq Z \leq Y\) we have
\[
\sum_{n \sim Y \atop P^*(n) < Z} 1 \ll Y e^{-u/2},
\]
where \(u := \log Y / \log Z\).

We also require the following elementary bound (see [5, Lemma 1], for instance).

Lemma 10. For every square-free integer \(n\) and every \(k \geq 1\) 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 11. 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 \(b_1, \ldots, b_{k-1}\) 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 [10] 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
\[
d \leq (b_1 \cdots b_{k-1})^{1/k} b_k \leq (b_1 \cdots b_{k-1} b_k^{1/k})^{1/k} \leq n^{1/k}
\]
and
\[
\tau(n) \leq (2 \tau(d_1) \cdots \tau(d_{k-1}) \tau(b_k))^{k^2} \leq 2^{k^2} \tau(d)^{k^3}.
\]

\[\square\]

3.3. Proof of Proposition 4. In this section all implicit constants depend on the parameter \(\eta\), that is, \(O\) and \(\ll\) stand for \(O_\eta\) and \(\ll_\eta\) throughout this section. Let us first consider the case \(j = 1\), and denote
\[
\Delta(b) := 1 - \kappa_1 \sum_{b=b_0^2+1} 2|b_0|.
\]

Then the claim is that for \(D_1 \leq X^{3/4-\eta'}\) we have for some \(\eta > 0\)
\[
\sum_{d \sim D_1} \alpha_d \sum_{a^2+b^2 \sim X \atop a^2+b^2 \equiv 0 (d)} \Delta(b) \ll X^{1-\eta}.
\]

We split the sums over \(a\) and \(b\) into short sums using Section 3.1 with \(\delta = X^{-\varepsilon}\) with \(\varepsilon = 4\eta\). Letting \(A, B \leq 4X^{1/2}\) with \(A^2 + B^2 \sim X\), we consider
\[
\sum_{d \sim D_1} \alpha_d \sum_{a^2+b^2 \sim X \atop a^2+b^2 \equiv 0 (d)} \Delta(b) \psi_A(a) \psi_B(b),
\]
which is trivially bounded by $\ll \delta^{-1} ABX^{\varepsilon}$ (using $\tau(a^2 + b^2) \ll X^\varepsilon$). The contribution from $A \leq X^{1/2-\eta-2\varepsilon}$ or $B \leq X^{1/2-\eta-2\varepsilon}$ is

$$\frac{1}{\delta^2} \int_{A, B \leq X^{1/2}} \int_{a, b \leq X^{1/2-\eta-2\varepsilon}} |\Delta(b)| \psi_A(a) \psi_B(b) \tau(a^2 + b^2) \frac{dAdB}{AB} \ll \varepsilon X^{1-\eta},$$

so that we may restrict to $A, B > X^{1/2-\eta-2\varepsilon}$. The condition $a^2 + b^2 \sim X$ is trivially true except in the diagonal cases when

$$A^2 + B^2 = X + O(X^{1-\varepsilon}) \quad \text{or} \quad A^2 + B^2 = 2X + O(X^{1-\varepsilon}),$$

and the total contribution from these diagonal cases is at most

$$\sum_{k \in \{1, 2\}} \frac{1}{\delta^2} \int_{\max(A, B) > X^{1/2-\eta-2\varepsilon}} \int_{a, b \leq X^{1/2-\eta-2\varepsilon}} |\Delta(b)| \psi_A(a) \psi_B(b) \tau(a^2 + b^2) \frac{dAdB}{AB} \ll \varepsilon X^{1-\varepsilon/2}.$$ 

Hence, it suffices to show that for all $A, B \in [X^{1/2-\eta}, 10X^{1/2}]$ we have

$$(3.1) \quad \sum_{d \sim D_1} \alpha_d \sum_{a^2 + b^2 \equiv 0 (d) \atop (a, b) = 1} \Delta(b) \psi_A(a) \psi_B(b) \ll X^{1-2\varepsilon - \eta}. $$

The sum can be written as

$$\sum_{d \sim D_1} \alpha_d \sum_b \Delta(b) \psi_B(b) \sum_{a^2 = -b^2 \equiv 0 (d) \atop (a, b) = 1} \psi_A(a) = \sum_{d \sim D_1} \alpha_d \sum_b \Delta(b) \psi_B(b) \sum_{\nu^2 + 1 \equiv 0 (d) \atop a \equiv 0 (c) \atop (a, b) = 1} \sum_{\nu \equiv 0 (d) \atop a \equiv 0 (c)} \psi_A(a)$$

Expanding the condition $(a, b) = 1$ using the Möbius function gives

$$S(A, B) := \sum_{d \sim D_1} \alpha_d \sum_b \Delta(b) \psi_B(b) \sum_{c | b \atop c > X^\varepsilon} \mu(c) \sum_{\nu^2 + 1 \equiv 0 (d) \atop a \equiv 0 (c)} \sum \psi_A(a).$$

The contribution from $c > X^\varepsilon$ is trivially bounded by

$$\ll \sum_b |\Delta(b)| \psi_B(b) \sum_{c | b \atop c > X^\varepsilon} \sum \psi_A(a) \tau(a^2 + b^2)^2 \ll X^{\varepsilon/4} \sum_b |\Delta(b)| \psi_B(b) \sum_{c | b \atop c > X^\varepsilon} \sum \psi_A(a) \ll X^{1 - 2\varepsilon - \varepsilon/2}. $$
For $c \leq X^\varepsilon$ in $S(A, B)$ we apply the Poisson summation formula (Lemma 7) with $H := X^{\varepsilon'}cD_1/A$ for some small $\varepsilon' > 0$ to the sum over $a$ to get

$$
\sum_{d \sim D_1} \alpha_d \sum_b \Delta(b) \psi_B(b) \sum_{c | b \leq X^\varepsilon} \mu(c) \frac{\hat{\psi}(0)A\rho_1(d)}{cd} \\
+ \sum_{d \sim D_1} \alpha_d \sum_b \Delta(b) \psi_B(b) \sum_{c | b \leq X^\varepsilon} \mu(c) \sum_{\nu^2 + 1 \equiv 0 (d)} \frac{X^{\varepsilon'}D_1}{H} \sum_{0 < |h| \leq H} \hat{\psi} \left( \frac{hA}{cd} \right) e_d(hb\nu)
$$

$$
= \sum_{d \sim D_1} \alpha_d \sum_{c \leq X^\varepsilon} \mu(c) \sum_b \Delta(bc) \psi_B(bc) \frac{\hat{\psi}(0)A\rho_1(d)}{d} \\
+ \sum_{c \leq X^\varepsilon} \alpha_d \sum_{d \sim D_1} \mu(c) \sum_b \Delta(bc) \psi_B(bc) \sum_{\nu^2 + 1 \equiv 0 (d)} \frac{X^{\varepsilon'}D_1}{H} \sum_{0 < |h| \leq H} \hat{\psi} \left( \frac{hA}{cd} \right) e_d(hb\nu).
$$

We can absorb the factor $D_1/d \approx 1$ into the coefficient $\alpha_d$ for the second sum.

3.3.1. **Bounding the main term.** We get by Poisson summation (Lemma 7) using $B > X^{1/2 - \eta}$

$$
\sum_b \Delta(bc) \psi_B(bc) = \sum_b \psi_B(bc) - \sum_{b_0} \psi_B(b_0(b_0^2 + 1))2b_0 = \hat{\psi}(0)B \left( \frac{1}{c} - \frac{\rho_1(c)}{c} \right) + O(X^{-10}).
$$

Summing over $c$ and recalling the definition (2.3) of $\kappa_1$ we see that the main terms cancel in

$$
\sum_{c \leq X^\varepsilon} \frac{\mu(c)}{c^2} (1 - \kappa_1 \rho_1(c)) = \sum_c \frac{\mu(c)}{c^2} - \kappa_1 \sum_c \frac{\mu(c)\rho_1(c)}{c^2} + O(X^{-\varepsilon}) \ll X^{-\varepsilon},
$$

so that the total contribution from the main term is

$$
\ll \hat{\psi}(0)^2ABX^{-\varepsilon/2} \ll X^{1-2\varepsilon-\varepsilon/2},
$$

which is sufficient for (3.1).

3.3.2. **Bounding the non-zero frequencies.** Recall that we need to show that

$$
\sum_{c \leq X^\varepsilon} \sum_{d \sim D_1} \sum_{\nu^2 + 1 \equiv 0 (d)} \left| \sum_b \Delta(bc) \psi_B(bc) \frac{X^{\varepsilon'}}{H} \sum_{0 < |h| \leq H} \hat{\psi} \left( \frac{hA}{cd} \right) e_d(hb\nu) \right| \ll X^{1-2\varepsilon-\eta}.
$$

We can remove the smooth weight $\hat{\psi} \left( \frac{hA}{cd} \right)$ by applying Mellin transform to separate the variables. Thus, we need to show for any bounded coefficients $c_h$ we have

$$
\sum_{c \leq X^\varepsilon} \sum_{d \sim D_1} \sum_{\nu^2 + 1 \equiv 0 (d)} \left| \sum_b \Delta(bc) \psi_B(bc) \sum_{h \leq H} c_h e_d(hb\nu) \right| \ll HX^{1-5\varepsilon-\eta}.
$$

Write $\Delta(b) = \Delta_1(b) - \Delta_2(b)$ with

$$
\Delta_1(b) = 1 \quad \text{and} \quad \Delta_2(b) = \kappa_1 \sum_{b = b_0 + 1} 2|b_0|.
$$
It then suffices to show that for all $c \leq X^{\varepsilon}$ we have (recalling $H = \varepsilon \ell D_1/A$ and $A \leq 2X^{1/2}$)

\begin{equation}
(3.2) \sum_{d \sim D_1} \sum_{\nu^2+1=0(d)} \left| \sum_{b} \Delta_k(bc) \psi_B(b) c_{h} \epsilon_d(h \nu) \right| \ll D_1 X^{1/2-3\varepsilon-\eta}
\end{equation}

Let us consider this for $k = 2$, the case $k = 1$ being similar but easier. Let $N := HB/c$ and denote

$$
\gamma_n := \sum_{n=hb} c_{h} \Delta_2(bc) \psi_B(b) \chi_{h \leq H}
$$

Then by Cauchy-Schwarz and Lemma 6, the left-hand side in (3.2) is at most

$$
\ll D_1^{1/2}(D_1 + HB)^{1/2} \left( \sum_{n \leq N} |\gamma_n|^2 \right)^{1/2}.
$$

We have

\begin{equation}
(3.3) \sum_{n \leq N} |\gamma_n|^2 \ll B^{1/2} \sum_{n \leq N} \tau(n) \gamma_n \ll B \sum_{b_0 \leq 2B^{1/2}} \sum_{h \leq 2H'} \tau(h(b_0^2 + 1)) \ll B^{3/2} H X^{\varepsilon}.
\end{equation}

Thus, the final contribution to (3.2) is (using $B \leq 2X^{1/2}$, $H \leq cX^{\varepsilon} D_1/A$, $c \leq X^{\varepsilon}$, and $A \geq X^{1/2-\varepsilon}$)

$$
\ll X^{\varepsilon}(D_1(B^{3/2} H)^{1/2} + D_1^{1/2} B^{5/4} H) \ll X^{6\varepsilon} D_1^{3/2} X^{1/8}.
$$

Hence, (3.2) is satisfied provided that

$$
D_1^{3/2} X^{1/8} \ll D_1 X^{1/2-9\varepsilon-\eta},
$$

which holds if

$$
D_1 \ll X^{3/4-3\eta'}.
$$

for $\varepsilon = 4\eta$ and $\eta' = 100\eta$.

3.3.3. Case of $a^2 + (c^3 + d^3)^2$. The details are essentially the same, with the only change that we let

$$
\Delta(b) := 1 - \kappa_2 \sum_{b = c_0^3 + d_0^3 \atop c_0, d_0 > X^{1/6-\varepsilon}} \Omega(c_0, d_0),
$$

bounding the contribution from $\min\{c_0, d_0\} \leq X^{1/6-\varepsilon}$ trivially. The contribution of the main term from Poisson summation is handled by the same argument as in Section 3.3.1, applying Poisson summation to the variables $c_0, d_0$ to extract $\rho_2(c)$.

For the non-zero frequencies, the argument is essentially the same as in Section 3.3.2. Let

$$
\Delta'(b) := \kappa_2 \sum_{b = c_0^3 + d_0^3 \atop c_0, d_0 > X^{1/6-\varepsilon}} \Omega(c_0, d_0) \quad \text{and} \quad \gamma'_n := \sum_{n=hb} c_{h} \Delta'(bc) \psi_B(b) \psi_{H'}(h).
$$

To get the bound corresponding to (3.3) one also has to use Lemma 8 to bound the number of representations as a sum of two cubes, which gives

$$
\sum_{n \sim N} |\gamma_n'|^2 \ll X^{\varepsilon^2} B^{1/3} \sum_{n \sim N} \tau(n) \gamma_n' \ll X^{\varepsilon^2} B^{2/3} \sum_{c_0, d_0 \leq 2B^{1/3}} \sum_{h \leq 2H'} \tau(h(c_0^3 + d_0^3)) \ll B^{4/3} H' X^{\varepsilon}.
$$
The end bound is (using $B \leq 2X^{1/2}$, $H_1 \leq cX^{1/2}D_1/A$, $c \leq X^{\varepsilon}$, and $A \geq X^{1/2-\varepsilon}$)
\[
\ll X^{2\varepsilon} \left(D_1 (B^{4/3}H')^{1/2} + D_1^{1/2} B^{7/6} H' \right) \ll X^{6\varepsilon} D_1^{3/2} X^{1/12},
\]
which is sufficient for the bound (3.2) if $D_2 \leq X^{5/6-\eta'}$ with $\varepsilon = 4\eta$ and $\eta' = 100\eta$. □

3.4. Proof of Proposition 5 To prove Proposition 5 we write
\[
1_{(n,P(W))=1} = \sum_{e \mid (n,P(W))} \mu(e) = \sum_{e \mid (n,P(W)) \leq X^\eta} \mu(e) + \sum_{e \mid (n,P(W)) > X^\eta} \mu(e).
\]
For $e \leq X^\eta$ we get by Proposition 4 with level $D_j X^\eta$
\[
\sum_{e \leq X^\eta/2} \mu(e) \sum_{d \sim D_j \atop n \sim X/d \text{ even}} \alpha(d) (a_{de,n}^{(j)} - b_{de,n}) \ll_{C} X \log^{-C} X,
\]
since the coefficient of $d' = ed$ is bounded by $\ll X^{\eta}/2$.

For $e > X^\eta$ we note that $e \mid P(W)$ implies that there is a divisor $e_1 \mid e$ with $e_1 \in [X^\eta, X^{2\eta}]$. Thus, by crude bounds, Lemma 11 and Proposition 4 we get
\[
\sum_{e > X^\eta \atop e \mid P(W)} \mu(e) \sum_{d \sim D_j \atop n \sim X/d \text{ even}} \alpha(d) (a_{de,n}^{(j)} - b_{de,n}) \ll \sum_{e_1 \in [X^\eta, X^{2\eta}] \atop P^+(e_1) < W} \tau(n) \tau(e_1) (a_{e_1,n}^{(j)} + b_{e_1,n}) \ll_{C} X \log^{-C} X
\]
by applying Lemma 9 to get the last bound. □

4. Type II information: initial reductions

Let
\[
\varphi_{Z[i]}(d) := |\{v \in \mathbb{Z}[i]/d\mathbb{Z}[i] : (v, d) = 1\}|.
\]
For coefficients $\beta(n)$ we denote $\beta_z := \beta(|z|^2)$ for $z \in \mathbb{Z}[i]$ and for any $Y > 0$ write $\psi_Y(y) := \psi(y/Y)$. It is convenient for us to give the following definition of the Siegel-Walfisz property over Gaussian integers, where the variable is weighted by a smooth function supported on a polar box.
Definition 1. (Siegel-Walfisz property on $\mathbb{Z}[i]$). We say that coefficients $\beta(n)$, supported on $n \sim N$, satisfies the Siegel-Walfisz property if the following holds for any $C > 0$ and any smooth weight $\psi$ as in Section 3.1 with $\delta^{-1} = \log^O(1) N$. For any $\theta \in [\pi, 5\pi]$, $N_1 \sim N$, $v \in \mathbb{Z}[i]$ and $d = \log^O(1) N$ we have

\[
\sum_{z \equiv v(d)} \beta_z \psi_N(\frac{|z|^2}) \psi_\theta(\arg z + 2\pi) = \frac{\psi_\theta^\cap(2\pi)}{\varphi_{\mathbb{Z}[i]}(d)} \sum_{z} \beta_z \psi_N(\frac{|z|^2}) + O_C(N \log^{-C} N).
\]

For all coefficients $\beta(n)$ that we consider this property follows by standard arguments (similar to [6 Section 16]). More precisely, we will need this for $\beta(n)$ which have a large prime factor, that is, for $\beta(n)$ such that for some $PK \asymp N$ with $P \gg W$ and for some divisor bounded coefficients $\gamma(k)$ supported on $(k, P(W)) = 1$ we have

\[
\beta(n) = \psi_N(n) \sum_{n=pk \atop p \equiv 1 \pmod{4}} \gamma(k).
\]

Fixing $|u|^2 = k$ and relabeling, it then suffices to show that for $P \gg W$ we have

\[
\sum_{z \equiv v(d)} \Lambda(\frac{|z|^2}) \psi_P(\frac{|z|^2}) \psi_\theta(\arg z + 2\pi) = \frac{\psi_\theta^\cap(2\pi)}{\varphi_{\mathbb{Z}[i]}(d)} \sum_{z} \Lambda(\frac{|z|^2}) \psi_P(\frac{|z|^2}) + O_C(P \log^{-C} N).
\]

Using Dirichlet characters to detect $z \equiv v(d)$ and the Fourier series expansion to $\psi_\theta$ the contribution from the principal character and the zeroth frequency in the Fourier expansion of $\psi_\theta$ cancels the main term on the right-hand side of (4.1). For non-principal characters or non-zero frequencies we use Melling inversion for $\psi_P$ and shift the contour past $s = 1$. Shifting the contour is justified by applying the Siegel-Walfisz bound [6 Lemma 16.1] with the character

\[
\psi(z) := \chi(z) \left(\frac{z}{|z|}\right)^k
\]

where $\chi$ a Dirichlet character of $(\mathbb{Z}[i]/d\mathbb{Z}[i])^\times$.

Our Type II information for Theorems 1 and 2 is given by the following.

Proposition 12. (Type II information). Let $W = X^{1/(\log \log X)^2}$. For $j \in \{1, 2\}$ let $M_j N_j \asymp X$ and for any small $\eta > 0$ let

\[
X^{1/4+\eta} \ll N_1 \ll X^{1/3-\eta}
\]

and

\[
X^{1/6+\eta} \ll N_2 \ll X^{2/9-\eta}.
\]

Let $\alpha(m)$ and $\beta(n)$ be bounded coefficients. Assume that $\beta(n)$ is supported on square-free numbers with $(n, P(W)) = 1$ and satisfies the Siegel-Walfisz property (4.1) and assume that $\alpha(m)$ is supported on $(m, P(W)) = 1$. Then for any $C > 0$

\[
\sum_{m \sim M_j \atop n \sim N_j} \alpha(m) \beta(n)(a_{mn}^{(j)} - b_{mn}) \ll_{\eta, C} X \log^{-C} X.
\]
We begin by reducing to a case where $\beta(n)$ is oscillatory. For this we make the following definition.

**Definition 2.** (Siegel-Walfisz property with main term equal to 0). We say that coefficients $\beta(n)$, supported on $n \sim N$, satisfies the Siegel-Walfisz property with main term equal to 0 if the following holds for any $C > 0$ and any smooth weight $\psi$ as in Section 3.1 with $\delta^{-1} = \log^O(1) N$. For any $\theta \in [\pi, 5\pi], N_1 \sim 2N, v \in \mathbb{Z}[\theta]$ and $d = \log^O(1) N$ we have

$$\sum_{z \equiv v \pmod{d}} \beta(z) \psi(N_1(|z|^2)) \psi(\arg z + 2\pi) \ll C N \log^{-C} N.$$  

(4.3)

We show in Section 4.1 that for Proposition 12 it suffices to prove the following.

**Proposition 13.** Let $W = X^{1/(\log \log X)^2}$. For $j \in \{1, 2\}$ let $M_j N_j \asymp X$ and for any small $\eta > 0$ let

$$X^{1/4 + \eta} \ll N_1 \ll X^{1/3 - \eta}$$

and

$$X^{1/6 + \eta} \ll N_2 \ll X^{2/9 - \eta}.$$  

Let $\alpha(m)$ and $\beta(n)$ be bounded coefficients. Assume that $\beta(n)$ is supported on square-free numbers with $(n, P(W)) = 1$ and satisfies the Siegel-Walfisz property with main term equal to 0, that is, (4.3). Assume that $\alpha(m)$ is supported on $(m, P(W)) = 1$. Then for any $C > 0$

$$\sum_{m \sim M_j} \alpha(m) \beta(n) \alpha^{(j)}_{mn} \ll_{\eta, C} X \log^{-C} X$$

and

$$\sum_{m \sim M_j} \alpha(m) \beta(n) b_{mn} \ll C X \log^{-C} X.$$  

We will give the proof of this proposition in Section 6.

4.1. **Proof that Proposition 13 implies Proposition 12** The idea is similar to one that appears in Heath-Brown’s work [9]. Let $W := X^{1/(\log \log x)^2}$. For a given $\delta > 0$ let $\delta' = \delta(\log X)^{-C}$ for some large $C > 0$ and let $\psi'$ be a smooth function as in Section 3.1 with parameter $\delta'$. Let

$$C(\beta, N') := \left( \sum_{(z, P(W)) = 1} \beta(|z|^2) \psi'_{N'}(|z|^2) \right) \left( \sum_{(z, P(W)) = 1} \psi'_{N'}(|z|^2) \right)^{-1}$$

(4.4)

Given the coefficients $\beta(n)$ we write

$$\beta^\#(n) := 1_{(n, P(W)) = 1} \int_{N' \sim N_j} \frac{1}{\delta'} \psi'_{N'}(n) C(\beta, N') \frac{dN'}{N'},$$

approximates $\beta(n)$ by $1_{(n, P(W)) = 1}$ while mimicking the distribution of $\beta(|z|^2)$ over short intervals near $n$. If $\beta(n)$ is bounded and supported on $(n, P(W)) = 1$, then the coefficient $\beta^\#(n)$ is also bounded since $|C(\beta, N')| \ll 1$ and $\int \psi'(y) \frac{dy}{y} = \delta'$. By Proposition 5 the
claim of Proposition 12 holds for $\beta$ replaced by $\beta^\#$, that is, if we let $S(\alpha, \beta)$ denote the Type II sum, then

$$S(\alpha, \beta^\#) = \sum_{m \sim M_j, n \sim N_j} \alpha(m)\beta^\#(n)(a_{mn} - b_{mn})$$

(4.5)

$$= \frac{1}{\delta} \int C(\beta, N') \sum_{m \sim M_j, n \sim N_j} \alpha(m)1_{(n, P(W)) = 1}(\psi'_N(n)(a_{mn} - b_{mn}) \frac{dN'}{N'})$$

$$\ll_C X \log^{-C} X,$$

since the weight $\psi'_N(n)$ may be removed by partial summation. Furthermore, by the fundamental lemma of the sieve (arguing similar to Section 3.4) it follows that $\beta^\#(n)$ satisfies the Siegel-Walfisz condition (4.1). Thus, if $\beta(n)$ also satisfies the Siegel-Walfisz property, then the coefficient $(\beta - \beta^\#)(n)$ satisfies the Siegel-Walfisz property (4.3) with a main term equal to 0, as can be seen from

$$\sum_{z \equiv e(d)} (\beta - \beta^\#)(|z|^2)\psi_{N_1}(|z|^2)\psi_\theta(\arg z + 2\pi)$$

$$= \frac{\psi_\theta(0)1_{(v,d) = 1}}{\varphi_{\mathbb{Z}/(d)}} \sum_{\beta - \beta^\#}(\psi'_{N_1}(|z|^2)) + O_C(N \log^{-C} N)$$

(4.6)

$$= \frac{\psi_\theta(0)1_{(v,d) = 1}}{\varphi_{\mathbb{Z}/(d)}} \int_{N' \sim N_j} \frac{1}{\delta'} \sum_z (\beta'(|z|^2) - C(\beta, N')1_{(z, P(W)) = 1})\psi'_{N_1}(|z|^2)\psi_{N_1}(|z|^2) \frac{dN'}{N'} + O_C(N \log^{-C} N)$$

$$\ll_C N \log^{-C} N,$$

where the last bound holds by the construction (4.4) since the parameter $\delta'$ (associated to $\psi'_{N_1}$) is small in comparison to the parameter $\delta$ (associated to $\psi_{N_1}$). Thus, the using decomposition

$$\beta = \beta^\# + (\beta - \beta^\#), \quad S(\alpha, \beta) = S(\alpha, \beta^\#) + S(\alpha, \beta - \beta^\#)$$

and (4.5) and (4.6), Proposition 12 follows from Proposition 13 applied to bound $S(\alpha, \beta - \beta^\#)$.

5. Applications of the Weil bound

For the proof of Proposition 13 we require the Weil bound both for counting points on certain curves over finite fields as well as for bounding algebraic exponential sums. Let $p$ be a prime and fix an algebraic closure $\overline{\mathbb{F}}_p$ of the finite field $\mathbb{F}_p$. Let $C$ be the projective curve defined by a polynomial equation $F(X, Y, Z) = 0$ with a homogeneous polynomial $F \in \mathbb{F}_p[X, Y, Z]$. Recall that $C$ is non-singular if for all $(x, y, z) \in \mathbb{P}^2(\overline{\mathbb{F}}_p)$ (that is, $(x, y, z) \neq (0, 0, 0)$) with $F(x, y, z) = 0$ we have $\partial_X F(x, y, z) \neq 0$ or $\partial_Y F(x, y, z) \neq 0$ or $\partial_Z F(x, y, z) \neq 0$.

For Theorem 1 we need to understand affine curves over $\mathbb{F}_p$ defined by the equation

$$x_1^2 + 1 = a(x_2^2 + 1)$$
with varying $a$. For $a \neq 0,1$ the homogenization $(1 - a)x_0^2 + x_1^2 - ax_2^2 = 0$ defines a non-singular curve. For Theorem 2 we end up with affine varieties defined by the equation

$$x_1^2 + x_2^2 = a(x_3^2 + x_4^2)$$

with varying $a$. By fixing $x_3$ and $x_4$ this reduces to understanding curves defined by the equation

$$x_1^2 + x_2^2 = \gamma,$$

which is non-singular for $\gamma \neq 0$.

5.1. Points on curves. We set

$$N_1(a; d) := \{|x_1, x_2(d) : x_1^2 + 1 \equiv a(x_2^2 + 1) (d)\}|,$$

and define

$$\varepsilon_p(a) := N_1(a; p) - p.$$

For $k \geq 2$ we define

$$\varepsilon_{pk}(a) := N_1(a, p^k) - pN_1(a, p^{k-1}) = N_1(a, p^k) - (p^k + p^{k-1}\varepsilon_p(a)) + \cdots + p\varepsilon_{pk-1}(a)$$

and for any $d$

$$\varepsilon_d(a) := \prod_{p^k|\,d} \varepsilon_{pk}(a).$$

Then by the Chinese remainder theorem

$$N_1(a, D) = \prod_{p^k\mid D} N_1(a, p^k) = \prod_{p^k\mid D} \left( \sum_{j=0}^{k} p^{k-j}\varepsilon_{pk}(a) \right) = \sum_{d\mid D} D d \varepsilon_d(a).$$

**Lemma 14.** For $a \neq 0,1 (p)$ we have

$$\varepsilon_p(a) \ll p^{1/2}.$$

For any $a$ and $k \geq 1$ we have

$$N_1(a; p^k) \leq 2p^k.$$

For any $d$ we have

$$|\varepsilon_d(a)| \leq d\tau(d)^{O(1)}.$$

**Proof.** For $a \neq 0,1 (p)$ the projective curve over $\mathbb{F}_p$ defined by

$$(1 - a)x_0^2 + x_1^2 - ax_2^2 = 0$$

is non-singular, so the claim follows from the classical Hasse-Weil bound ([21] or [ in Exercise V.1.10], for instance). The last claim follows from

$$|\varepsilon_{pk}(a)| = |N_1(a, p^k) - pN_1(a, p^{k-1})| \leq \max\{N_1(a, p^k), pN_1(a, p^{k-1})\} \leq 2p^k.$$

□

Recall that an integer $c$ is said to be powerful if $p|c$ implies $p^2|c$. The next lemma gives an evaluation of $N_1(a; D)$ for a generic $a$ unless $D$ has a large powerful factor. When we apply this we take $Y = \log^C X$ and make use of the fact that $D$ which have a large powerful factor are very sparse.
Lemma 15. Let $Y \geq 1$. Let $P_Y(d)$ denote the characteristic function of $d$ such that there exists a powerful divisor $c|d$ with $c > Y$. Then

$$N_1(a; D) = \sum_{d|D, d < Y^4} \frac{D}{d} \varepsilon_d(a) + D\tau(D)O(1)O(Y^{-1} + P_Y(D) + 1_{(a(a-1), D > Y)}).$$

Proof. By the Chinese remainder theorem we have

$$N_1(a; D) = \prod_{p^k|D} N_1(a; p^k).$$

If $P_Y(D) = 1$ or $(a(a-1), D) > Y$, then we use the bound $N_1(a; p^k) \leq 2p^k$ to get

$$N_1(a; D) \ll D\tau(D).$$

Suppose then that $P_Y(D) = 0$ and $(a(a-1), D) \leq Y$. We have

$$N_1(a; D) = \prod_{p^k|D} (p^k + \varepsilon_{p^k}(a)) = \sum_{d|D} \frac{D}{d} \varepsilon_d(a) = \sum_{d|D, d < Y^4} \frac{D}{d} \varepsilon_d(a) + \sum_{d|D, d > Y^4} \frac{D}{d} \varepsilon_d(a).$$

Let us write $D = D_1D_2$ where $D_1$ is the largest square-free divisor with $(D_1, D/D_1) = 1$ so that $D_2$ is powerful. Then for any $d|D$ with $d \geq Y^4$ we have $d_1 := (d, D_1) \geq Y^3$ and

$$d_0 := \frac{d_1}{(a(a-1), d_1)} \geq Y^2.$$

By Lemma 14 we have $\varepsilon_{d_0}(a) \ll \tau(D)O(1)d_0^{1/2}$, so that for $d > Y^4$ we have

$$\varepsilon_{d_0}(a)/d_0 \ll \tau(D)O(1)Y^{-1}.$$

and by the trivial bound $\varepsilon_{d/d_0}d_0/d \ll \tau(D)O(1)$ we get

$$\frac{\varepsilon_d(a)}{d} = \frac{\varepsilon_{d_0}(a)\varepsilon_{d/d_0}(a)}{d/d_0} \ll \tau(D)O(1)Y^{-1}.$$

Therefore, the contribution from $d \geq Y^4$ is

$$\ll \sum_{d|D, d > Y^4} \frac{D}{d} \varepsilon_d(a) \ll D\tau(D)O(1)Y^{-1}.$$

\[\square\]

We now give a similar treatment for the curve corresponding to Theorem 2. Similarly as in Lemma 14 we get the following by the Weil bound.

Lemma 16. Denote

$$N_2(\gamma; D) := |\{x_1, x_2 (D) : x_1^3 + x_2^3 \equiv \gamma (D)\}|$$

Let $\gamma \not\equiv 0 (p)$. Then

$$N_2(\gamma; p) = p + O(p^{1/2})$$

and

$$N_2(0; p) \ll p.$$

Furthermore, $N_2(\gamma, p^k) \ll p^k$. 

Lemma 17. Let
\[ N_3(a; D) := |\{x_1, x_2, x_3, x_4 : x_1^3 + x_2^3 \equiv a(x_3^3 + x_4^3) (D)\}| \]
and define \( \varepsilon'_p(a) \) by
\[ p^2 \varepsilon'_p(a) := N_3(a; p) - p^3, \quad p^{2k} \varepsilon'_p(a) := N_3(a; p^k) - p^3 N_3(a; p^{k-1}). \]
Set
\[ \varepsilon'_d(a) := \prod_{p^k || d} \varepsilon'_p(a). \]
For \( a \not\equiv 0 (p) \) we have
\[ N_3(a; p) = p^3 + O(p^{5/2}), \]
and for any \( a \) we have \( N_3(a, p) \ll p^3 \). Furthermore,
\[ \varepsilon'_d(a) \ll d \tau(d)^{O(1)}. \]
Proof. Clearly
\[ N_3(a; p) = \sum_{x_3, x_4 (p)} N_2(a(x_3^3 + x_4^3); p). \]
The contribution from \( x_3^3 + x_4^3 \equiv 0 (p) \) is trivially bounded by \( \ll p^2 \), and from \( x_3 + x_4 \not\equiv 0 (p) \) we get a contribution
\[ (p^2 + O(p))(p + O(p^{1/2})) = p^3 + O(p^{5/2}). \]
The last claim follows since by \( N_2(\gamma, p^k) \ll p^k \) we have \( N_3(a, p^k) \ll p^{3k} \) and therefore \( \varepsilon'_p(a) \ll p^k \).

Note that by the Chinese remainder theorem
\[ N_3(a; D) = \prod_{p^k || D} N_3(a, p^k) = D^2 \prod_{p^k || D} \left( \sum_{j=0}^{k} p^{k-j} \varepsilon'_p(a) \right) = \sum_{D || d} D^3 \frac{D^3}{d} \varepsilon'_d(a). \]
Then by a similar argument as in Lemma [15] we get the following.

Lemma 18. We have for \( \varepsilon'_d(a) \ll d \tau(d)^{O(1)} \) that
\[ N_3(a; D) = \sum_{D || d} D^3 \frac{D^3}{d} \varepsilon'_d(a) + D^3 \tau(D)^{O(1)} O(Y^{-1} + P_Y(D) + 1_{(a, D) > Y}). \]

5.2. Algebraic exponential sums.

Lemma 19. Let \( h = (h_1, h_2) \). We have for \( a \not\equiv 0, 1 (p) \)
\[ \sum_{x_1, x_2 (p) \atop x_1^2 + 1 \equiv a(x_2^2 + 1) (p)} e_p(h \cdot x) \ll (h_1, h_2, p)^{1/2} p^{1/2}, \]
and for \( a \equiv 1 (p) \) we have
\[ \sum_{x_1, x_2 (p) \atop x_1^2 + 1 \equiv a(x_2^2 + 1) (p)} e_p(h \cdot x) \ll (h_1 - h_2, p) + (h_1 + h_2, p) \]
Proof. For \( a \equiv 1 (p) \) we have
\[
\sum_{x_1, x_2 (p)} e_p(h \cdot x) = -1 + \sum_{x (p)} e_p((h_1 + h_2)x) + e_p((h_1 - h_2)x)
\]
\[
\ll (h_1 - h_2, p) + (h_1 + h_2, p).
\]
Consider \( a \not\equiv 0, 1 (p) \) and fix \( j \in \{1, 2\} \). If \((h_1, h_2, p) \neq 1\) we apply the trivial bound, so we may assume \((h_1, h_2, p) = 1\). The claim now follows from the Weil bound (see [1, Theorem 5], for instance), applied to the projective (hence complete) curve \( X \) defined by the (homogenized) polynomial
\[
(1 - a)x_0^2 + x_1^2 - ax_2^2,
\]
which is non-singular for \( a \not\equiv 0, 1 (p) \).

By using the Chinese remainder theorem as in the proof of Lemma 19, this implies the following.

Lemma 20. Let \( D_1 \) be the largest square-free divisor of \( D \) such that \((D_1, D/D_1) = 1\). Let \((a, D) = 1\). Then
\[
\sum_{x_1, x_2 (D)} e_D(h \cdot x) \ll \tau(D)^{O(1)} DD_1^{-1/2} (h_1, h_2, D_1)^{1/2} (a - 1, D_1)^{1/2}
\]
Proof. By the Chinese remainder theorem we have
\[
\sum_{x_1, x_2 (D)} e_D(h \cdot x) = \prod_{p^k|D} \sum_{x_1, x_2 (p^k)} e_{p^k}((D/p^k)^{-1} h \cdot x),
\]
where \((D/p^k)^{-1}\) denotes the multiplicative inverse modulo \( p^k \). For \( p^k|\langle D/D_1 \rangle \) we use the trivial bound
\[
\sum_{x_1, x_2 (p^k)} e_{p^k}((D/p^k)^{-1} h \cdot x) \ll p^k.
\]
For \( p|D_1 \) we have by Lemma 19
\[
\sum_{x_1, x_2 (p)} e_p((D/p)^{-1} h \cdot x) \ll p^{1/2} (h_1, h_2, p)^{1/2} (a - 1, p)^{1/2}
\]
We now turn to the corresponding sum for Theorem 2. Here we apply a slightly different argument to the above.

Lemma 21. Let \( h = (h_1, h_2, h_3, h_4) \). We have for \( a \not\equiv 0 (p) \)
\[
\sum_{x_1, x_2, x_3, x_4 (p)} e_p(h \cdot x) \ll (h_1, h_2, h_3, h_4, p)^p
\]
Proof. If \((h_j, p) \neq 1\) for all \(j \leq 4\) then we apply the trivial bound, so we may assume that \((h_1, h_2, h_3, h_4, p) = 1\). We expand the congruence condition to get

\[
\sum_{x_1, x_2, x_3, x_4 (p) \atop x_1^2 + x_2^2 \equiv a(x_3^3 + x_4^3)} e_p(h \cdot x) = \frac{1}{p} \sum_{w (p)} \sum_{x_1, x_2, x_3, x_4 (p) \atop x_1^2 + x_2^2 \equiv a(x_3^3 + x_4^3)} e_p(h \cdot x + w(x_1^3 + x_2^3 - a(x_3^3 + x_4^3)))
\]

\[
= \frac{1}{p} \sum_{w (p)} \prod_{j=1}^4 \left( \sum_{x (p)} e_p(h_j x + wa_j x^3) \right),
\]

where \(a_1 = a_2 = 1\) and \(a_3 = a_4 = -a\). The contribution from \(w = 0\) is 0 (as \((h_j, p) = 1\) for some \(j \in \{1, 2, 3, 4\}\), and for \(w \neq 0\) we apply the Weil bound ([1, Theorem 5] for \(X = \mathbb{P}^1\) or [20]) to each of the four sums to get

\[
\sum_{x_1, x_2, x_3, x_4 (p) \atop x_1^2 + x_2^2 \equiv a(x_3^3 + x_4^3)} e_p(h \cdot x) \ll \frac{1}{p} \sum_{w (p)} (p^{1/2})^4 \ll p^2.
\]

Remark 3. It should be possible to obtain the Deligne bound \(\ll p^{3/2}\) for the generic \(h\) and \(a\), but since this is not necessary for our application nor sufficient for an asymptotic version of Theorem 2, we do not pursue this issue here.

Using Lemma 21 we get the following by the Chinese remainder theorem.

**Lemma 22.** Let \(D_1\) be the largest square-free divisor of \(D\) such that \((D_1, D/D_1) = 1\). Let \((a, D) = 1\). Then

\[
\sum_{x_1, x_2, x_3, x_4 (\mathbb{Z}/D\mathbb{Z}) \atop x_1^2 + x_2^2 \equiv a(x_3^3 + x_4^3)} e_D(h \cdot x) \ll \tau(D)^{O(1)} D^3 D_1^{-1}(h_1, h_2, h_3, h_4, D_1).
\]

**Proof.** By the Chinese remainder theorem we have

\[
\sum_{x_1, x_2, x_3, x_4 (\mathbb{Z}/D\mathbb{Z}) \atop x_1^2 + x_2^2 \equiv a(x_3^3 + x_4^3)} e_D(h \cdot x) = \prod_{p \mid D} \sum_{x_1, x_2, x_3, x_4 (\mathbb{Z}/p \mathbb{Z}) \atop x_1^2 + x_2^2 \equiv a(x_3^3 + x_4^3)} e_{p^k}((D/p^k)^{-1}h \cdot x),
\]

where \((D/p^k)^{-1}\) denotes the multiplicative inverse modulo \(p^k\). For \(p^k||D/D_1\) we use the trivial bound

\[
\sum_{x_1, x_2, x_3, x_4 (\mathbb{Z}/p^k \mathbb{Z}) \atop x_1^2 + x_2^2 \equiv a(x_3^3 + x_4^3)} e_{p^k}((D/p^k)^{-1}h \cdot x) \ll p^{3k}.
\]

For \(p|D_1\) we have by Lemma 19

\[
\sum_{x_1, x_2, x_3, x_4 (\mathbb{Z}/p \mathbb{Z}) \atop x_1^2 + x_2^2 \equiv a(x_3^3 + x_4^3)} e_p((D/p)^{-1}h \cdot x) \ll p^2(h_1, h_2, h_3, h_4, p).
\]
6. Proof of Proposition 13

We need the following simple bound.

Lemma 23. Let $\Delta = \Delta(z_1, z_2) := \text{Im}(z_2 \overline{z_1})$ for $z_1, z_2 \in \mathbb{Z}[i]$. Then

$$\sum_{|z_1|^2, |z_2|^2 \sim N \atop \Delta \sim D} 1 \ll N \log^2 N.$$  

Proof. Writing $z_j = r_j + is_j$, we have $\Delta = r_2 s_1 - r_1 s_2$. Since $\Delta \neq 0$ we have $s_1 \neq 0$ or $s_2 \neq 0$, so that we can assume that $s_2 \neq 0$. Summing over $r_2$ with $r_2 s_1 \equiv D(s_2)$ we have

$$\sum_{|z_1|^2, |z_2|^2 \sim N \atop \Delta \sim D} 1 \ll N^{1/2} \sum_{s_1, s_2 \leq 2N^{1/2}} (s_1, s_2) s_2 \ll N \log^2 N.$$  

□

6.1. Application of Cauchy-Schwarz. Since $mn$ is not divisible by any prime $p \equiv 3 (4)$ by the support of $a^{(j)}_m$ and $b_m$, we can write $m = |w|^2$ and $n = |z|^2$ for some Gaussian integers $w, z \in \mathbb{Z}[i]$. Since the expression $\alpha(m) \beta(n)$ is supported on $(mn, P(W)) = 1$, the conditions $(a, b^2 + 1) = 1$, $(a, c^3 + d^3) = 1$, and $(a, b) = 1$ in the definitions of $a_n^{(1)}$, $a_n^{(2)}$, and $b_n$ may be dropped with a negligible error term. Let

$$S_1^{(1)}(w, z) = \sum_{\text{Re}(\overline{w}z) = b^2 + 1 \atop b > 0} b,$$

$$S_1^{(2)}(w, z) = \sum_{\text{Re}(\overline{w}z) = c^3 + d^3 \atop c, d > 0} \Omega(c, d),$$

$$S_2^{(1)}(w, z) = S_2^{(2)}(w, z) = 1_{\text{Re}(\overline{w}z) > 0}.$$  

Denote $\alpha_w = \alpha(|w|^2)$ and $\beta_z = \beta(|z|^2)$. Note that since $\beta(n)$ is supported on square-free numbers, $\beta$ is supported on $z \in \mathbb{Z}[i]$ with $(z, \overline{z}) = 1$. Then Proposition 13 follows once we show that for $j, k \in \{1, 2\}$

$$\sum_{|w|^2 \sim M_j \atop |z|^2 \sim N_j} \alpha_w \beta_z S_k^{(j)}(w, z) \ll X \log^{-C} X.$$  

Let $F_M(m) = F(m/M_j)$ denote a $C^\infty$-smooth majorant of $m \sim M_j$, for some fixed smooth $F$ supported on $[1/2, 4]$. Applying Cauchy-Schwarz we get

$$(6.1) \sum_{|w|^2 \sim M_j \atop |z|^2 \sim N_j} \alpha_w \beta_z S_k^{(j)}(w, z) \ll M_j^{1/2} (U_k^{(j)})^{1/2}$$

for

$$U_k^{(j)} := \sum_{|z_1|^2, |z_2|^2 \sim N_j} \beta_{z_1} \beta_{z_2} \sum_w F_M_j(|w|^2) S_k^{(j)}(w, z_1) S_k^{(j)}(w, z_2).$$  

Then our task is reduced to showing that for $j, k \in \{1, 2\}$ we have

$$(6.2) U_k^{(j)} \ll_C X N_j \log^{-C} X.$$
Let us first restrict only to the case \( j = k = 1 \). Similarly as in \([6, \text{Section 6}]\), let
\[
\Delta = \Delta(z_1, z_2)^{\text{def}} = \text{Im}(\overline{z_1}z_2) = r_1s_2 - r_2s_1 = |z_1z_2|\sin(\arg z_2 - \arg z_1).
\]
Let \( a_i, b_i \) denote integers such that
\[
\text{Re}(wz_i) = (b_i^2 + 1) \quad \text{and} \quad \text{Im}(wz_i) = a_i.
\]
Then
\[
z_2\text{Re}(wz_1) - z_1\text{Re}(wz_2) = \frac{1}{2}w(z_2\overline{z_1} - z_1\overline{z_2}) = i\Delta w,
\]
that is,
\[
i\Delta w = z_2(b_1^2 + 1) - z_1(b_2^2 + 1).
\]
We first bound separately the contribution from the diagonal terms where \( \Delta = 0 \) or \( (z_1, z_2) > 1 \).

### 6.2. Diagonal terms with \( \Delta = 0 \)

By \( \Delta = 0 \) we have
\[
z_1(b_1^2 + 1) = z_2(b_2^2 + 1).
\]
Multiplying both sides by \( \overline{w} \) and taking the imaginary parts we find
\[
a_1(b_1^2 + 1) = a_2(b_2^2 + 1).
\]
Hence, from \( \arg z_1 = \arg z_2 \) we get a contribution to \( U_1^{(1)} \) of size
\[
\sum_{a_1, a_2 \ll X^{1/2}} \sum_{b_1, b_2 \ll X^{1/4}} \sum_{\ell \ll X^{3/4}} \left( \sum_{\ell = a(b^2 + 1)} 1 \right)^2 \ll \varepsilon X^{5/4 + \varepsilon} \ll X^{1 - \eta} N_j.
\]
by using the bound \( \tau(\ell) \ll \varepsilon \ell^\varepsilon \) and the assumption \( N_1 \gg X^{1/4 + \eta} \). The same bound holds for \( U_2^{(1)} \), so that the diagonal contribution is sufficiently small in terms of \( \eta \).

Similar arguments apply in the case \( j = 2 \), with \( b_1^2 + 1 \) replaced by \( c_1^3 + d_1^3 \), and by using the trivial bound \( \Omega(c_j, d_j) \ll \max\{c_j, d_j\} \ll X^{1/6} \) we get a contribution
\[
\sum_{c_1, c_2, d_1, d_2 \ll X^{1/6}} \Omega(c_1, d_1)\Omega(c_2, d_2) \ll X^{1/3} \sum_{c_1, c_2, d_1, d_2 \ll X^{1/6}} \left( \sum_{a_1, a_2 \ll X^{1/2}} \sum_{c_1, d_1, d_2 \ll X^{1/6}} \sum_{a_1(c_1^3 + d_1^3) = a_2(c_2^3 + d_2^3)} 1 \right)^2.
\]
Using Lemma \( \S \) to bound the number of representations as a sum of two cubes we get
\[
\sum_{a_1, a_2 \ll X^{1/2}} \Omega(c_1, d_1)\Omega(c_2, d_2) \ll \varepsilon X^{7/6 + \varepsilon} \ll X^{1 - \eta} N_2
\]
since \( N_2 \gg X^{1/6 + \eta} \).
6.3. Diagonal terms with \((z_1, z_2) > 1\). Consider first the case \(j = 1\). Recall that, by assumptions in Proposition \[^{13}\] \(\beta(n)\) is supported on \(n\) with \((n, P(W)) = 1\) with \(W = X^{1/(\log \log X)}\). Therefore, if \(z_0 := (z_1, z_2) > 1\), then \(|z_0|^2 \geq W\). Making the change of variables \(z_i \mapsto z_0 z_i\), we have

\[
\text{Re}(\overline{w} z_0 z_i) = (b_i^2 + 1).
\]

Denoting \(\Delta' = \text{Im}(\overline{w} z_0 z_2) \neq 0\) and combining \(w' = w z_0\) we have

\[
i \Delta' w' = z_2 (b_1^2 + 1) - z_1 (b_2^2 + 1) \neq 0,
\]

so that fixing \(z_1, z_2, b_1, b_2\) fixes \(w'\). We have

\[
\sum_{W \leq |z_0|^2 \leq N_1} 1_{|w|w} \ll 2 \log X / \log W \ll W^{1/2}.
\]

Hence, by a dyadic decomposition we see that the part \((z_1, z_2) > 1\) contributes at most

\[
\ll \sum_{W \leq |z_0|^2 \leq N_1} \sum_{|w|2 \sim M_1 |z_0|^2} \sum_{|z_1|^2, |z_2|^2 \sim N_1 / |z_0|^2} S_{(1)}^1 (w', z_1) S_{(1)}^1 (w', z_2)
\]

\[
\ll X^{1/2} W^{1/2} (\log X) \sup_{w \leq N_1} \sum_{b_1, b_2 \ll X^{1/4}} \sum_{|z_1|^2, |z_2|^2 \sim N_1 / P} 1_{\Delta' \not\mid z_2 (b_1^2 + 1) - z_1 (b_2^2 + 1) \neq 0}.
\]

Let \(z_i := r_i + is_i\). Then we have

\[
\Delta' \mid s_2 (b_1^2 + 1) - s_1 (b_2^2 + 1) \quad \text{and} \quad \Delta' \mid r_2 (b_1^2 + 1) - r_1 (b_2^2 + 1)
\]

and \(z_2 (b_1^2 + 1) - z_1 (b_2^2 + 1) \neq 0\) so that by symmetry we may assume \(s_2 (b_1^2 + 1) - s_1 (b_2^2 + 1) \neq 0\) and thus by symmetry we may also assume that \(s_2 \neq 0\). For a given \(\Delta', s_1, s_2\) we have

\[
\sum_{r_1, r_2 \ll N_1^{1/2} / P^{1/2}} \sum_{r_1 s_2 - r_2 s_1 = \Delta' s_2} 1 \ll \sum_{r_2 \ll N_1^{1/2} / P^{1/2}} \ll N_1^{1/2} (s_1, s_2) / P^{1/2} s_2.
\]

Hence, denoting \(s_0 = (s_1, s_2)\) and \(s'_i = s_0 s_i\), we get

\[
\ll X^{1/2} N_1^{1/2} W^{1/2} \log X / P^{1/2} \sum_{s_0 \ll N_1^{1/2}} \tau(s_0) \sum_{s'_1, s'_2 \ll N_1^{1/2} / P^{1/2} / s_0} \frac{1}{s_2'}
\]

\[
\sum_{b_1, b_2 \ll X^{1/4}} \tau(s'_2 (b_1^2 + 1) - s'_1 (b_2^2 + 1) \neq 0).
\]

By Lemma \[^{11}\] there exists \(d \mid s'_2 (b_1^2 + 1) - s'_1 (b_2^2 + 1)\) such that \(d \ll X^\varepsilon\) and

\[
\tau(s'_2 (b_1^2 + 1) - s'_1 (b_2^2 + 1)) \ll \tau(d)^{O(1)}.
\]
Thus, the last expression is bounded by
\[ \ll \frac{X^{1/2}N_{1}^{1/2}W^{1/2} \log X}{P^{1/2}} \sum_{s_{0} \ll N_{1}^{1/2}} \tau(s_{0}) \sum_{d \ll X^{s}} \frac{\tau(d)O_{s}(1)}{d} \sum_{s'_{1},s'_{2} \ll N_{1}^{1/2}p^{-1/2}/s_{0}} \frac{1}{s'_{2}} \sum_{b_{1},b_{2} \ll X^{1/4}} \frac{1}{s'_{2}(b_{1}^{2}+1)-s'_{1}(b_{2}^{2}+1)} \equiv 0(d) \]
\[ \ll \frac{XN_{1}^{1/2}W^{1/2} \log X}{P^{1/2}} \sum_{s_{0} \ll N_{1}^{1/2}} \tau(s_{0}) \sum_{d \ll X^{s}} \frac{\tau(d)O_{s}(1)}{d} \sum_{s'_{1},s'_{2} \ll N_{1}^{1/2}p^{-1/2}/s_{0}} \frac{\tau(s'_{2})O_{s}(1)}{s'_{2}} \]
\[ \ll W^{-1/2}XN_{1} \log O_{1}X \ll_{C} XN_{1} \log^{-C} X, \]
where we have used the bounds
\[ \sum_{b_{1},b_{2} \ll X^{1/4}} \frac{1}{s'_{2}(b_{1}^{2}+1)-s'_{1}(b_{2}^{2}+1) \equiv 0(d)} \ll X^{1/2}(d, s'_{2})/d \]
and
\[ \sum_{d \sim D} \tau(d)O_{1}(d, s) \leq \sum_{v|s} \sum_{d \sim D} \tau(d)O_{1} \ll \tau(s)O_{1}D \log O_{1}D. \]

Similar arguments again apply when \( j = 2 \) or \( k = 2 \).

6.4. The off-diagonal for \( U_{1}^{(1)} \). Let \( U_{11}^{(1)} \) denote the part of \( U_{1}^{(1)} \) where \( \Delta \neq 0 \) and \( (z_{1}, z_{2}) = 1 \), so that we have
\[ (z_{1}, z_{2}) = (z_{1}, \overline{z}_{1}) = (z_{2}, \overline{z}_{2}) = (\Delta, |z_{2}z_{1}|^{2}) = 1. \]
To see the last equality, suppose that \( p|\Delta, |z_{1}|^{2} \). Then using \( (z_{1}, \overline{z}_{1}) = 1 \) we have \( p \equiv 1 \pmod{4} \) so that \( p = \pi \overline{\pi} \) for some Gaussian prime \( \pi|z_{1} \). By \( (z_{1}, z_{2}) = 1 \) we have \( \pi \not| z_{2} \), so that \( \overline{\pi}|z_{2}\overline{z}_{1} \) but \( p \not| z_{2}\overline{z}_{1} \). This is a contradiction, since \( p|\text{Im}(z_{2}\overline{z}_{1}) \) and \( \overline{\pi}|z_{2}\overline{z}_{1} \) imply that \( p|\text{Re}(z_{2}\overline{z}_{1}) \) and thus \( p|z_{2}\overline{z}_{1} \). Similarly we see that \( (\Delta, |z_{2}|^{2}) = 1. \)

In particular, we have
\[ 1 \leq |\Delta| \ll N_{1}. \]
Recall that
\[ i\Delta w = z_{2}(b_{1}^{2}+1) - z_{1}(b_{2}^{2}+1). \]
That is, \( w \) is fixed for given \( z_{1}, z_{2}, b_{1}, b_{2} \), and we have
\[ z_{2}(b_{1}^{2}+1) \equiv z_{1}(b_{2}^{2}+1) \pmod{(|\Delta|)}. \]
Note that \( (\overline{z}_{1} \overline{z}_{2}) \) denoting the complex conjugate
\[ a = a(z_{1}, z_{2}) := z_{2}/z_{1} = \frac{\overline{z}_{1}z_{2}}{|z_{1}|^{2}} \equiv \frac{r_{1}r_{2}+s_{1}s_{2}}{r_{1}^{2}+s_{1}^{2}} \pmod{|\Delta|}. \]
is a congruence in $\mathbb{Z}$ despite the fact that $z_1, z_2 \in \mathbb{Z}[i]$, and note that $(\alpha, \Delta) = 1$. Hence, we have

$$U_{11}^{(1)} = \sum_{|z_1|^2 | z_2|^2 \sim N_1} \beta_{z_1} \beta_{z_2} \sum_{\begin{subarray}{c} b_1, b_2 \\ \Delta \neq 0 \\ (z_1, z_2) = 1 \end{subarray}} b_1 b_2 F_{M_1} \left( \frac{|z_2(b_1^2 + 1) - z_1(b_2^2 + 1)|}{\Delta} \right).$$

Let $\Delta_2$ denote the largest powerful divisor of $\Delta$ and let $Y = (\log X)^{C'}$ for some large $C' > 0$. Note that $(\alpha - 1, \Delta) = (z_2 - z_1, \Delta)$. We now wish to discard the part where $\Delta_2 > Y$ or $(z_2 - z_1, \Delta) > Y$ and write

$$U_{11}^{(1)} = U_{11 \leq}^{(1)} + U_{11 >}^{(1)}$$

with

$$U_{11 \leq}^{(1)} := \sum_{|z_1|^2 | z_2|^2 \sim N_1} \beta_{z_1} \beta_{z_2} \sum_{\begin{subarray}{c} b_1, b_2 \\ \Delta \neq 0 \\ (z_1, z_2) = 1 \\ \max(\Delta_2, (z_2 - z_1, \Delta)) \leq Y \end{subarray}} b_1 b_2 F_{M_1} \left( \frac{|z_2(b_1^2 + 1) - z_1(b_2^2 + 1)|}{\Delta} \right),$$

$$U_{11 >}^{(1)} := \sum_{|z_1|^2 | z_2|^2 \sim N_1} \beta_{z_1} \beta_{z_2} \sum_{\begin{subarray}{c} b_1, b_2 \\ \Delta \neq 0 \\ (z_1, z_2) = 1 \\ \max(\Delta_2, (z_2 - z_1, \Delta)) > Y \end{subarray}} b_1 b_2 F_{M_1} \left( \frac{|z_2(b_1^2 + 1) - z_1(b_2^2 + 1)|}{\Delta} \right).$$

We will show in Section 6.4.5 by crude estimates that

$$(6.3) \quad U_{11 \leq}^{(1)} \ll C X N_1 \log^{-C} X.$$

For $U_{11 \leq}^{(1)}$ we further separate terms near the diagonal by writing

$$U_{11 \leq}^{(1)} = U_{11}^{(1)} + U_{11_2}^{(1)},$$

where $U_{11_2}$ is the part where for $\delta = \log^{-C_1} X$ with some large $C_1 > 0$

$$\arg z_1 = \arg z_2 + O(\sqrt{\delta}) \pmod{\pi}.$$ 

We will show in Section 6.4.4 that for $C_1$ large enough in terms of $C$ we have

$$(6.4) \quad U_{11_2}^{(1)} \ll C X N_1 \log^{-C} X.$$ 

For now we consider $U_{11_1}^{(1)}$, where $|\arg z_1 - \arg z_2 \pmod{\pi}| > \sqrt{\delta}$. To simplify the evaluation we split the sums smoothly into finer-than-dyadic intervals. Let $\delta = \log^{-C_1} X$ be as above and let $\psi$ be a $C'^*\text{-smooth}$ function as in Section 3.1. We will apply the procedure of Section 3.1 six times, to introduce smooth weights for the variables $z_1, z_2, \arg z_1, \arg z_2, b_1, b_2$, which will cost us a factor of $\delta^{-6}\log^2 X$. For any $Y \geq 1$ let $\psi_Y(y) := \psi(y/Y)$. Let $N_{11}, N_{12} \asymp N_1, \theta_1, \theta_2 \in [\pi, 5\pi)$, and $B_1, B_2 \ll X^{1/4}$ and denote

$$\beta_z^{(2)} := \beta_z \psi_{N_{11}}(|z|^2) \psi_{B_z}(\arg z + 2\pi).$$

For $z_1, z_2$ such that $\beta_{z_1}^{(1)} \beta_{z_2}^{(2)} \neq 0$ we have using $|\theta_1 - \theta_2 \pmod{\pi}| > \sqrt{\delta}$

$$|\Delta| = |z_1 z_2| |\sin(\arg z_2 - \arg z_1)| = (N_{11} N_{12})^{1/2} |\sin(\theta_2 - \theta_1)| (1 + O(\delta)) \gg \sqrt{\delta} N.$$
Hence, for any fixed tuple \((B, N, \theta)\) there is a constant \(F(B, N, \theta)\) such that for all \(z, b\) in the support of \(\beta^{(1)}_{z_1} \beta^{(2)}_{z_2} \psi_{b_1}(b_1) \psi_{b_2}(b_2)\) we have (using \(|\Delta| \gg \sqrt{\delta N}\))

\[
\begin{align*}
    b_1 b_2 F_{M_1} \left( \left| \frac{z_2 (b_1^2 + 1) - z_1 (b_2^2 + 1)}{\Delta} \right|^2 \right) = B_1 B_2 \left( F(B, N, \theta) + O(\sqrt{\delta}) \right).
\end{align*}
\]

Accordingly we write

\[
U_{111}^{(1)} = U_{111}^{(1)} + O(\sqrt{\delta} U_{1112}^{(1)}),
\]

where in \(U_{1112}^{(1)}\) we have used the triangle inequality to put absolute values around \(\beta\).

We will show in Section 6.4.3 that

\[
U_{1112}^{(1)} \ll X N_1 \log^{O(1)} X,
\]

so that by taking \(C_1\) to be large enough in terms of \(C\) we get

\[
\sqrt{\delta} U_{111}^{(1)} \ll C X N_1 \log^{-C} X.
\]

Before this we consider the main term \(U_{1111}^{(1)}\), for which we need to bound sums of the form

\[
U_1^{(1)}(B, N, \theta) := B_1 B_2 \sum_{\substack{x_1, x_2 \in \mathbb{Z}^2 \mid \Delta \neq 0 \mid (z_1, z_2) = 1}} \beta^{(1)}_{z_1} \beta^{(2)}_{z_2} \psi_{b_1}(b_1) \psi_{b_2}(b_2).
\]

Since the cost from introducing smooth weights is bounded by \(\ll \delta^6 \log^6 X = \log^{O(1)} X\), to handle \(U_{1111}^{(1)}\) it suffices to show that for any \(C > 0\)

\[
U_1^{(1)}(B, N, \theta) \ll C X N_1 \log^{-C} X.
\]

Let \(H_j := X^\epsilon |\Delta|/B_j\) and define

\[
\begin{align*}
    N_1(a; \Delta) &:= |\{x_1, x_2(\Delta) : x_1^2 + 1 = a(x_2^2 + 1) (|\Delta|)\}|, \\
    S_1(a, h; |\Delta|) &:= \sum_{x_1, x_2(\Delta) : x_1^2 + 1 = a(x_2^2 + 1) (|\Delta|)} e_{|\Delta|}(h \cdot x),
\end{align*}
\]

and

\[
\hat{U}_1^{(1)}(B, N, \theta) := B_1 B_2 \sum_{\substack{x_1, x_2 \in \mathbb{Z}^2 \mid \Delta \neq 0 \mid (z_1, z_2) = 1}} \frac{|\beta^{(1)}_{z_1} \beta^{(2)}_{z_2}|}{H_1 H_2} \sum_{h_1, h_2 \neq (0, 0)} |S_1(a, h; \Delta)|.
\]

Then applying the Poisson summation formula (Lemma 7) twice we get

\[
U_1^{(1)}(B, N, \theta) = B_2^2 \sum_{\substack{x_1, x_2 \in \mathbb{Z}^2 \mid \Delta \neq 0 \mid (z_1, z_2) = 1}} \frac{\beta^{(1)}_{z_1} \beta^{(2)}_{z_2}}{\Delta^2} \hat{\psi}(0)^2 N_1(a; \Delta) + O(X^2 \hat{U}_1^{(1)}(B, N, \theta)) + O(\epsilon)(X^{-10}).
\]

Note that here \(\hat{\psi}(0) \ll \delta\).
6.4.1. Bounding the main term. We have to show that for any $C > 0$

\begin{equation}
\delta^2 B_1^2 B_2^2 \sum_{\substack{z_1, z_2 \\ \Delta \neq 0 \\ (z_1, z_2) = 1}} \beta_{z_1}^{(1)} \beta_{z_2}^{(2)} \frac{1}{\Delta} N_1(a; \Delta) \ll_C X N \log^{-C} X.
\end{equation}

Recall that $(a, \Delta) = 1$ for all $z_1, z_2$ appearing in the above sum and note that $(a-1, \Delta) = (z_2 - z_1, \Delta)$. Using Lemma 23 with $Y = \log^C X$ as above, noting that $\max\{\Delta_2, (z_2 - z_1, \Delta)\} \leq Y$, the left-hand side in (6.8) is bounded by $T_1 + O(T_2)$ for

\begin{align*}
T_1 &:= \delta^2 B_1^2 B_2^2 \sum_{\substack{z_1, z_2 \\ \Delta \neq 0 \\ (z_1, z_2) = 1}} \beta_{z_1}^{(1)} \beta_{z_2}^{(2)} \frac{1}{\Delta} \sum_{d \leq Y} \frac{\varepsilon_d(a)}{d}, \\
T_2 &:= Y^{-1} \delta^2 B_1^2 B_2^2 \sum_{\substack{z_1, z_2 \\ \Delta \neq 0 \\ (z_1, z_2) = 1}} |\beta_{z_1}^{(1)} \beta_{z_2}^{(2)}| \frac{1}{\Delta} \tau(\Delta)^{O(1)}.
\end{align*}

Using Lemma 23 we see (taking $Y$ to be sufficiently large power of $\log X$)

\begin{equation*}
T_2 \ll_C B_1^2 B_2^2 N_1 \log^{-C} X \ll_C X N_1 \log^{-C} X.
\end{equation*}

In $T_1$ the condition $d|\Delta$ is equivalent to $z_2 \equiv a z_1 (d)$ for some $a (d)$, so that

\begin{equation*}
T_1 = \delta^2 B_1^2 B_2^2 \sum_{d \leq Y} \sum_{(a, d) = 1} \frac{\varepsilon_d(a)}{d} \sum_{\substack{z_2 \equiv a z_1 (d) \\ \Delta \neq 0 \\ (z_1, z_2) = 1}} \beta_{z_1}^{(1)} \beta_{z_2}^{(2)} \frac{1}{\Delta}.
\end{equation*}

The condition $(z_1, z_2) = 1$ can be dropped with a negligible error term by trivial bounds, recalling that $\beta(n)$ is supported on $(n, P(W)) = 1$. Similarly, the condition $\max\{\Delta_2, (z_2 - z_1, \Delta)\} \leq Y$ may also be dropped with a negligible error term by crude bounds.

Recall that in the support of $\beta_{z_1}^{(1)} \beta_{z_2}^{(2)}$ we have for some constant $\sqrt{\delta} N_1 \ll D \leq 2N_1$

\begin{equation*}
\Delta = D(1 + O(\sqrt{\delta})).
\end{equation*}

The contribution from $O(\sqrt{\delta})$ is bounded by trivial bounds, recalling that $\beta_2^{(f)}$ are supported on small polar boxes. Applying (1.3) (the Siegel-Walfisz property) we get

\begin{equation*}
T_1 \ll_C B_1^2 B_2^2 N_1 (\delta^{-1/2} \log^{-C} X + \delta^6 \sqrt{\delta}) \ll_C X N_1 (\log^{-C} X + \delta^6 \sqrt{\delta}),
\end{equation*}

which is sufficient for (6.7) since the cost from introducing the smooth weights was $\delta^{-6}$.

Remark 4. At first it may seem surprising that in our problem the evaluation of the main term is vastly easier than in the situation in [6]. This is because here with $N(a; \Delta)$ we are computing solutions to a non-singular equation. This allowed us to bound the large moduli $d$ by using the Hasse-Weil bound (Lemma 14).
6.4.2. Bounding the non-zero frequencies. It suffices to show that for $N_1 \ll X^{1/3-\eta}$ we have for some $\eta > 0$

$$U^{(1)}_1(B, N, \theta) \ll X^{1-\eta}N_1.$$ 

For $h_1 = 0$ we apply Lemma [20] with $h_2 \neq 0$ to get a bound

$$B_1B_2 \sum_{z_{1,2}} \frac{1}{H_1H_2} \sum_{0 < |h_2| \leq H_2} |S_1(a, 0, h_2; \Delta)| \ll YO(1)B_1B_2 \sum_{z_{1,2}} \frac{|\Delta|^{O(1)}}{\Delta^{1/2}} \sum_{0 < |h_2| \leq H_2} \tau(h_2, \Delta_1)$$

Using $H_j = |\Delta|/B_j$ and Lemma [23] the last expression is bounded by

$$\ll YO(1) \sum_{z_{1,2}} \frac{|\Delta|^{O(1)}}{\Delta^{1/2}} B_1B_2 \ll YO(1)B^2B_2 \sum \frac{\tau(\Delta)^{O(1)}}{\Delta^{1/2}} \ll YO(1)B^2_2N_1^{3/2} \ll X^{1-\eta}N_1,$$

since $B_j \ll X^{1/4}$ and $N \ll X^{1/3-\eta} < X^{1/2-\eta}$. Similarly, we get a sufficient bound for the contribution from $h_2 = 0$ with $h_1 \neq 0$.

For the contribution from $(h_1, h_2) \neq 0$ we get by Lemma 8 a bound

$$\ll YO(1) \sum_{z_{1,2}} \frac{\tau(\Delta)^{O(1)}}{\Delta^{1/2}} \sum_{0 < |h_2| \leq H_2} ((h_1, \Delta_1) + (h_1 + h_2, \Delta_1)^{1/2} + (h_1 - h_2, \Delta_1)^{1/2})$$

$$\ll YO(1)B_1B_2 \sum \frac{\tau(\Delta)^{O(1)}}{\Delta^{1/2}} \left( \frac{|\Delta|}{\Delta^{1/2}} + \frac{|\Delta|}{H_1} \right) \ll B_1B_2 \sum \frac{\tau(\Delta)^{O(1)}}{\Delta^{1/2}} \left( \frac{|\Delta|}{\Delta^{1/2}} + B_1 \right).$$

By Lemma [23] the last expression is at most

$$\ll X^\varepsilon(B_1B_2N_1^{5/2} + B^2B_2N_1^{3/2}) \ll X^\varepsilon(X^{1/2}N_1^{5/2} + X^{3/4}N_1^{3/2}) \ll X^{1-\eta}N_1$$

by using $N_1 \ll X^{1/3-\eta}$.

Remark 5. Morally speaking we have in the above applied the bound $S_1(a; \Delta) \ll \Delta^{1/2}$. The corresponding exponential sum in [6] is bounded in terms of Ramanujan sums, so that in there one morally gets a bound $\ll 1$ for the exponential sums. This loss is the reason why our Type II range is narrower than in [6].

6.4.3. Bounding the error from removing cross-conditions. We now explain how to get [6.6]. To bound $U^{(1)}_{1112}$ we consider

$$V^{(1)}_1(B, N, \theta) := B_1B_2 \sum_{|z_{1,2}|^2 \sim N_1 \Delta \neq 0} |\beta^{(1)}_{z_{1}}||\beta^{(2)}_{z_{2}}| \sum_{b_{1,2} \Delta \neq 0, \Delta_1 = 1} \psi_{B_1}(b_1) \psi_{B_2}(b_2)$$

$$\sum_{b_{1,2} \Delta \neq 0, \Delta_1 = 1} \psi_{B_1}(b_1) \psi_{B_2}(b_2)$$
For fixed $B$ and $N$ we can write
\[ V_{1}^{(1)}(B, N, \theta) \ll B_{1}B_{2} \sum_{|z_1|^2,|z_2|^2 \sim N_1 \atop \Delta \neq 0 \atop \max \{ \Delta_2, |z_2 - z_1, \Delta \} \leq Y} \psi_{N_1}(|z_1|^2)\psi_{N_2}(|z_2|^2)\psi_{\theta_1}(\arg z_1 + 2\pi)\psi_{\theta_2}(\arg z_2 + 2\pi) \]
\[ \sum_{b_1, b_2 \atop b_2^2 + 1 = a(b_1^2 + 1) (|\Delta|)} \psi_{1}(b_1)\psi_{2}(b_2). \]

Then applying Poisson summation and carrying out the same arguments as in Sections 6.4.1 and 6.4.2, replacing the use of the Siegel-Walfisz property with $\beta_{z_1}^{(1)}\beta_{z_2}^{(2)}$ by a trivial point count with $z_1$ and $z_2$ in small polar boxes, we can evaluate the sum to obtain for any $C > 0$
\[ V_{1}^{(1)}(B, N, \theta) \ll_{C} (\delta^6 X N_1 \log^{O(1)} X + X N_1 \log^{C} X). \]

Since the integration over the tuple $(B, N, \theta)$ is weighted by $\delta^{-6} \log^{O(1)} X$ (as in Section 3.1), we get
\[ U_{112}^{(1)} \ll C X N_1 \log^{O(1)} X. \]

6.4.4. Bounding the contribution from $\arg z_1 = \arg z_2 + O(\sqrt{\delta})$ (mod $\pi$). To complete the proof for $a_n^{(1)}$ we need to show the bound (6.3) for the contribution from $U_{112}^{(1)}$, where $\arg z_1 = \arg z_2 + O(\sqrt{\delta})$ (mod $\pi$). The idea is similar as in the previous section but we cannot ignore the weight
\[ F_{M_1} \left( \frac{|z_2(b_1^2 + 1) - z_1(b_2^2 + 1)|^2}{\Delta} \right), \]
which now restricts the variables $b_j$ to a narrow subset. We split $U_{112}^{(1)}$ dyadically according to the size of $\Delta \sim \delta_1 N$ to get
\[ U_{112}^{(1)} \ll \sum_{\delta_1 = 2^{-j} \ll \sqrt{\delta}} U(\delta_1), \]
where
\[ U(\delta_1) = \sum_{|z_1|^2,|z_2|^2 \sim N_1 \atop \max \{ \Delta_2, |z_2 - z_1, \Delta \} \leq Y \atop \Delta \sim \delta_1 N} |\beta_{z_1}| |\beta_{z_2}| \sum_{b_1, b_2 \atop b_2^2 + 1 = a(b_1^2 + 1) (\Delta)} b_1 b_2 F_{M_1} \left( \frac{|z_2(b_1^2 + 1) - z_1(b_2^2 + 1)|^2}{\Delta} \right). \]

We introduce finer-than-dyadic weights for $b_2$ using Section 3.1 with $\delta_1$, denoting $\psi = \psi^{(\delta_1)}$, to get
\[ U(\delta_1) \ll \int_{1/2}^{2X^{1/4}} |V(\delta_1, B_1)| \frac{dB_1}{B_1}, \]
where
\[ V(\delta_1, B_1) = \frac{1}{\delta_1} \sum_{|z_1|^2,|z_2|^2 \sim N_1 \atop \max \{ \Delta_2, |z_2 - z_1, \Delta \} \leq Y \atop \Delta \sim \delta_1 N} |\beta_{z_1}| |\beta_{z_2}| \sum_{b_1, b_2 \atop b_2^2 + 1 = a(b_1^2 + 1) (\Delta)} b_1 b_2 \psi^{(\delta_1)}(b_1) F_{M_1} \left( \frac{|z_2(b_1^2 + 1) - z_1(b_2^2 + 1)|^2}{\Delta} \right). \]
The weight $F_{M_1}$ is supported on

$$|b_2^2 + 1 - z_2/ z_1(b_1^2 + 1)| \ll \delta_1 M_1^{1/2} N_1^{1/2} = \delta_1 X^{1/2},$$

Denoting $B_2^2 := -1 + z_2/ z_1(B_1^2 + 1)$, for $b_1 = B_1(1 + O(\delta_1))$ this is contained in

$$|b_2^2 - B_2^2| \ll \delta_1 X^{1/2} + \delta_1 B_1^2 \ll \delta_1 X^{1/2},$$

which is contained in

$$|b_2 - B_2| \ll \delta_1 X^{1/2}/B_1.$$

Thus, if we let $\psi^{(\delta_2)}_{B_2}(b_2)$ denote a suitable smooth majorant to this with

$$\delta_2 = \min \left\{ 1, \delta_1 X^{1/2}/B_1 B_2 \right\},$$

we get

$$V(\delta_1, B_1) \ll \frac{1}{\delta_1} \sum_{|z_1|, |z_2|, \delta = \delta_1 N_1} |\beta_{z_1}| |\beta_{z_2}| B_1 B_2 \sum_{b_1, b_2} \psi^{(\delta_1)}_{B_1}(b_1) \psi^{(\delta_2)}_{B_2}(b_2).$$

We now evaluate the inner sum by Poisson summation (Lemma 7) to get

$$V(\delta_1, B_1) = \frac{1}{\delta_1} \sum_{|z_1|, |z_2|, \delta = \delta_1 N_1} |\beta_{z_1}| |\beta_{z_2}| B_1^2 B_2^2 \frac{\hat{\psi}(\delta_1)(0) \hat{\psi}(\delta_2)(0)}{\Delta^2} N_1(a; \Delta) + X^2 \hat{V} + O_\epsilon(\hat{V}^{10})$$

The non-zero frequencies $\hat{V}$ are bounded just as before, making use of bound

$$\frac{\hat{\psi}(\delta_j)}{\delta_j} \ll \delta_j$$

to cancel the factor $\frac{1}{\delta_1}$ and $\Delta \sim \delta_1 N_1$ to cancel the factor $\frac{1}{\delta_1}$. We get similarly as in Section 6.4.2

$$\hat{V} \ll X^\epsilon(B_1 B_2 N_1^{5/2} + B_1^2 B_2 N_2^{3/2}) \ll X^\epsilon(X^{1/2} N_1^{5/2} + X^{3/4} N_1^{3/2}) \ll X^{1-\eta} N_1$$

by using $N_1 \ll X^{1/3-\eta}$.

For the main term we get using the trivial bound $N_1(a; \Delta) \ll \tau(\Delta)^{O(1)}|\Delta|$,

$$\frac{1}{\delta_1} \sum_{|z_1|, |z_2|, \delta = \delta_1 N_1} |\beta_{z_1}| |\beta_{z_2}| B_1^2 B_2^2 \frac{\hat{\psi}(\delta_1)(0) \hat{\psi}(\delta_2)(0)}{\Delta^2} N_1(a; \Delta) \ll \delta_1 X N_1^2 \ll \sqrt{\delta} X N_1^2,$$

which is $\ll X N_1 \log^{-C} X$ once we choose a sufficiently large $C_1$ in $\delta = \log^{-C_1} X$.

Remark 6. When we split $b_1$ into intervals of length $\delta_1 B_1$ it can happen that $\delta_1 B_1 < 1$ for small $\delta_1$, so that the sum over $b_1$ may be empty. This is not a problem since we are not trying to show that the error term $\hat{V}$ from the Poisson summation is smaller than the main term, only that both terms are smaller than $\ll X N_1 \log^{-C} X$. 

6.4.5. Bounding \( U_{11}^{(1)} \). Recall that we are aiming to show (6.3) for

\[
U_{11}^{(1)} = \sum_{\substack{|z_1|^2, |z_2|^2 \sim N_1 \\ \Delta \neq 0 \\ (z_1, z_2) = 1 \atop \max\{\Delta_2, (z_2-z_1, \Delta)\} > Y}} \beta_{z_1} \beta_{z_2} \sum_{b_1, b_2} b_1 b_2 F_{M_1} \left( \frac{|z_2(b_1^2+1) - z_1(b_2^2+1)|^2}{\Delta} \right)
\]

\[\ll X^{1/2} \sum_{\substack{|z_1|^2, |z_2|^2 \sim N_1 \\ \Delta \neq 0 \\ (z_1, z_2) = 1 \atop \Delta > Y}} \sum_{b_1, b_2 < X^{1/4}} 1 + X^{1/2} \sum_{\Delta \neq 0} \sum_{D \mid \Delta} \sum_{b_1, b_2 < X^{1/4}} 1 \]

\[=: W_{11}^{(1)} + W_{11}^{(1)} \]

where \( \Delta_2 \) is the powerful part of \( \Delta \). For both of the terms we will apply a similar argument as in Section 6.3. We first consider \( W_{11}^{(1)} \), which is the more difficult of the two. Denoting \( z_j = r_j + is_j \), since \( \Delta = r_1 s_2 - r_2 s_1 \neq 0 \), by symmetry we may assume that \( s_1 \neq s_2 \) and thus by symmetry we can further assume that \( s_2 \neq 0 \). Let \( D_0 | \Delta \) stand for \((z_2 - z_1, \Delta)\), so that we have \( D_0 | s_2 - s_1 \) and \( D_0 | r_2 - r_1 \). Hence,

\[
(6.9) \quad W_{11}^{(1)} \ll X^{1/2} \sum_{D_0 > Y} \sum_{b_1, b_2 < X^{1/4}} 1 \sum_{\substack{|s_1|, |s_2| < N_1^{1/2} \\ s_1 \neq s_2 \neq 0}} \sum_{\Delta \neq 0} \sum_{D_0 | \Delta} \sum_{b_1, b_2 < X^{1/4}} 1 \ll 1 + \frac{N_1^{1/2}(D_0 s_2, (s_2 - s_1))}{D_0 s_2}
\]

For fixed \( s_1, s_2, \Delta \) with \( s_2 \neq 0, |s_2| \ll N_1^{1/2} \), denoting \( r_1 = r_2 + r_3 D_0 \) we see that

\[
\sum_{\substack{|r_1|, |r_2| \ll N_1^{1/2} \\ r_2 s_2 - r_2 s_1 = \Delta}} 1 \ll \sum_{\substack{|r_2| \ll N_1^{1/2} \\ r_2 (s_2 - s_3) \equiv \Delta \pmod{D_0}}} 1 \ll 1 + \frac{N_1^{1/2}(D_0 s_2, (s_2 - s_1))}{D_0 s_2}
\]

Thus, denoting \( s_0 = (s_1, s_2) \) and \( s_j = s_0 s_j' \) we have

\[
\frac{(D_0 s_2, (s_2 - s_1))}{D_0 s_2} = \frac{(D_0, (s_2' - s_1'))}{D_0 s_2'},
\]

so that

\[
W_{11}^{(1)} \ll X^{1/2} \sum_{D_0 > Y} \sum_{b_1, b_2 < X^{1/4}} \sum_{s_0 \ll N_1^{1/2}} \tau(s_0)
\]

\[
(6.10) \quad \sum_{\substack{|s_1|, |s_2| \ll N_1^{1/2} \div s_0 \\ (s_1', s_2') = 1 \\ s_1' \neq s_2' \neq 0}} \sum_{s_0 s_2'(b_1^2+1) \equiv s_0 s_2'(b_2^2+1) \pmod{D_0}} \tau(s_2'(b_1^2+1) - s_2'(b_2^2+1)) = \tau(1) D_0^{O(1)}
\]

By using Lemma 11 to find a divisor \( d \leq X^\varepsilon \) of

\[
\frac{1}{D_0} (s_0 s_2'(b_1^2+1) - s_0 s_2'(b_2^2+1))
\]

with

\[
\tau \left( \frac{1}{D_0} (s_0 s_2'(b_1^2+1) - s_0 s_2'(b_2^2+1)) \right) \ll \tau(d)^{O(1)}
\]
we get

\[ W_{11}^{(1)} \ll X^{1/2} \sum_{D_0 > Y} \tau(D_0) \sum_{d \leq X^s} \tau(d) \sum_{s_0 \ll N_1^{1/2}} \tau(s_0) \sum_{|s'_1|, |s'_2| \ll N_1^{1/2}/s_0} \left(1 + \frac{N_1^{1/2} (D_0, (s'_2 - s'_1))}{D_0 s'_2} \right) \sum_{b_1, b_2 \ll X^{1/4}} 1. \]

Using the trivial bound

\[ \sum_{s_0 s'_2 (b_2^2 + 1) \equiv s_0 s'_1 (b_2^2 + 1) (dD_0)} 1 \ll \tau(D_0) \tau(d) \left(1 + \frac{X^{1/4}(dD_0, s_0 s'_2)}{dD_0} \right) \]

we get

\[ W_{11}^{(1)} \ll X^{1/2} \sum_{D_0 > Y} \tau(D_0)^2 \sum_{d \leq X^s} \tau(d) \sum_{s_0 \ll N_1^{1/2}} \tau(s_0) \sum_{|s'_1|, |s'_2| \ll N_1^{1/2}/s_0} \left(1 + \frac{N_1^{1/2} (D_0, (s'_2 - s'_1))}{D_0 s'_2} \right) X^{1/4} \left(1 + \frac{X^{1/4}(dD_0, s_0 s'_2)}{dD_0} \right) \]

\[ \ll X^{3/4} \sum_{d \leq X^s} \tau(d) \sum_{s_0 \ll N_1^{1/2}} \tau(s_0) \sum_{|s'_1|, |s'_2| \ll N_1^{1/2}/s_0} \left(1 + \frac{N_1^{1/2} (D_0, (s'_2 - s'_1))}{dD_0} \right) \frac{N_1^{1/2}}{s'_2} \frac{X^{1/4}(dD_0, s_0 s'_2)}{dD_0 s'_2} (D_0, (s'_2 - s'_1)) (dD_0, s_0 s'_2).
\]

Note that by \((s'_1, s'_2) = 1\) and \(D_0 | s_0 (s'_2 - s'_1)\) we have

\((dD_0, s_0 s'_2) \leq (d, s_0 s'_2)(D_0, s_0),\)

and

\((D_0, (s'_2 - s'_1)) (D_0, s_0 s'_2) = (D_0, (s'_2 - s'_1))(D_0, s_0) \leq D_0 (s_0, (s'_2 - s'_1))\)

so that

\[ W_{11}^{(1)} \ll X^{3/4} \sum_{d \leq X^s} \tau(d) \sum_{s_0 \ll N_1^{1/2}} \tau(s_0) \sum_{|s'_1|, |s'_2| \ll N_1^{1/2}/s_0} \left(1 + \frac{X^{1/4}(dD_0, s_0 s'_1)}{dD_0} \right) \frac{N_1^{1/2}}{s'_2} \frac{X^{1/4}(d, s_0 s'_2)(s_0, (s'_2 - s'_1))}{dD_0 s'_2} \]

\[ \sum_{D_0 > Y} \tau(D_0)^2 \left(1 + \frac{X^{1/4}(dD_0, s_0 s'_1)(D_0, s_0)}{dD_0} \right) \frac{N_1^{1/2}}{s'_2} + \frac{N_1^{1/2} X^{1/4}(d, s_0 s'_2)(s_0, (s'_2 - s'_1))}{dD_0 s'_2}. \]
By using the bounds
\[
\sum_{D_0 > Y \atop D_0 \mid s_0 (s_2' - s_1')} \frac{\tau(D_0)^2 (D_0, s_0)}{D_0} \leq (Y^{-1/2} + 1_{s_0 > Y^{-1/2}}) \tau(s_0)^3 \tau(s_2' - s_1')^3
\]
and
\[
\sum_{D_0 > Y \atop D_0 \mid s_0 (s_2' - s_1')} \frac{\tau(D_0)^2}{D_0} \leq Y^{-1} \tau(s_0)^3 \tau(s_2' - s_1')^3
\]
we get
\[
W_{11}^{(1)} \ll X^{3/4} \sum_{d \leq X^\varepsilon} \tau(d) O(1) \sum_{s_0 \ll N_1^{1/2}} \tau(s_0)^4
\]
\[
\sum_{|s_1|, |s_2'| \ll N_1^{1/2} / s_0 \atop s_1 \neq s_2' \neq 0} \tau(s_2' - s_1')^3 \left(1 + \frac{N_1^{1/2}}{s_2'} \right) \frac{X^{1/4} (d, s_0 s_2')}{d} \left(Y^{-1/2} + 1_{s_0 > Y^{-1/2}}\right)
\]
\[
+ \frac{N_1^{1/2} X^{1/4} (s_0, (s_2' - s_1')) (d, s_0 s_2')}{Y ds_2'}
\]
\ll \varepsilon \cdot X^{3/4 + 2\varepsilon} N_1 + \frac{X N_1 \log O(1) X}{Y^{1/2}},
\]
where the last line follows by first computing the sum over \(d \leq X^\varepsilon\) using the bound
\[
\sum_{d \sim D} \tau(d)^O(1) (d, s) \leq \sum_{v \mid s} \sum_{d \sim D} \tau(d)^O(1) \ll \tau(s)^O(1) D \log O(1) D.
\]
Thus, by taking \(\varepsilon > 0\) small and \(Y = \log C' X\) with \(C' \gg C\) we get
\[
W_{11}^{(1)} \ll X N_1 \log^{-C} X.
\]
To bound \(V_{11}^{(1)}\), we note that similarly as in (6.9) we have
\[
V_{11}^{(1)} \ll X^{1/2} \sum_{Y < \Delta_2 \ll N_1} \sum_{b_1, b_2 \ll X^{1/4} \atop \Delta_2 \text{ powerful}} \sum_{|s_1|, |s_2| \ll N_1^{1/2} \atop s_1 \neq s_2 \neq 0} \sum_{\Delta \neq 0 \atop D_0 \mid \Delta} \sum_{r_{12} \ll N_1^{1/2} \atop \Delta_2 \equiv \Delta (b_1^2 + 1) \equiv s_1 (b_2^2 + 1) \Delta (s_2 (b_2^2 + 1) - s_1 (b_2^2 + 1))}
\]
And similar to (6.10) we get
\[
V_{11}^{(1)} \ll X^{1/2} \sum_{Y < \Delta_2 \ll N_1} \sum_{b_1, b_2 \ll X^{1/4} \atop \Delta_2 \text{ powerful}} \sum_{s_0 \ll N_1^{1/2}} \tau(s_0)
\]
\[
\sum_{|s_1'|, |s_2'| \ll N_1^{1/2} / s_0 \atop s_1' \neq s_2' \neq 0} \left(1 + \frac{N_1^{1/2}}{s_2'} \right) \tau(s_2' (b_1^2 + 1) - s_1' (b_2^2 + 1)).
\]
By using Lemma 11 to find a suitable divisor $d \leq X^\varepsilon$ of $\frac{1}{\Delta_2}(s_0 s'_2 (b_2^2 + 1) - s_0 s'_1 (b_1^2 + 1))$ and using the trivial bound for the sum over $b_1, b_2$ we get

$$V^{(1)}_{11} \ll X^{1/2} \sum_{Y < \Delta_2 \leq N_1} \frac{\tau(\Delta_2)}{\Delta_2} \sum_{d \leq X^\varepsilon} \frac{\tau(d)}{\log \Delta_2} \sum_{s_0 \ll N_1^{1/2}} \tau(s_0) \sum_{b_1, b_2 \ll X^{1/4}} \frac{1}{s_0 (b_1^2 + 1) s_1 (b_2^2 + 1)} (d \Delta_2)$$

$$\ll X^{1/2} \sum_{Y < \Delta_2 \leq N_1} \frac{\tau(\Delta_2)^2}{\Delta_2} \sum_{d \leq X^\varepsilon} \frac{\tau(d)}{\log \Delta_2} \sum_{s_0 \ll N_1^{1/2}} \tau(s_0)$$

$$\ll X^{3/4} \sum_{Y < \Delta_2 \leq N_1} \frac{\tau(\Delta_2)^2}{\Delta_2} \sum_{d \leq X^\varepsilon} \frac{\tau(d)}{\log \Delta_2} \sum_{s_0 \ll N_1^{1/2}} \tau(s_0)$$

Thus, we get

$$V^{(1)}_{11} \ll X^{3/4} (\log O(1) X) \sum_{Y < \Delta_2 \leq N_1} \frac{\tau(\Delta_2)}{\Delta_2} \left( X^{\varepsilon} N_1 + \frac{N_1 X^{1/2}}{\Delta_2} \right)$$

(6.11)

$$\ll X^{3/4 + 2\varepsilon} N_1^{3/2} + \frac{X N_1 \log O(1) X}{Y}.$$

Since $N_1 \leq X^{1/2 - 10\varepsilon}$, by taking $\varepsilon > 0$ small and $Y = \log C' X$ with $C' \gg C$ we get

$$V^{(1)}_{11} \ll X N_1 \log^{-C} X.$$

6.5. The off-diagonal for $U^{(2)}_1$. Here we apply the same arguments as with $U^{(1)}_{11}$, except that we count solutions to

$$c_1^3 + d_1^3 \equiv a(c_2^3 + d_2^3) \ (|\Delta|),$$

with $c_1, c_2, d_1, d_2 \ll X^{1/6}$, so that in place of Lemmas 15 and 20 we use Lemmas 18 and 22. First we separate the part where $\Delta_2 > Y$ by writing

$$U^{(2)}_{11} = U^{(2)}_{11, \leq} + U^{(2)}_{11, >}.$$
By similar arguments as in Section 6.4.5 in place of (6.11) we get

\[ U^{(2)}_{11} \ll X^{5/6 + 2\varepsilon} N_{2}^{3/2} + \frac{X N_{2} \log O(1)}{Y}, \]

which suffices since \( N_{2} \ll X^{1/3 - 10\varepsilon}. \) Note that it is not necessary to separate large values of \((z_{2} - z_{1}, \Delta)\) since the bounds from Section 5 we will use here do not care about the common factor \((a - 1, D)\).

For \( U^{(2)}_{11} \) the technical issues of removing the smooth cross-condition \( F_{M_{2}} \) and bounding the part \( \theta_{1} = \theta_{2} + O(\sqrt{\delta}) \) (mod \( \pi \)) are handled similarly as for \( U^{(1)}_{11} \) (cf. Sections 6.4.3 and 6.4.4), so that we need to consider sums of the form

\[ U^{(2)}_{1}(C, D, N, \theta) := \Omega(C_{1}, D_{1}) \Omega(C_{2}, D_{2}) \sum_{\beta_{z_{1}}^{(1)} \beta_{z_{2}}^{(2)}} \beta_{z_{1}}^{(1)} \beta_{z_{2}}^{(2)} \]

\[ \sum_{\frac{1}{x_{1} + x_{2} = a(x_{1}^{2} + x_{2}^{2}) (\Delta)}} \psi_{C_{1}}(c_{1}) \psi_{C_{2}}(c_{2}) \psi_{D_{1}}(d_{1}) \psi_{D_{2}}(d_{2}) \]

with \( C_{1}, C_{2}, D_{1}, D_{2} \ll X^{1/6} \). Let \( H_{1} := X^{\varepsilon} |\Delta| / C_{1}, H_{2} := X^{\varepsilon} |\Delta| / C_{2}, H_{3} := X^{\varepsilon} |\Delta| / D_{1}, \) and \( H_{4} := X^{\varepsilon} |\Delta| / D_{2}. \) Applying Poisson summation four times and denoting

\[ S_{2}(\mathbf{h}; \Delta) := \sum_{x_{1}, x_{2}, x_{3}, x_{4} (\Delta)} e_{\Delta} (\mathbf{h} \cdot \mathbf{x}), \]

the contribution from \((h_{1}, h_{2}, h_{3}, h_{4}) \neq 0\) is by Lemma 22 bounded by (using \((a - 1, \Delta) \leq Y \) and \((a, \Delta) = 1\))

\[ \Omega(C_{1}, D_{1}) \Omega(C_{2}, D_{2}) \sum_{\frac{1}{\Delta \neq 0}} \frac{1}{H_{1} H_{2} H_{3} H_{4}} \sum_{\frac{|\beta_{|H_{j}|} \leq H_{j}}{h_{1}, h_{2}, h_{3}, h_{4} \neq 0}} S_{2}((\mathbf{h}; \Delta)) \ll (X^{1/6}) \sum_{\frac{\arg z_{1} \neq \arg z_{2}}{\Delta \neq 0}} \frac{Y^{O(1)} \tau(\Delta)^{O(1)} \Delta^{3}}{\Delta} \ll X^{1/3 + \varepsilon} N^{4} \ll X^{1 - \eta} N, \]

since \( N \ll X^{2/9 - \eta}. \) For the main term (with \((h_{1}, h_{2}, h_{3}, h_{4}) = 0\))

\[ \Omega(C_{1}, D_{1}) \Omega(C_{2}, D_{2}) C_{1} C_{2} D_{1} D_{2} \sum_{\frac{\beta_{z_{1}}^{(1)} \beta_{z_{2}}^{(2)}}{\Delta \neq 0}} \frac{\beta_{z_{1}}^{(1)} \beta_{z_{2}}^{(2)} \psi(0)^{4}}{\Delta^{4}} N_{3}(a; \Delta) \]

we get a sufficient bound by using Lemma 18 and the Siegel-Walfisz property with main term 0 [13] similarly as in Section 6.4.1.
6.6. The off-diagonal for $U_2^{(j)}$. Here we can apply the same arguments as in the above sections, but the evaluation is of course much easier. We are counting $b_1, b_2$ which satisfy the simple equation (with $(a, \Delta) = 1$)
\[ b_2 = ab_1 (|\Delta|). \]
Here the lengths $B_j$ of $b_j$ satisfy $B_j > X^{1/2-\eta} > N X^\eta \geq |\Delta| X^\eta$, so that after Poisson summation we only get a contribution from the frequency $(0, 0)$. The point count modulo $\Delta$ corresponding to $N_1(a; \Delta)$ is trivially equal to $|\Delta|$. Thus, applying the Siegel-Walfisz property (4.3) we get
\[ U_2^{(j)} \ll_C X N \log^{-C} X. \]
Again the technical issues of the contribution from $\max\{\Delta_2, (z_2 - z_1, \Delta)\} > Y$, removing $F_{M_1}$, and a bounding the part $\theta_1 = \theta_2 + O(\sqrt{\delta}) \pmod{\pi}$ are handled similarly as above.

7. The sieve argument for Theorems 1 and 2

In this section we give the proofs of Theorems 1 and 2 by applying Harman’s sieve method with the arithmetic information given by Propositions 5 and 12. Define
\[ S(A^{(j)}_d, Z) := \sum_{n \sim X/d, (n, P(Z)) = 1} a^{(j)}_{dn} \] and \[ S(B_d, Z) := \sum_{n \sim X/d, (n, P(Z)) = 1} b_{dn}, \]
and denote $A_1^{(j)} = A^{(j)}$ and $B_1 = B$. Then Theorems 1 and 2 follow from the following quantitative version. The lower bounds have not been optimized and could be improved by more Buchstab iterations, especially for $A^{(2)}$. The upper bounds in Theorems 1 and 2 follow from a standard sieve upper bound, using Proposition 4 or even just the trivial level of distribution $X^{1/2-\eta}$.

**Theorem 24.** For $X$ sufficiently large we have
\[ S(A^{(1)}, 2X^{1/2}) \geq 0.6 \cdot S(B, 2X^{1/2}) \] and
\[ S(A^{(2)}, 2X^{1/2}) \geq 0.1 \cdot S(B, 2X^{1/2}). \]
Let $\eta > 0$ be small and define the widths of the Type II ranges (Proposition 12)
\[ \gamma_1 := 1/3 - 1/4 - 2\eta = 1/12 - 2\eta \text{ and } \gamma_2 := 2/9 - 1/6 - 2\eta = 1/18 - 2\eta. \]
While our arithmetic information is not sufficient to give an asymptotic formula for primes, we can still give an asymptotic formula for certain sums of almost-primes with no prime factors below $X^{\gamma_2}$, as the following Proposition shows.

**Proposition 25.** Let $D_1 := X^{3/4-\eta}$ and $D_2 := X^{5/6-\eta}$ and let $U_j \leq D_j$. Let $W = X^{1/(\log \log x)^2}$. Then for any bounded coefficients $\alpha(m)$ supported on $(m, P(W)) = 1$ we have
\[ \sum_{m \sim U_j} \alpha(m) S(A^{(j)}_m, X^{\gamma_2}) = \sum_{m \sim U_j} \alpha(m) S(B_m, X^{\gamma_2}) + O_{\eta,C}(X/\log^C X). \]
Proof. Let \( C = \{c_n\} \in \{A^{(j)}, B\} \). Then by expanding using the Möbius function we get
\[
\sum_{m=U_j} d \, \alpha(m) S(C_m, X^{1/2}) = \sum_{m=U_j} \alpha(m) \sum_{n \sim X/m} \mu(d) \sum_{(n, P(X^{1/2})) = 1} c_{mn}
\]
\[
= \sum_{m=U_j} \alpha(m) \sum_{d \mid P(W, X^{1/2})} \mu(d) \sum_{n \sim X/dm} \sum_{(n, P(W)) = 1} c_{dnmn}.
\]

We split the sum in two parts, \( du_j \leq D_j \) and \( du_j > D_j \), and show that in each part we can get an asymptotic formula by Type I information and Type II information, respectively.

For the first part we write
\[
\alpha'(m') := \sum_{m' = dm \atop m' \sim U_j} \alpha(m) \mu(d) = \alpha'_1(m') + \alpha'_2(m'),
\]
where for some large constant \( C > 0 \)
\[
\alpha'_1(m') = \alpha'(m')1_{\tau(m') \leq \log^C X} \quad \text{and} \quad \alpha'_2(m') = \alpha'(m')1_{\tau(m') > \log^C X}.
\]
Since \( |\alpha'_1(m')| \leq \log^C X \), we get by Proposition [5]
\[
\sum_{m' \leq 2D_j} \alpha'_1(m') \sum_{n \sim X/m'} a_{m'n} = \sum_{m' \leq 2D_j} \alpha'_1(m') \sum_{n \sim X/m'} b_{m'n} + O_C(X/\log^C X).
\]
For \( \alpha'_2(m') \leq \tau(m') \) we get by Lemma [11] with \( k = 10 \) and by trivial bounds
\[
\left| \sum_{m' \leq 2D_j} \alpha'_2(m') \sum_{n \sim X/m'} c_{m'n} \right| \leq \sum_{\tau(n) \leq \log^C X} \sum_{\tau(n) \geq \log^C X} \tau(n)c_{en}.
\]
(7.1) \[
\leq \sum_{\tau(n) \geq \log^C X} \tau(n)\mathcal{O}(1) \sum_{n \sim X/ef} c_{efn} \ll X \sum_{\tau(n) \geq \log^C X} \tau(n)\mathcal{O}(1) \frac{\tau(n)\mathcal{O}(1)}{ef} \leq X \log^{-C+O(1)} X,
\]
where we have plugged in the factor \( \tau(n)\mathcal{O}(1) / \log^C X \geq 1 \) to get the penultimate step.

Similarly, in the part \( du_j > D_j \) we get an asymptotic formula by Proposition [12]. To see this, let us write \( d = p_1 \cdots p_k \) to get (noting that \( k \neq 0 \) since \( d > D_j/U_j > 1 \))
\[
\sum_{m=U_j \atop dU_j > D_j} \alpha(m) \mu(d) \sum_{n \sim X/dm} c_{dnmn} = \sum_{1 \leq k < \log X} (-1)^k \sum_{m \sim U_j} \alpha(m) \sum_{W \sim \log X / p_1 \cdots p_{k-1} U_j} \sum_{n \sim X/p_1 \cdots p_k m \atop (n, P(W)) = 1} c_{p_1 \cdots p_k mn}.
\]
Since \( p_1 \cdots p_k U_j > D_j, U_j \leq D_j \) and \( p_k < X^{1/2} \), by the greedy algorithm there is a unique \( 0 \leq \ell \leq k \) such that
\[
p_1 \cdots p_{\ell-1} U_j < D_j X^{1/2} \quad \text{and} \quad p_1 \cdots p_{\ell} U_j < D_j X^{1/2},
\]
Note that $[D_jX^{-\gamma_i}, D_j]$ is now exactly the admissible range for $M$ in Proposition \[\text{Proposition 12}\]. We obtain that the part $dU_j > D_j$ is partitioned into $\ll \log^2 X$ sums of the form

$$\sum_{m \sim U_j} \alpha(m) \sum_{W \leq p_1 < \cdots < p_k < X^\gamma} \sum_{n \sim X/p_1 \cdots p_k m \atop (n, P(W)) = 1} \sum_{p_1 \cdots p_k U_j > D_j} c_{p_1 \cdots p_k m m}.$$ 

The cross-conditions $p_{\ell+1} > p_\ell$ and $p_1 \cdots p_k U_j > D_j$ are easily removed by applying a finer-than-dyadic decomposition (Section \[3.1\] with $\delta = \log^{-C} X$) to the four variables $p_\ell, p_{\ell+1}, p_1 \cdots p_\ell$, and $p_{\ell+1} \cdots p_k$. Hence, writing

$$m' := mp_1 \cdots p_\ell \quad \text{and} \quad n' := np_{\ell+1} \cdots p_k,$$

we obtain for some coefficients $\alpha'(m')$ and $\beta'(n')$ supported on $(m'n', P(W)) = 1$ sums of the form

$$\sum_{m'n' \sim X} \alpha'(m') \beta'(n') c_{m'n'}.$$ 

Here $\alpha'(m')$ is supported on $m' \in [D_jX^{-\gamma_i}, 2D_j]$ with $(m', P(W)) = 1$. The cross-condition $m'n' \sim X$ can be removed by a finer-than-dyadic decomposition and the coefficients $\alpha'(m'), \beta'(n')$ can be reduced to bounded coefficients by a similar argument as in \[7.1\]. Hence, by Proposition \[12\] we get

$$\sum_{m'n' \sim X} \alpha'(m') \beta'(n') a_{m'n'}^{(j)} = \sum_{m'n' \sim X} \alpha'(m') \beta'(n') b_{m'n'} + O(C/X) + O(C/X \log C X),$$

provided that we show that $\beta'(|z|^2)$ satisfies the Siegel-Walfisz property \[1.1\]. To this end we write for any $m'$

$$\beta'(n') = \sum_{p \geq W} \beta'(pn) 1_{p+(n) \leq p} = \sum_{p \geq W} \beta'(pn) 1_{p+(n) < p} 1_{p|m'n'} + \sum_{p \geq W} \beta'(pn) 1_{p+(n) < p} 1_{p|m'n'}.$$ 

In the first term we see by the support of $c_{m'n'}$ that $p$ is restricted to $p \equiv 1 (4)$ by the sum-of-two-squares theorem. Hence, by dropping $p \nmid m'n'$ we get

$$\beta'(n') = \sum_{p \geq W \atop p \equiv 1 (4)} 1_{p+(n) < p} \beta'(pn) + O\left( \sum_{p \geq W} |\beta'(pn)| 1_{p|m'n'} \right).$$

The contribution from the second term is negligible by trivial bounds since we get a square factor $\geq W$. Similarly, we may reduce to the case where $\beta'(n')$ is supported on square-free integers. By the construction of $\beta'(n')$ we can write

$$n' = pn = pq_1 \cdots q_K, \quad \text{for} \quad W < q_1 < \cdots < q_K < p,$$

and remove any cross-conditions involving $p$ by a finer-than-dyadic decomposition (Section \[3.1\] with $\delta$) applied to $p$ and at most $O(1)$ variables. Hence, for some $\beta''(n')$ the coefficient $\beta'(n')$ is up to a negligible error term replaced by coefficients of the form

$$\beta''(n') \sum_{p \equiv 1 (4) \atop p|m'} \psi(p) \psi(p \equiv 1 (4)),$$

where $\psi$ is a suitable Fourier coefficient.
with $P \geq W/2$. The condition $p \nmid n$ may be dropped with a negligible error term as $p \gg W$. Let $\psi'$ be as in Section 3.1 with $\delta' = \delta(\log X)^{-C}$. Then applying Section 3.1 to the variable $n$ with $\psi'$ we get coefficients of the form

$$
\sum_{n' = pn} \beta''(n) \psi'_{N_0}(n) \sum_{p \equiv 1 (4)} \psi(p),
$$

and since $\delta'$ is small compared to $\delta$, this can be replaced with a negligible error term by

$$
\beta''(n') := \psi_{PN_0}(n') \sum_{n' = pn} \beta''(n) \psi'_{N_0}(n) \sum_{p \equiv 1 (4)} 1
$$

which is as in (4.2) and thus satisfies the Siegel-Walfisz property (4.1). □

The general idea of Harman’s sieve is to use Buchstab’s identity to decompose the sum $S(C, 2\sqrt{X})$ (in parallel for $C = A^{(j)}$ and $C = B$) into a sum of the form $\sum_{k} \epsilon_k S_k^{(j)}(C)$, where $\epsilon_k \in \{-1, 1\}$ and $S_k^{(j)}(C) \geq 0$ are sums over almost-primes. Since wish to obtain a lower bound, for $C = A^{(j)}$ we can insert the trivial estimate $S_k^{(j)}(A^{(j)}) \geq 0$ for any $k$ such that the sign $\epsilon_k = 1$ – we say that these sums are discarded. For the remaining $k$ we will obtain an asymptotic formula by using Propositions 12 and 25. That is, if $\mathcal{K}$ is the set of indices of the sums that are discarded, then we will show

$$
S(A^{(j)}, 2\sqrt{X}) = \sum_{k} \epsilon_k S_k^{(j)}(A^{(j)}) \geq \sum_{k \in \mathcal{K}} \epsilon_k S_k^{(j)}(A^{(j)})
$$

$$
= (1 + o(1)) \sum_{k \notin \mathcal{K}} \epsilon_k S_k^{(j)}(B) = (1 + o(1)) S(B, 2\sqrt{X}) - \sum_{k \in \mathcal{K}} S_k^{(j)}(B).
$$

We are successful if we can then show that $\sum_{k \in \mathcal{K}} S_k^{(j)}(B) < c_1 S(B, 2\sqrt{X})$ for $c_1 = 0.4$ and $c_2 = 0.9$.

To bound the error terms corresponding to $k \in \mathcal{K}$ we need a lemma which converts sums over almost primes over $b_n$ into integrals which can be bounded numerically. Let $\omega(u)$ denote the Buchstab function (see [7, Chapter 1] for the properties below, for instance), so that by the Prime number theorem for $Y^\epsilon < Z < Y$

$$
\sum_{Y < n \leq 2Y} 1_{(n, P(Z)) = 1} = (1 + o(1)) \omega \left( \frac{\log Y}{\log Z} \right) \frac{Y}{\log Z}.
$$

Similarly, using the Prime number theorem for primes $p \equiv 1 (4)$ we get

$$
\sum_{Y < n \leq 2Y} 1_{(n, P(Z)) = 1} \sum_{n = a^2 + b^2} 1 = (1 + o(1)) \omega \left( \frac{\log Y}{\log Z} \right) \frac{Y}{\log Z}.
$$

Note that for $1 < u \leq 2$ we have $\omega(u) = 1/u$. In the numerical computations we will use the following upper bound for the Buchstab function (see [10, Lemma 5], for instance)

$$
\omega(u) \leq \begin{cases} 
0, & u < 1 \\
1/u, & 1 \leq u < 2 \\
(1 + \log(u - 1))/u, & 2 \leq u < 3 \\
0.5644, & 3 \leq u < 4 \\
0.5617, & u \geq 4.
\end{cases}
$$
Proof. Recall that by the change of variables $u$

$$\sum_{(p_1,\ldots,p_k)\in U} S(B_{p_1,\ldots,p_k}) = S(B,2\sqrt{X})(1+o(1)) \int \omega(\beta) \frac{d\beta_1\cdots d\beta_k}{\beta_1\cdots \beta_{k-1}^2},$$

where the integral is over the range

$$\{\beta : (X^{\beta_1},\ldots,X^{\beta_k}) \in U\}$$

and $\omega(\beta) = \omega(\beta_1,\ldots,\beta_k) := \omega((1-\beta_1-\cdots-\beta_k)/\beta_k)$.

Lemma 26. Let $U \subset [X^{2\delta},X]^k$. Then

$$\sum_{(p_1,\ldots,p_k)\in U} 1_{(q,P(p_k))=1} b_{q,p_1\cdots p_k}$$

By (7.2) and by the Prime number theorem for primes $p \equiv 1 \pmod{4}$, the left-hand side in the lemma is

$$= (1+o(1)) \frac{X}{2} \sum_{(n_1,\ldots,n_k)\in U} \frac{1}{n_1\cdots n_k (\log n_1)\cdots (\log n_{k-1})\log^2 n_k} \omega \left( \frac{\log X/(n_1\cdots n_k)}{\log n_k} \right)$$

$$= (1+o(1)) \frac{X}{2} \int_U \omega \left( \frac{\log X/(u_1\cdots u_k)}{\log u_k} \right) \frac{du_1\cdots du_k}{u_1\cdots u_k (\log u_1)\cdots (\log u_{k-1})\log^2 u_k}$$

$$= (1+o(1)) \frac{X}{2\log X} \int \omega(\beta) \frac{d\beta_1\cdots d\beta_k}{\beta_1\cdots \beta_{k-1}^2} = (1+o(1)) S(B,2\sqrt{X}) \int \omega(\beta) \frac{d\beta_1\cdots d\beta_k}{\beta_1\cdots \beta_{k-1}^2}$$

by the change of variables $u_j = X^{\beta_j}$.

7.1. Buchstab iterations. We are now ready to give the proof of Theorem 24. We will apply similar Buchstab decompositions in both cases $j \in \{1,2\}$. Let $C = (c_n) \in \{A^{(j)},B\}$. We get by applying Buchstab’s identity twice

$$S(C,2X^{1/2}) = \sum_{X^{\gamma_j} \leq p_1 < 2X^{1/2}} S(C_{p_1},X^{\gamma_j}) + \sum_{X^{\gamma_j} \leq p_2 < p_1 < 2X^{1/2}} S(C_{p_1,p_2},p_2)$$

$$= S_1(C) - S_2(C) + S_3(C).$$

By Proposition 25 we get for $j,k \in \{1,2\}$

$$S_k(A^{(j)}) = S_k(B) + O_C(X/\log^C X).$$
For the third sum we first separate the part where $p_1 p_2^2 > 4X$. In this part $S(C_{p_1 p_2}, p_2)$ counts numbers $n$ of size $\leq 2X/p_1 p_2 < p_2/2$ with $(n, P(p_2)) = 1$, so that we must have $n = 1$ and $p_1 p_2 \sim X$. Hence, we get

$$\sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop p_1 p_2^2 > 4X} S(B_{p_1 p_2}, p_2) \ll \sum_{p_1, p_2 \ll X^{1/2}} b_{p_1 p_2} \ll X/\log^2 X.$$  

and for $C = A^{(j)}$ we use positivity and the trivial bound

$$\sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop p_1 p_2^2 > 4X} S(A_{p_1 p_2}^{(j)}, p_2) \geq 0.$$  

Let $\tilde{S}_3(C)$ denote the part of $S_3(A)$ with $p_1 p_2^2 \leq 4X$. We split it into two parts

$$\tilde{S}_3(C) = S_{31}(C) + S_{32}(C),$$

where

$$S_{31}(C) := \sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop p_1 p_2^2 \leq D_j} S(C_{p_1 p_2}, p_2) \text{ and } S_{32}(C) := \sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop D_j < p_1 p_2^2 \leq 4X} S(C_{p_1 p_2}, p_2).$$

Consider first $S_{32}(C)$. Let $D_j := X^{\alpha_j}$ and denote the Type II range (Proposition 12) by

$$I(j) := [X^{1-\alpha_j}, X^{1-\alpha_j+\gamma_j}] \cup [X^{\alpha_j-\gamma_j} \cup X^{\alpha_j}] \text{ and } J(j) := \log I(j) = [1-\alpha_j, 1-\alpha_j + \gamma_j] \cup [\alpha_j - \gamma_j, \alpha_j].$$

By Proposition 12 (after removal of cross-conditions) we get an asymptotic formula whenever a combination of the variables lies in $I(j)$, that is,

$$\sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop D_j < p_1 p_2^2 \leq 4X} S(A_{p_1 p_2}^{(j)}, p_2) = \sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop D_j < p_1 p_2^2 \leq 4X\atop p_1, p_2, \text{or } p_1 p_2 \in I(j)} S(B_{p_1 p_2}, p_2) + O(X/\log^C X).$$

For the remaining part we write by positivity $S_{32}(A^{(j)}) \geq 0$ and by Lemma 26

$$\sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop D_j < p_1 p_2^2 \leq 4X\atop p_1, p_2, \text{and } p_1 p_2 \notin I(j)} S(B_{p_1 p_2}, p_2) = (\Omega_2(j) + o(1))S(B, 2X^{1/2})$$

(7.3)

for

$$\Omega_2(j) := \int_{U_2(j)} \omega((1-\beta_2 - \beta_2)/\beta_2) \frac{d\beta_1 d\beta_2}{\beta_1 \beta_2^2}$$

with the two dimensional range

$$U_2(j) := \{ (\beta_1, \beta_2) : \gamma_j < \beta_2 < \beta_1 < 1/2, \alpha_j < \beta_1 + 2\beta_2 < 1, \beta_1, \beta_2, \beta_1 + \beta_2 \notin J(j) \}.$$  

In the part $S_{31}(C)$ we can apply Buchstab’s identity to get

$$S_{31}(C) = \sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop p_1 p_2^2 \leq D_j} S(C_{p_1 p_2}, X^\gamma) - \sum_{X^\gamma \leq p_2 < p_1 < 2X^{1/2}\atop p_1 p_2^2 \leq D_j} S(C_{p_1 p_2 p_3}, p_3)$$
If $p_1p_2p_3^2 > 4X$ then $S(C_{p_1p_2p_3}, p_3) = 0$ except if $p_1p_2p_3 \sim X$, which contradicts the fact that $p_1p_2p_3 < p_1p_2^2 \leq D_j$. Thus, in the second sum we always have $p_1p_2p_3^2 \leq 4X$. Applying Buchstab’s identity once more we get

\[ S_{31}(C) = S_{311}(C) - S_{312}(C) + S_{313}(C) \]

with

\[ S_{311}(C) := \sum_{X^{\gamma_j} < p_3 < p_1 < 2X^{1/2} \atop p_1p_2^2 \leq D_j} S(C_{p_1p_2}, X^{\gamma_j}), \]

\[ S_{312}(C) := \sum_{X^{\gamma_j} < p_3 < p_2 < p_1 < 2X^{1/2} \atop p_1p_2^2 \leq D_j \atop p_1p_2p_3^2 \leq 4X} S(C_{p_1p_2p_3}, X^{\gamma_j}), \quad \text{and} \]

\[ S_{313}(C) := \sum_{X^{\gamma_j} < p_4 < p_3 < p_2 < p_1 < 2X^{1/2} \atop p_1p_2^2 \leq D_j \atop p_1p_2p_3^2 \leq 4X} S(C_{p_1p_2p_3p_4}, p_4). \]

Since $p_1p_2p_3 \leq p_1p_2^2 \leq D_j$, we have asymptotic formulas for the first two sums, that is, for $j, k \in \{1, 2\}$

\[ S_{31k}(A^{(j)}) = S_{31k}(B) + O_C(X/\log^C X). \]

For the third sum we use a similar treatment as for $S_{32}(C)$ above. First, we separate the contribution from $p_1p_2p_3p_4^2 > 4X$, which implies $p_1p_2p_3p_4 \sim X$. By $p_1p_2p_3^2 \leq 4X$ and $p_4 < p_3$ we infer $p_4 = (1 + O(1))p_3$. Hence, we have

\[ S(B_{p_1p_2p_3p_4}, p_4) \ll \sum_{X^{\gamma_j} < p_4 < p_3 < p_2 < p_1 < 2X^{1/2} \atop p_1p_2^2 \leq D_j \atop p_1p_2p_3^2 \leq 4X} b_{p_1p_2p_3p_4} \ll X/\log^2 X, \]

so that we may restrict to $p_1p_2p_3p_4^2 \leq 4X$. Here we use Proposition [12] to give an asymptotic formula whenever a combination of the variables is in the Type II range $I(j)$ and we discard the rest. This gives us an error contribution (by Lemma [20])

\[ (7.4) \quad \sum_{X^{\gamma_j} < p_4 < p_3 < p_2 < p_1 < 2X^{1/2} \atop p_1p_2^2 \leq D_j \atop p_1p_2p_3^2 \leq 4X} S(B_{p_1p_2p_3p_4}, p_4) = (\Omega_4(j) + o(1)) S(B, 2X^{1/2}), \]

where

\[ \Omega_4(j) := \prod_{i \in I} \omega((1 - \beta_i - \beta_2 - \beta_3 - \beta_4)/\beta_4) d\beta_1d\beta_2d\beta_3d\beta_4/\beta_1\beta_2\beta_3\beta_4^2 \]

with the four dimensional range $U_4(j)$ defined by

\[ \gamma_j < \beta_4 < \beta_3 < \beta_2 < \beta_1 < 1/2, \quad \beta_1 + 2\beta_2 < \alpha_j, \quad \beta_1 + \beta_2 + 2\beta_3 < 1, \]

\[ \beta_1 + \beta_2 + \beta_3 + 2\beta_4 < 1, \quad \text{and} \quad \sum_{i \notin J(j)} \beta_i \neq J(j) \forall I \subseteq \{1, 2, 3, 4\}. \]
Combining the above asymptotic evaluations and the error terms \((7.3)\) and \((7.4)\), we have shown that for \(j \in \{1, 2\}\)
\[S(A(j), 2X^{1/2}) \geq (1 - \Omega_2(j) - \Omega_4(j))S(B, 2X^{1/2}).\]
To complete the proof of Theorem \([24]\) we compute the numerical upper bounds
\[\Omega_2(1) + \Omega_4(1) \leq 0.38 + 0.017 < 0.4\]
and
\[\Omega_2(2) + \Omega_4(2) \leq 0.38 + 0.49 < 0.9.\]
Python codes which compute rigorous upper bounds for these integrals can be found at the following links.
\[
\begin{align*}
\Omega_2(1) & \quad \text{http://codepad.org/lO5xHGuO} \\
\Omega_4(1) & \quad \text{http://codepad.org/Ygi8kNUF} \\
\Omega_2(2) & \quad \text{http://codepad.org/6hNTZtYt} \\
\Omega_4(2) & \quad \text{http://codepad.org/spA3ENVz}
\end{align*}
\]
This completes the proof of Theorem \([24]\) and as mentioned above, we then get Theorems \([1]\) and \([2]\).

8. The set-up for Theorem \([3]\)

8.1. The set-up. Let \(K/Q\) be a Galois extension of degree \(k\) and fix a basis \(\omega_1, \ldots, \omega_k\) for the ring of integers \(O_K\). Define the form
\[N(b_1, \ldots, b_k) := N_{K/Q}(b_1\omega_1 + \cdots + b_k\omega_k)\]
and the incomplete form
\[N(b_1, \ldots, b_k-1) := N(b_1, \ldots, b_k, 0)\]
Let \(\psi\) be as in Section \([3.1]\) with \(\delta = \log^{-C} X\) for some large \(C > 0\). We define
\[\Omega(b_1, \ldots, b_{k-1}) := \int_{0 \leq b_i \leq X^{1/2k}} \frac{1}{\delta} \int_{1/2 \leq B \leq 2X\eta_k} \psi_B(N(b_1, \ldots, b_{k-1})) \frac{\int_{u \in [0, X^{1/2k}k^{-1}]} \psi_B(N(u_1, \ldots, u_{k-1})) du}{dB},\]
where \(C_k > 0\) is a large enough constant so that
\[\int_{u \in [0, C_k^{1/k}k^{-1}]} \psi_2(N(u_1, \ldots, u_{k-1})) du \neq 0,\]
so that the denominator in the definition of \(\Omega(b_1, \ldots, b_{k-1})\) is always non-zero. We also define the arithmetic factor
\[\kappa_k := \sum_{c \geq 1} \frac{\mu(c)}{c^2} \left( \sum_{c \geq 1} \frac{\mu(c) \rho_k(c)}{c^k} \right)^{-1},\]
where
\[\rho_k(c) = |\{b_1, \ldots, b_{k-1} : N(b_1, \ldots, b_{k-1}) \equiv 0 (c)\}|\]
We set
\[a_n^{(k)} := \kappa_k \sum_{n = a^2 + N(b_1, \ldots, b_{k-1})^2 \atop (a, N(b_1, \ldots, b_{k-1})) = 1} \Omega(b_1, \ldots, b_{k-1}),\]
For typical \(b_1, \ldots, b_{k-1}\) we have
\[
\Omega(b_1, \ldots, b_{k-1}) \asymp X^{1/(2k)},
\]
where the upper bound holds for all \(b_1, \ldots, b_{k-1}\), so that
\[
\sum_{n \sim X} d_n^{(k)} \asymp X.
\]
For the comparison sequence we set
\[
b_n := \sum_{n=a^2+b^2} \frac{1}{\phi(n)} \int_{1/2 \leq B \leq \sqrt{X} \over \sqrt{k}} \psi_B(b) \frac{dB}{B},
\]
which is the same as before except that we essentially restrict to \(b \leq 2C_k^{-1}\sqrt{X}\). When we write \(a_n^{(k)}\) we have suppressed the fact that the sequence depends also on \(K\) and the choice of basis \(\omega_j\). We will use similar convention below to other quantities, and with the exception of Section 11 all implied constants are allowed to depend on \(K\) and the \(\omega_j\).

### 8.2. Lemmas

We need the following lemma for handling the diagonal terms after Cauchy-Schwarz with the Type I and Type II sums. Basically, this ensures that the density of the numbers \(a^2 + N(b_1, \ldots, b_{k-1})^2 \sim X\) is \(\asymp X^{-1/(2k)} + o(1)\), as expected, so that we get an optimal control for the diagonal terms. Here we do not need to restrict to \(b_k = 0\), and the lemma is essentially an analogue of the divisor bound \(\tau(n) \ll n^{\varepsilon}\) for the the number field \(K\).

**Lemma 27.** The number of representations
\[
N(b_1, \ldots, b_k) = b
\]
with \(|b_j| \leq B\) is \(\ll B^\varepsilon\), where the implied constant may depend on \(K\) and the choice of \(\omega_j\).

**Proof.** Let \(b_1, \ldots, b_k\) be such that \(N(b_1, \ldots, b_k) = b\). The principal ideal factorizes uniquely into prime ideals
\[
(b_1 \omega_1 + \cdots + b_k \omega_k)O_K = p_1 \cdots p_m.
\]
We have
\[
N_{K/Q}(b'_1 \omega_1 + \cdots + b'_k \omega_k) = b
\]
only if
\[
(b'_1 \omega_1 + \cdots + b'_k \omega_k)O_K = \sigma_1(p_1) \cdots \sigma_m(p_m)
\]
for some \(\sigma_j \in \text{Gal}(K/Q)\). The number of choices for \(\sigma_j\) which give a different ideal is at most
\[
\leq \prod_{p^l \| b} (\ell + k)^k \ll \tau(b)^{O_k(1)} \ll B^\varepsilon
\]
by a divisor bound. To see this, by multiplicativity of \(N_{K/Q}\) it suffices to consider the case that
\[
b = p^\ell
\]
so that \(m \leq \ell\) and for some \(efg = k\) we have \(pO_K = P_1^{e_1} \cdots P_g^{e_g}\) for some prime ideals \(P_i \subseteq O_K\). Then the number of choices for \(\sigma_j\) which give a distinct ideal is bounded by
the number of ways choosing \( \ell \) (possibly repeating) elements from a set of size \( k \), which is
\[
\leq \binom{\ell + k - 1}{\ell} \leq (\ell + k)^k.
\]
Thus, it suffices to show that the number of units \( \varepsilon_0 \) of the ring of integers \( O_K \) with
\[
\varepsilon_0(b_1 \omega_1 + \cdots + b_k \omega_k) = b'_1 \omega_1 + \cdots + b'_k \omega_k, \quad |b'_j| \leq B
\]
is \( \ll \varepsilon B^\varepsilon \). Consider the Minkowski embedding
\[
M : \mathbb{R} \ni (\log |\sigma(\beta)|)_{\sigma} \rightarrow \mathbb{R}^r,
\]
indexed by the \( r = r_1 + r_2 \) complex embeddings \( \sigma : K \rightarrow \mathbb{C} \). For every such embedding we have
\[
|\sigma(b_1 \omega_1 + \cdots + b_k \omega_k)| = |b_1 \sigma(\omega_1) + \cdots + b_k \sigma(\omega_k)| \ll B,
\]
and
\[
|\sigma(b_1 \omega_1 + \cdots + b_k \omega_k)| = b \prod_{\sigma \neq \sigma'} |\sigma'(b_1 \omega_1 + \cdots + b_k \omega_k)| \gg B^{1-k}.
\]
Thus, our task is reduced to enumerating units \( \varepsilon_0 \) such that
\[
\|M(\varepsilon_0)\| \ll \log B.
\]
The set of units of \( O_K \) is mapped by \( M \) to a rank \( r - 1 \) lattice in \( \mathbb{R}^r \) of determinant \( \ll 1 \), so that by considering a Minkowski-reduced basis (see [16, Lemma 4.1], for instance) we see that the number of such units is
\[
\ll \log^{r-1} B \ll \log^k B \ll \varepsilon B^\varepsilon.
\]
\[\square\]

For a matrix \( A \) denote by \( A^{ij} \) the matrix which is obtained by deleting \( i \)th row and \( j \)th column. For the exponential sums in the next section we need the following two lemmas.

**Lemma 28.** Let \( \text{Gal}(K/\mathbb{Q}) = \{\sigma_1, \ldots, \sigma_k\} \) and define
\[
A := (a_{ij}), \quad a_{ij} = \sigma_i(\omega_j).
\]
Then \( \det A \neq 0 \), that is, \( A : K^k \rightarrow K^k \) is invertible. Furthermore, for any fixed \( i \) the numbers \( \det A^{ij} \in K \), \( j \leq k \) are linearly independent over \( \mathbb{Q} \), so that in particular \( \det A^{ij} \neq 0 \) for all \( i, j \leq k \).

**Proof.** We have \( \det(A) \neq 0 \) since \( \det(A)^2 \) is the discriminant of the number field \( K/\mathbb{Q} \), which is always a non-zero integer (cf. [14, Chapter III.3], for instance).

For the second claim suppose that for some \( i \) there are \( q_1, \ldots, q_k \in \mathbb{Q} \) not all zero such that
\[
\sum_{j \leq k} q_j \det A^{ij} = 0.
\]
For all \( \ell \leq k \) there is some \( \sigma^{(i,\ell)} \in \text{Gal}(K/\mathbb{Q}) \) and \( \varepsilon_{i\ell} \in \{\pm\} \) such that
\[
\det(\sigma^{(i,\ell)}(A^{ij})) = \varepsilon_{i\ell} \det A^{ij}, \quad j \leq k,
\]
since any \( \sigma \in \text{Gal}(K/Q) \) the permutes the row vectors of \( A \) and so we may choose \( \sigma^{(i,\ell)} = \sigma \ell \sigma_i^{-1} \). Note that \( \varepsilon_{i\ell} \) is independent of \( j \) as it only depends on the sign of the permutation of rows defined by \( \sigma^{(i,\ell)} \). Thus, we get
\[
\sum_{j \leq k} q_j \varepsilon_{i\ell} \det A^{ij} = \sum_{j \leq k} q_j \det(\sigma^{(i,\ell)}(A^{ij})) = \sigma^{(i,\ell)} \left( \sum_{j \leq k} q_j \det A^{ij} \right) = 0.
\]

Then the column vectors of the matrix \( D = (d_{\ell j}) = (\varepsilon_{i\ell} \det A^{ij}) \) are linearly dependent, that is, for \( c_j = (\varepsilon_{i\ell} \det A^{ij}) \) we have
\[
\sum_{j \leq k} q_j c_j = 0.
\]

Hence \( \det D = 0 \), which implies that also for the cofactor matrix \( C = ((-1)^{\ell+j} \det A^{ij}) \) we have \( \det C = 0 \), since \( C \) can be obtained from \( D \) by the row/column operations of multiplying the rows by \( \varepsilon_{i\ell}(-1)^{\ell} \) and the columns by \((-1)^j \). Then \( \det A^{-1} = (\det A)^{-1} \det C = 0 \), which is a contradiction, since in the above we have shown that \( \det A \neq 0 \).

We include the following lemma for completeness, even though we will not need it.

**Lemma 29.** Let \( F \in K[X_1, \ldots, X_k] \) be the form defined by
\[
F(X_1, \ldots, X_k) := \prod_{i \leq k} \left( \sum_{j \leq k} (-1)^{i+j} X_j \det A^{ij} \right).
\]
Then \( F \in \mathbb{Z}[X_1, \ldots, X_k] \).

**Proof.** We have \( \text{Gal}(K/Q) \cong \text{Gal}(K(X_1, \ldots, X_k)/Q(X_1, \ldots, X_k)) \), where the isomorphism is given by letting \( \sigma \in \text{Gal}(K/Q) \) act on the rational functions \( P/Q \in K(X_1, \ldots, X_k) \) coefficient-wise. By the definition of \( A \) we have
\[
F(X_1, \ldots, X_k) = \prod_{\sigma \in \text{Gal}(K/Q)} \left( \sum_{j \leq k} (-1)^{i+j} X_j \sigma(\det A^{ij}) \right).
\]
Hence, for any \( \sigma \in \text{Gal}(K(X_1, \ldots, X_k)/Q(X_1, \ldots, X_k)) \)
\[
\sigma F = F,
\]
so that \( F \in \mathbb{Q}[X_1, \ldots, X_k] \). Clearly then \( F \in \mathbb{Z}[X_1, \ldots, X_k] \). \( \square \)

**9. A Deligne-type bound**

Throughout this section we denote
\[
h_{2k} = (h_1, \ldots, h_{2k}) \quad \text{and} \quad h = (h_1, \ldots, h_{k-1}, h_{k+1}, \ldots, h_{2k-1}).
\]

We shall first consider the exponential sum with the complete norm form with \( k \) variables
\[
\sum_{b_1, c_1, \ldots, b_k, c_k (p)} e_p(h_{2k} \cdot (b, c)) N(b_1, \ldots, b_k) \equiv a N(c_1, \ldots, c_k) (p)
\]
and in the proof of Lemma 33 we reduce the case of incomplete forms with \( k-1 \) variables to this. The square-root bound would be \( \ll p^{k-3/2} \), whereas we lose a factor of \( p^{3/2} \) and
get only $\ll p^k$ for the generic $h$. This does not matter in the application, where we take a large $k$ so that the relative loss is quite small.

As is often the case with such exponential sums, it turns out to be helpful to consider a more general sum over finite fields $F_{p^m}$, where the additive character $e_p(x)$ on $F_p$ is replaced by the additive character on $F_{p^m}$

$$e_p(\text{Tr}_{F_{p^m}/F_p}(x)).$$

We need the following lemma, which is equivalent to the rationality of the $L$-function associated to the exponential sum.

Lemma 30.

$$\sum_{b_1,c_1,\ldots,b_k,c_k \in F_{p^m} \atop N(b_1,\ldots,b_k)=aN(c_1,\ldots,c_k)} e_p(\text{Tr}_{F_{p^m}/F_p}(h_{2k} \cdot (b,c))) = \sum_{j \leq g} \epsilon_j \lambda_j^m,$$

where for some fixed $g$ depending only on $k,a,N,h_{2k}$ and $\lambda_j \in \mathbb{C}$, $\epsilon_j \in \{\pm 1\}$ depending only on $k,p,a,N,h_{2k}$.

Proof. We partition the sum into

$$\sum_{I,J \subseteq \{1,\ldots,k\}} S(I,J),$$

where

$$S(I,J) = \sum_{b_1,c_1,\ldots,b_k,c_k \in F_{p^m} \atop N(b_1,\ldots,b_k)=aN(c_1,\ldots,c_k) \atop b_i=c_j=0 \Leftrightarrow (i,j) \in I \times J} e_p(\text{Tr}_{F_{p^m}/F_p}(h_{2k} \cdot (b,c))).$$

Here $b_i = c_j = 0$ means that these variables are 0 as elements of the field $F_{p^m}$ so that for $(i,j) \notin I \times J$ we have $b_j, c_j \in F_{p^m}^{\times}$. We expand the condition $N(b_1,\ldots,b_k) = aN(c_1,\ldots,c_k)$ to get

$$\frac{1}{p^m} \sum_{w \in F_{p^m}^{\times}} \sum_{b_1,c_1,\ldots,b_k,c_k \in F_{p^m} \atop b_i=c_j=0 \Leftrightarrow (i,j) \in I \times J} e_p(\text{Tr}_{F_{p^m}/F_p}(w(N(b_1,\ldots,b_k) - aN(c_1,\ldots,c_k)) + h_{2k} \cdot (b,c))).$$

The contribution from $w = 0$ is

$$\frac{1}{p^m} \sum_{I,J \subseteq \{1,\ldots,k\}} \sum_{b_1,c_1,\ldots,b_k,c_k \in F_{p^m} \atop b_i=c_j=0 \Leftrightarrow (i,j) \in I \times J} e_p(\text{Tr}_{F_{p^m}/F_p}(h_{2k} \cdot (b,c))).$$

which is 0 or $p^{m(2k-1)}$ if $p|h_j$ for all $j \leq 2k$, so that it is of a suitable form. For the contribution from $w \neq 0$ we have by \[11\] Theorem 1\] (since the $b_i, c_j$ with $(i,j) \notin I \times J$ run over $F_{p^m}^{\times}$)

$$\frac{1}{p^m} \sum_{w \in F_{p^m}^{\times}} \sum_{b_1,c_1,\ldots,b_k,c_k \in F_{p^m} \atop b_i=c_j=0 \Leftrightarrow (i,j) \in I \times J} e_p(\text{Tr}_{F_{p^m}/F_p}(w(N(b_1,\ldots,b_k) - aN(c_1,\ldots,c_k)) + h_{2k} \cdot (b,c)))$$

$$= \sum_{i=1}^r \omega_i^m - \sum_{j=1}^s \eta_j^m.$$
for some fixed $r, s$ depending only on $k, a, N, h_{2k}, I, J$ and $\omega_i, \eta_j \in \mathbb{C}$ depending only on $k, p, a, N, h_{2k}, I, J$. \hfill \Box

We also need the following standard lemma (see [13] Lemma 4.15], for instance).

**Lemma 31.** Let $n, g \geq 1$ and let $\lambda_1, \ldots, \lambda_g \in \mathbb{C}$. Suppose that there are constants $A, B > 0$ such that for every $m \geq 1$ we have for

$$\left| \sum_{j \leq g} \lambda_j^m \right| \leq AB^m.$$ 

Then $|\lambda_j| \leq B$ for all $j \in \{1, \ldots, g\}$.

The purpose of the above two lemmas is that they allow us to reduce bounding an exponential sum over $\mathbb{F}_p$ to bounding the corresponding sum over $\mathbb{F}_{p^r}$ for some fixed $q = p^f$. The benefit of this is that after a suitable extension $\mathbb{F}_q/\mathbb{F}_p$, the norm form $N(b_1, \ldots, b_k)$ factors into a product of $k$ linear forms.

We need to set up some notations for the next lemma. Recall that for any prime $p$ there are integers $e, f, g$ with $efg = k$ such that

$$pO_K = (P_1 \ldots P_g)^e$$

for some prime ideals $P_j$ of $K$. The integer $f$ is called the inertia degree of $p$, and we have for all $j \leq g$

$$O_K/P_j \cong \mathbb{F}_{p^f}.$$ 

We will denote $q = p^f$ and choose some $P_j$ for each prime $p$ (the exact choice is not important, see Remark [7]). For each prime $p$ we denote the reduction map by

$$\pi_p : O_K \rightarrow \mathbb{F}_q.$$ 

We then identify $\mathbb{F}_p = \pi_p(\mathbb{Z})$ as a subfield of $\mathbb{F}_q$.

We have the following Deligne-type bound, which is an explicit version for a special case of the more general stratification result of Fouvry and Katz [41 Theorem 1.2].

**Lemma 32.** Let $p$ be sufficiently large in terms of $K$ and $\omega_j$. Let $a \in \mathbb{F}_p$, $t \in \mathbb{F}_p^*$, and $q = p^f$. Let $A$ be as in Lemma 28. Let $F$ be as in Lemma 29 and let $Y_1 \subseteq \mathbb{A}^{2k}_K$ be defined by

$$F(x_1, \ldots, x_k) = (-1)^{k-1} a F(x_{k+1}, \ldots, x_{2k}).$$

Let $h_{2k} = (h_{2k,1}, h_{2k,2}) \in \mathbb{F}_p^{2k}$.

There are algebraic sets $X_j$ with $\mathbb{A}^{2k}_K = X_0 \supseteq X_1 \supseteq \cdots \supseteq X_{2k} \supseteq X_{2k+1} = \emptyset$ such that the following hold.

1. For $j \geq 1$ and $h_{2k} \in X_j(\mathbb{F}_q) \setminus X_{j+1}(\mathbb{F}_q)$ we have

$$\sum_{b_1, c_1, \ldots, b_k, c_k \in \mathbb{F}_p \atop N(b_1, \ldots, b_k) = a N(c_1, \ldots, c_k)} e_p(t h_{2k} \cdot (b, c)) \ll (a, p)^{1/2} p^{k-1+j/2}.$$ 

For $h_{2k} \in X_0(\mathbb{F}_q) \setminus (Y_1(\mathbb{F}_q) \cup X_1(\mathbb{F}_q))$ we have

$$\sum_{b_1, c_1, \ldots, b_k, c_k \in \mathbb{F}_p \atop N(b_1, \ldots, b_k) = a N(c_1, \ldots, c_k)} e_p(t h_{2k} \cdot (b, c)) \ll (a, p)^{1/2} p^{k-1/2}.$$
For $h_{2k} \in Y_1(F_q) \setminus X_1(F_q)$ we have
\[
\sum_{b_1, c_1, \ldots, b_k, c_k \in \mathbb{F}_p}
\sum_{N(b_1, \ldots, b_k) = a N(c_1, \ldots, c_k)} e_p(t \cdot (h_{2k} \cdot (b, c))) \ll p^k.
\]

(2) For each $j \geq 1$ we have
\[
X_j = \bigcup_{j = r + s, r,s \leq k} Z_r \times Z_s,
\]
where $Z_r \subseteq \mathbb{A}_K^k$ is a union of the $\binom{k}{r}$ hyperplanes of dimension $k - r$ defined by the $r$ independent linear equations
\[
\sum_{j \leq k} (-1)^j \det(A^{i,j}) h_j = 0
\]
for some distinct $i_1, \ldots, i_r \in \{1, \ldots, k\}$ with $A^{i,j}$ as in Lemma 28.

Proof. Suppose first that $a \neq 0$. It suffices to consider
\[
\sum_{b_1, c_1, \ldots, b_k, c_k \in \mathbb{F}_p}
\sum_{N(b_1, \ldots, b_k) = a N(c_1, \ldots, c_k)} e_p(t \cdot (h \cdot (b, c)))
\]
for all $n \geq 1$ with $q = p^g$, since by Lemma 30 for some $g \geq 1$
\[
\sum_{b_1, c_1, \ldots, b_k, c_k \in \mathbb{F}_p}
\sum_{N(b_1, \ldots, b_k) = a N(c_1, \ldots, c_k)} e_p(t \cdot (h \cdot (b, c))) = \sum_{j \leq g} \epsilon_j \lambda_j^{f_n}
\]
for some $\lambda_j \in \mathbb{C}, \epsilon_j \in \{\pm 1\}$ depending only on $p, a, t h_{2k}$, and $N$, so that the bound for $F_p$ follows by Lemma 31 from the corresponding bound for all $n \geq 1$.

Since $N(b_1, \ldots, b_k)$ is a norm form, it splits into linear factors over $K$ and hence over $F_q$
\[
N(b_1, \ldots, b_k) = \prod_{j=1}^k L_j(b_1, \ldots, b_k),
\]
where $L_j(b_1, \ldots, b_k)$ is the $j$th coordinate of $A b$ (with $A$ as in Lemma 28). Note that the linear map
\[
b \mapsto A b
\]
is invertible for $p$ sufficiently large, since $\det A \neq 0$ has finitely many prime factors. Thus, after the linear change of variables $(b, c) \mapsto (A b, A c) =: (b', c')$, it suffices to consider
\[
\sum_{b_1', c_1', \ldots, b_k', c_k' \in \mathbb{F}_p}
\sum_{b_1' - b_k' = a c_1' \cdots c_k'} e_p(t \cdot (h' \cdot (b', c')))
\]
where $h' = (h_1', h_2')$ is obtained from $h_{2k} = (h_{2k,1}, h_{2k,2})$ by the linear map
\[
h' = (A^{-1})^T h_{2k,j},
\]

since
\[
h_1' \cdot b' = (h_1')^T b' = ((A^{-1})^T h_{2k,1})^T A b = h_{2k,1} \cdot b.
\]
We split into separate cases depending on whether \( b'_1 \cdots b'_k = 0 \) or \( \neq 0 \). In the latter case the contribution is
\[
\sum_{x \in \mathbb{F}_q^k} \sum_{b'_1, \ldots, b'_k \in \mathbb{F}_q^k} e_p(t \text{Tr}_{\mathbb{F}_q^n/\mathbb{F}_p}(h'_1 \cdot b')) \sum_{c'_1, \ldots, c'_k \in \mathbb{F}_q} e_p(t \text{Tr}_{\mathbb{F}_q^n/\mathbb{F}_p}(h'_2 \cdot c'))
\]
Suppose first that all of the coordinates of \( h'_1 \) and \( h'_2 \) are \( \neq 0 \). Denote
\[
a' = a \prod_{i \notin J_1} h'_{1,i} \prod_{i \in J_1} (h'_{2,i})^{-1}
\]
and \( \psi(y) = e_p(t \text{Tr}_{\mathbb{F}_q^n/\mathbb{F}_p}(y)) \). After a change of variables the above sum is
\[
\sum_{x \in \mathbb{F}_q^k} \sum_{b''_1, \ldots, b''_k \in \mathbb{F}_q} \psi(b''_1 + \cdots + b''_k) \sum_{c''_1, \ldots, c''_k \in \mathbb{F}_q} e_p(t \text{Tr}_{\mathbb{F}_q^n/\mathbb{F}_p}(h'_1 \cdot a'x))
\]
If \( a' \neq (-1)^{k-1} \), then this is \( \ll (q^n)^{k-1/2} \) by [17] (6.12)]. If \( a' = (-1)^{k-1} \), then we use a point-wise bound for the hyper-Kloosterman sums and bound the sum over \( x \) trivially, which gives \( \ll (q^n)^k \). The equation \( a' = (-1)^{k-1} \) holds precisely when [12] holds, so that this is covered by \( Y'_1 \).

Suppose then that exactly \( j_1 \) and \( j_2 \) of the coordinates in \( h'_1 \) and \( h'_2 \) are 0, and denote the sets of such indices \( J_1 \) and \( J_2 \). By symmetry we may suppose that \( j_1 \geq 1 \). Then after a change of variables and denoting
\[
a' = a \prod_{i \notin J_1} h'_{1,i} \prod_{i \in J_1} (h'_{2,i})^{-1},
\]
we get
\[
\sum_{x \in \mathbb{F}_q^k} \sum_{b''_1, \ldots, b''_k \in \mathbb{F}_q} \psi(b''_{j_1+1} + \cdots + b''_k) \sum_{c''_1, \ldots, c''_k \in \mathbb{F}_q} e_p(t \text{Tr}_{\mathbb{F}_q^n/\mathbb{F}_p}(h'_1 \cdot a'x))
\]
Making the change of variables
\[
b''_1 \mapsto \frac{a' x}{b''_1 b''_2 \cdots b''_{j_1}}, \quad x \mapsto xc''_1 \cdots c''_{j_2}
\]
the sum becomes
\[
(q^n)^{j_1+j_2-1} \left( \sum_{b''_{j_1} \in \mathbb{F}_q} \text{Kl}_{k-j_1}(b''_{j_1} \cdot q^n) \right) \left( \sum_{x \in \mathbb{F}_q^k} \text{Kl}_{k-j_2}(x \cdot q^n) \right),
\]
which by [17] (6.11)] is
\[
\ll (q^n)^{j_1+j_2-1}(q^n)^{(k-j_1)/2}(q^n)^{(k-j_2)/2} \ll (q^n)^{-k+2(j_1+j_2)/2}.
\]
We now note that (with \( A^{ij} \) as in Lemma 28)
\[
(A^{-1})^T = \frac{1}{\det A}((-1)^{i+j} \det A^{ij})_{i,j \leq k},
\]
so that the equations $h'_1 = 0$ are precisely of the form (9.3). By Lemma 28 the equations are independent (for $p$ sufficiently large), so that $Z_r$ has dimension $k - r$ for each $r$.

Consider then the part where $b'_1 \cdots b'_k = 0$. If $a \neq 0$, we also have $c'_1 \cdots c'_k = 0$. Thus, this part of the sum is

$$
\sum_{I,J \subseteq \{1,\ldots,k\}} \sum_{I,J \neq \emptyset} \sum_{b'_1, \ldots, b'_k, c'_1, \ldots, c'_k \in \mathbb{F}_q} e_p(t \Tr_{\mathbb{F}_q^n/\mathbb{F}_p}(h' \cdot (b', c'))) \ll \min\{(p^{mn})^{2k-2}, \prod_{i \leq 2k} (h'_i, p)^{mn}\},
$$

which is bounded by $\max\{p^{k-1/2}, p^{k-1+j/2}\}$. Finally, if $a = 0$, then $b'_1 \cdots b'_k = 0$ and $c'_1, \ldots, c'_k$ are free and we have

$$
\sum_{I \subseteq \{1,\ldots,k\}} \sum_{I \neq \emptyset} \sum_{b'_1, c'_1, \ldots, c'_k \in \mathbb{F}_q} e_p(t \Tr_{\mathbb{F}_q^n/\mathbb{F}_p}(h' \cdot (b', c'))) \ll \min\{(p^{mn})^{2k-1}, \prod_{i \leq 2k} (h'_i, p)^{mn}\},
$$

which is

$$
\ll (a, p)^{mn/2} \max\{(p^{mn})^{k-1/2}, (p^{mn})^{k-1+j/2}\}
$$

with equality at $j = 2k - 1, 2k$. □

**Remark 7.** The coefficients $A^I$ of the equation (9.3) are in $\mathbb{F}_q$ by the reduction to $O_K/P_j$, where $pO_K = P_1 \cdots P_g$ and $q = p^{l_j}$. If $f = 1$, then they are in $\mathbb{F}_p$ and we get a system of $r$ linear equations for $h \in \mathbb{F}_q$. Since any $\sigma \in \Gal(K/\mathbb{Q})$ permutes the primes $P_1, \ldots, P_g$, we see that changing $P_j$ in $\mathbb{F}_q \cong O_K/P_j$ simply changes the set of $r$ indices $j \leq k$ such that $h'_j = 0$, so that the choice of the reduction modulus $P_j$ is inconsequential for our application, since we get the same bound for any set of $r$ indices. We can eliminate the use of the (countable) axiom of choice simply by averaging over all choices $P_j \in \{P_1, \ldots, P_g\}$. If $f = k$, then by Lemma 28 if one of the coordinates of $h'_j$ is 0, then all are 0 (for $p \gg 1$), so that the only solution $h \in \mathbb{F}_q^k$ is $h = 0$.

**Remark 8.** Using the Hasse-Davenport relation one can show that the function

$$
K(a; q) := \sum_{b_1, \ldots, b_h \in \mathbb{F}_q \colon N_{\mathbb{F}_q/\mathbb{F}_p}(b_1, \ldots, b_h) \equiv a (p)} \psi(b_1 + \cdots + b_h)
$$

is a constant multiple of the hyper-Kloosterman sum $\text{Kl}_k(a; q)$ (see [2] Applications de la formule des traces aux sommes trigonométriques (7.2.5)), thanks to Emmanuel Kowalski for pointing this out).

We need to set up some more notations for the following lemma, which uses the Chinese remainder theorem to combine Lemma 2 for different primes $p$. Then it may happen that $h$ lie in different $X_j$ for different primes $p$, where $X_j$ are as in Lemma 2. For any prime $p$ we extend the reduction map $\pi_p : O_K \to \mathbb{F}_q$ defined in (9.1) to a map

$$
\pi_p : O_K^{2k} \to \mathbb{F}_q^{2k} : (z_1, \ldots, z_{2k}) \mapsto (\pi_p(z_1), \ldots, \pi_p(z_{2k})).
$$

Let $D_1 = p_1 \cdots p_\ell$ be a square-free integer and denote $q_i = p_i^{\ell_i}$. For $(j_1, \ldots, j_\ell) \in \mathbb{F}_q$, set $j_i \leq k$,

$$
X(j_1, \ldots, j_\ell)(D_1) := \{z \in O_K^{2k} : \pi_{p_i}(z) \in X_{j_i}(\mathbb{F}_{q_i}) \forall i \leq \ell\}
$$
That is, we have
\[ X_{(j_1, \ldots, j_\ell)}(D_1) = \bigcap_{i < \ell} \pi_{p_i}^{-1} X_{j_i}(\mathbb{F}_q). \]

Then we have the following lemma.

**Lemma 33.** Let \( h_1, \ldots, h_{k-1}, h_{k+1}, \ldots, h_{2k-1} \in \mathbb{Z}. \) Let \( D_1 = p_1 \cdots p_{\ell} \) denote the largest square-free divisor such that \( (D/D_1, D_1) = 1 \) and let \( X_j \) be as in Lemma 32. For \( j \in \mathbb{Z} \)

denote
\[ j^+ := \max\{2, j\}. \]

Then
\[
\sum_{\substack{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv 0 \pmod{D_1} \in D_1 \n j_1, \ldots, j_{k-1} \equiv 0 \pmod{D_1} \in D_1}} e_D(h \cdot (b, c)) \ll (a, D_1)^{1/2}\tau(D)^{O(1)} \left( \frac{D}{D_1} \right)^{2k-2} \frac{1}{D_1^2} \sum_{h_k, h_{2k} \equiv 0 \pmod{D_1}} \sum_{h_{2k} \in X_{(j_1, \ldots, j_{k-1})}(D_1)} 1.
\]

**Proof.** By the Chinese remainder theorem we get
\[
\prod_{p^m \mid D} \sum_{\substack{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv 0 \pmod{D_1} \in D_1 \n j_1, \ldots, j_{k-1} \equiv 0 \pmod{D_1} \in D_1}} e_{p^m} \left( \left( \frac{D}{p^m} \right)^{-1} h \cdot (b, c) \right),
\]
where the inverse \( (\frac{D}{p^m})^{-1} \) is computed modulo \( p^m \). By expanding the conditions \( b_k \equiv c_k \equiv 0 \pmod{D_1} \), this is bounded by
\[
\frac{1}{D_1^2} \sum_{h_k, h_{2k} \equiv 0 \pmod{D_1}} \left( \prod_{p^m \mid D/D_1} \sum_{\substack{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv 0 \pmod{D_1} \in D_1 \n j_1, \ldots, j_{k-1} \equiv 0 \pmod{D_1} \in D_1}} 1 \right)
\times \left( \prod_{p \mid D_1} \sum_{\substack{b_1, c_1, \ldots, b_k, c_k \equiv 0 \pmod{D_1} \in D_1 \n j_1, \ldots, j_{k-1} \equiv 0 \pmod{D_1} \in D_1}} e_p \left( \left( \frac{D}{p} \right)^{-1} h_{2k} \cdot (b, c) \right) \right).
\]

For \( p^m \mid D/D_1 \) we use the trivial bound
\[
\sum_{\substack{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv 0 \pmod{D_1} \in D_1 \n j_1, \ldots, j_{k-1} \equiv 0 \pmod{D_1} \in D_1}} 1 \ll (p^m)^{2k-2}
\]
for \( p \mid D_1 \) we apply Lemma 32 (with \( t = (D/p)^{-1} \)) to get for some \( j = j(p) \)
with \( \pi_p(h_{2k}) \in X_j(\mathbb{F}_q) \) that
\[
\sum_{\substack{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv 0 \pmod{D_1} \in D_1 \n j_1, \ldots, j_{k-1} \equiv 0 \pmod{D_1} \in D_1}} e_p \left( \left( \frac{D}{p} \right)^{-1} h_{2k} \cdot (b, c) \right) \ll (a, p)^{1/2} p^{k-1+j^+/2}.
\]
Thus, for some \((j_1, \ldots, j_\ell)\) with \(h_{2k} \in X_{(j_1, \ldots, j_\ell)}(D_1)\) we have

\[
\left( \prod_{p|D_1} \sum_{b_1, c_1, \ldots, b_k, c_k \in (p)} \varepsilon_p \left( \frac{D}{p} \right)^{-1} h_{2k} \cdot (b, c) \right) \ll \tau(D)^{(1)} \prod_{p_i|D_1} p_i^{k-1+j_i^*/2} \ll \tau(D)^{(1)} D_1^{k-1} \prod_{p_i|D_1} p_i^{j_i^*/2}.
\]

\[\Box\]

Remark 9. Note that here we have not used the fact that in \(X_0 \setminus Y_1\) we get the superior bound \(D_1^{k-1/2}\). This is not crucial for our application and using this weaker bound simplifies greatly the upcoming computations. We have included the characterization of \(Y_1\) in Lemma 32 just for the sake of completeness. Making use of this would only marginally widen our Type II range (Proposition 11).

Let \(H_r\) denote any of the hyper-planes defining \(Z_r\) in Lemma 32. We need the following bound for the number of \((h_1, \ldots, h_k) \in \mathbb{Z}_k \subseteq O_K^k\) which are in \(H_r(\mathbb{F}_q)\) after reduction modulo some \(P_i\), where as before \(pO_K = P_1 \cdots P_q\) and \(q = p^j\). For a square-free integer \(D_1 = p_1 \cdots p_\ell\) and \(r = (r_1, \ldots, r_\ell) \in \mathbb{Z}_{\geq 0}^\ell\), denote

\[H_{(r_1, \ldots, r_\ell)}(D_1) := \{ z \in O_K^k : \pi_{p_i}(z) \in H_{r_i}(\mathbb{F}_q) \forall i \leq \ell \} \]

Then the following lemma is simply a generalization of the elementary bound

\[\sum_{h \leq H} 1_{h \equiv a(q)} \ll 1 + \frac{H}{q},\]

Lemma 34. Let \(H \geq 1\) and let \(D_1 = p_1 \cdots p_\ell\) be square-free. Let \(0 \leq r_1, \ldots, r_\ell \leq k\). Let

\[D_1(r) := \prod_{i \leq \ell p_i} \]

and define

\[r' = (r'_1, \ldots, r'_\ell), \quad r'_i = \max\{r_i - 1, 0\}.\]

Then for all \(i \leq \ell\) there are \(r'_{ij} \in \{0, 1\}, j \leq k-1\) such that

\[(9.5) \quad \sum_{j \leq k-1} r'_{ij} = r'_i\]

and

\[\sum_{h_{1, \ldots, h_k} \mid h_{1, \ldots, h_k} \leq H, i \leq k-1, h_k \leq D_1} 1_{h \in H_{(r_1, \ldots, r_\ell)}(D_1)} \ll D_1(r) \prod_{j \leq k-1} \left( 1 + \frac{H}{p_1^{r'_{1j}} \cdots p_{\ell}^{r'_{\ell j}}}. \right)\]

Furthermore, if \(I\) is the set of indices such that for \(i \in I\) we have \(r_i = k\), then

\[\sum_{h_{1, \ldots, h_k} \mid h_{1, \ldots, h_k} \leq H, i \leq k-1, h_k \leq D_1, (h_1, \ldots, h_{k-1}) \neq 0} 1_{h \in H_{(r_1, \ldots, r_\ell)}(D_1)} \ll 1_{H \geq \prod_{i \in I} p_i} D_1(r) \prod_{j \leq k-1} \left( 1 + \frac{H}{p_1^{r'_{1j}} \cdots p_{\ell}^{r'_{\ell j}}} \right)\].
where we have $r'_{ij} = 1$ for $i \in I$.

Proof. If we fix $h_1, \ldots, h_{k-1}$, then for every $i$ such that $r_i \neq 0$ and $p_i \gg 1$ there is at most one $h_k \in F_q$ with $(h_1, \ldots, h_k) \in H_{r_i}(F_q)$, since by Lemma 22 the coefficients det $A^{ik} \neq 0$ for $p_i \gg 1$. For $p_i \ll 1$ the number of $h_k \in F_q$ is trivially $\ll 1$. Thus, the number of $h_k \leq D_1$ is $\ll D_1(r)$, and the numbers $h_1, \ldots, h_{k-1}$ are restricted to a set of type

$$H'_{r_i}(D_1) \subseteq O^{k-1}_K,$$

which is a set such that $z \in H'_{r_i}(D_1)$ if for every $i \leq \ell$

$$\pi_{r_i}(z) \in H'_{r_i}(F_q),$$

for some $H'_{r_i} \subseteq H^{k-1}$ which is a hyperplane of dimension $k - r'_i$. That is, in this first step we have simply solved for $h_k$ in terms of $h_1, \ldots, h_{k-1}$ in one of the equations and substituted this to the remaining equations to get $r'_i - 1$ independent equations in $k - 1$ variables.

It then suffices to show that

$$\sum_{h_1, \ldots, h_{k-1}} 1_{h \in H'_{r_i}(D_1)} \ll \prod_{j \leq k} \left(1 + \frac{H}{p_1^{r_{ij}} \cdots p_{\ell}^{r_{ij}}}ight),$$

which we will do using induction on the number of variables $\ell' := k - 1$. The case of $\ell' = 1$ is trivially true, so consider the case of general $\ell'$. Fixing $h_1, \ldots, h_{\ell'-1}$, for each $i \leq \ell$ there are two possibilities, the set of $h_{\ell'} \in F_q$ such that

$$(h_1, \ldots, h_{\ell'}) \in H'_{r_i}(F_q),$$

(that is, the fibre) is either zero dimensional (a point) or one dimensional (a line) as a subvariety of $A^n_{F_q}$. Denoting this dimension by $d_{ik'}$, we let $r'_{ik'} = 1 - d_{ik'}$, and then the number of $|h_{\ell'}| \leq H$ with $(h_1, \ldots, h_{\ell'}) \in H'_{r_i}(F_q)$ is

$$\ll 1 + \frac{H}{p_1^{r'_{ik'}} \cdots p_{\ell}^{r'_{ik'}}},$$

and then $(h_1, \ldots, h_{\ell'-1})$ live in a hyperplane

$$H''_{r_i}(D_1) \subseteq O^{\ell'-1}_K,$$

where $r'' = (r'_1 - r'_{ik'}, \ldots, r'_\ell - r'_{ik'})$, and we can apply the induction hypothesis. Clearly we get (4.3), since if $\ell' \leq r'_i$ we must have $d_{ik'} = 0$ as we have $r'_i$ independent equations.

To get the second bound we note that $r_i = k$ means that $h'_j \equiv 0 (p_i)$ for all $j \leq k$, which implies that $h_j \equiv 0 (p_i)$ for all $j \leq k$. We sum over $(h_1, \ldots, h_{k-1}) \neq 0$, so suppose by symmetry that $h_{k-1} \neq 0$. Then $h_{k-1} \equiv 0 (\prod_{i \in I} p_i)$ is non-trivial and there are no solutions if $H < \prod_{i \in I} p_i$. Hence, the number of $h_{k-1}$ for fixed $h_1, \ldots, h_{k-2}$ is

$$\ll 1_{H \geq \prod_{i \in I} p_i} \left(1 + \frac{H}{p_1^{r_{ij}} \cdots p_{\ell}^{r_{ij}}}ight).$$

\qed
Remark 10. It is important for our applications that we get full savings in the longest sum \(h_k \leq D_1\), so that for large \(r_i\) we can cancel the losses from expanding the condition \(b_k = 0\) in Lemma 33.

**Lemma 35.** Let \(D_1 = p_1 \cdots p_\ell\) and \(\beta := \log H/\log D_1\). With the same notations as above, we have
\[
\prod_{j \leq k-1} \left(1 + \frac{H}{p_1 r_{ij} \cdots p_\ell r_{ij}}\right) \ll \tau(D_1)^{O(1)} H^{k-1} \prod_{i \leq \ell} p_i^{-\min\{1, \beta\} r'_i}.
\]

**Proof.** Let \(\alpha_i = \log p_i/\log D_1\), so that
\[
\alpha_1 + \cdots + \alpha_\ell = 1.
\]
We have (by \(\sum_{j \leq k-1} r'_{ij} = r'_i\))
\[
\prod_{j \leq k-1} \left(1 + \frac{H}{p_1 r_{ij} \cdots p_\ell r_{ij}}\right) = \prod_{j \leq k-1} \left(1 + \frac{H^{\alpha_1 + \cdots + \alpha_\ell}}{p_1 r_{ij} \cdots p_\ell r_{ij}}\right) \leq \prod_{j \leq k-1, i \leq \ell} \left(1 + \frac{H^{\alpha_i}}{p_i r_{ij}}\right)^{k-1 - r'_i}
\]
\[
= \prod_{i \leq \ell} \prod_{j \leq k-1} \left(1 + \frac{H^{\alpha_i}}{p_i r_{ij}}\right) = \prod_{i \leq \ell} \left(1 + H^{\alpha_i} r'_i\right) \left(1 + H^{\alpha_i}\right)^{k-1 - r'_i}
\]
\[
\ll \tau(D_1)^{O(1)} \prod_{i \leq \ell} \left(1 + H^{\alpha_i} r'_i\right) H^{\alpha_i (k-1) - r'_i} = \tau(D_1)^{O(1)} \prod_{i \leq \ell} \left(H^{\alpha_i (k-1) - r'_i} + \frac{H^{\alpha_i (k-1)}}{p_i r'_i}\right)
\]
\[
= \tau(D_1)^{O(1)} \prod_{i \leq \ell} \left(p_i^{\beta (k-1) - r'_i} + p_i^{\beta (k-1) - r'_i}\right) \ll \tau(D_1)^{O(1)} H^{k-1} \prod_{i \leq \ell} p_i^{-\min\{1, \beta\} r'_i}.
\]

\[\square\]

Combining Lemmas 33, 34, and 35 we finally get our key bound, using the fact that if \(h \neq 0\) then \((h_1, \ldots, h_{k-1}) \neq 0\) or \((h_{k+1}, \ldots, h_{2k-1}) \neq 0\).

**Lemma 36.** Let \(\beta := \log H/\log D_1\). We have
\[
\sum_{|h| \leq H, j \in \{1, \ldots, k-1, k+1, \ldots, 2k-1\}} \left| \sum_{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \in (D)} e_D(h \cdot (b, c)) \right|
\]
\[
\ll (a, D_1)^{1/2} \tau(D_1)^{O(1)} \left(\frac{D}{D_1}\right)^{2k-2} H^{2k-2} D_1^{k-1} \times (D_1 + D_1^{k(1/2-\beta)+\beta-1} + D_1^{(2k-1)(1/2-\beta)+2\beta-2} H^{1/2-\beta} + D_1 H^{2k(1/2-\beta)+2\beta-3} + D_1^{k(1/2-\beta)+\beta-1} H^{k(1/2-\beta)+\beta-1}).
\]

**Proof.** By applying Lemma 33 it suffices to show that
\[
\sum_{|h| \leq H, j \in \{1, \ldots, k-1, k+1, \ldots, 2k-1\}} \sum_{|h_k| \leq D_1, h_k \neq 0} p_i^{j_{i1}/2} \cdots p_\ell^{j_{i\ell}/2} \frac{1}{D_1^2} \sum_{h_{2k} \in X(j, \ldots, j, D_1)} 1
\]
\[
\ll \tau(D_1)^{O(1)} H^{2k-2} H^{\max\{1, k(1/2-\beta)+\beta-1, (2k-1)(1/2-\beta)+2\beta-2\}} (D_1/H)^{\max\{1, k(1/2-\beta)+\beta-1, (2k-1)(1/2-\beta)+2\beta-2\}}.
\]
Thus, we get
\[ \text{For } \beta \text{ we get } \]
\[ p_i = \max_{h_{2k}, \ldots, h_{k+1} \neq 0} H^{2k-2} \prod_{i \in I} p_i^{(r_i+s_i)} = \prod_{i \leq \ell} p_i^{\hat{r}_i + \hat{s}_i} \]
\[ \text{By Lemma 35 we reduce to bounding} \]
\[ \max_{0 \leq r_1, \ldots, r_{\ell}, s_1, \ldots, s_{\ell} \leq k} \frac{\tau(D_1)^{(1)}}{D_1^2} \prod_{i \leq \ell} p_i^{(r_i+s_i)^+ / 2} \sum_{h_{2k}, \ldots, h_{k+1} \neq 0} 1_{h_{2k}, \ldots, h_{k+1} \neq 0} H_{r_1, \ldots, r_{\ell}}(D_1) 1_{h_{2k+1} \neq 0} H_{s_1, \ldots, s_{\ell}}(D_1) \]
\[ \text{where the maximum runs over hyperplanes } H_{r_i}, H_{s_i} \text{ which define } Z_{r_i} \times Z_{s_i} \text{ as in (9.3)} \]
\[ \text{when applying Lemma 32 with } p = p_i. \text{ By symmetry we may assume that } (h_1, \ldots, h_{k-1}) \neq 0. \]
\[ \text{By applying Lemma 33 this is then bounded by} \]
\[ \max_{0 \leq r_1, \ldots, r_{\ell}, s_1, \ldots, s_{\ell} \leq k} \frac{\tau(D_1)^{(1)}}{D_1^2} \prod_{i \leq \ell} p_i^{(r_i+s_i)^+ / 2} D_1(r) D_1(s) \prod_{j \leq k-1} \left( 1 + \frac{H}{p_{i_j} \cdots p_{i_{j+1}}} \right) \]
\[ \prod_{j \leq k-1} \left( 1 + \frac{H}{p_{s_{j}} \cdots p_{s_{j+1}}} \right), \]
\[ \text{where } I \text{ denotes the set of indices } i \leq \ell \text{ such that } r_i = k. \]
\[ \hat{r}_i = 1_{r_i=0}, \quad \hat{s}_i := 1_{s_i=0}, \]
\[ \text{so that} \]
\[ D_1(r) D_1(s) = \prod_{i \leq \ell} p_i^{\hat{r}_i + \hat{s}_i} \]
\[ \text{By Lemma 35 we reduce to bounding} \]
\[ \max_{0 \leq r_1, \ldots, r_{\ell}, s_1, \ldots, s_{\ell} \leq k} \frac{\tau(D_1)^{(1)}}{D_1^2} H^{2k-2} \prod_{i \leq \ell} p_i^{(r_i+s_i)^+ / 2 + \hat{r}_i + \hat{s}_i - (r'_i + s'_i) \min(1, \beta) - 2}. \]

For \( \beta > 1 \) the exponent of \( p_i \) is
\[ (r_i + s_i)^+ / 2 + \hat{r}_i + \hat{s}_i - (r'_i + s'_i) - 2 \leq 1 \]
and we get
\[ \ll \tau(D_1)^{(1)} H^{2k-2} D_1 \]
which is sufficient. Thus, let \( \beta \leq 1 \). For \( r_i + s_i \leq 2 \) we have
\[ (r_i + s_i)^+ / 2 + \hat{r}_i + \hat{s}_i - (r'_i + s'_i) \beta - 2 \leq 1 \]
For \( r_i + s_i > 2 \) and \( r_i = 0 \) or \( s_i = 0 \) we get
\[ (r_i + s_i)^+ / 2 + \hat{r}_i + \hat{s}_i - (r'_i + s'_i) \beta - 2 \leq \max\{r_i(1/2 - \beta) + \beta - 1, s_i(1/2 - \beta) + \beta - 1\}. \]
For \( r_i + s_i > 2 \) and \( r_i \neq 0 \neq s_i \) we get
\[ (r_i + s_i)^+ / 2 + \hat{r}_i + \hat{s}_i - (r'_i + s'_i) \beta - 2 = (r_i + s_i)(1/2 - \beta) + 2, \beta - 2. \]
Thus, we get
\[ \ll \max_{0 \leq r_1, \ldots, r_{\ell}, s_1, \ldots, s_{\ell} \leq k} \tau(D_1)^{(1)} H^{2k-2} \prod_{i \leq \ell} p_i^{\max\{1, r_i(1/2 - \beta) + \beta - 1, s_i(1/2 - \beta) + \beta - 1, (r_i + s_i)(1/2 - \beta) + 2, \beta - 2\}}. \]
If \( \beta \in [1/2, 1] \), then the exponent is \( \leq 1 \) so we get
\[ \tau(D_1)^{(1)} H^{2k-2} D_1. \]
For $\beta < 1/2$ we may assume that $s_i \geq r_i$ (since $H \geq \prod_{i \in I} p_i$ with $I = \{ i : r_i = k \}$ ) and we get 

$$
\max_{0 \leq r_1, \ldots, r_s, t \leq k} \tau(D_1) H^{2k-2} \prod_{i \in I} p_i^\max\{1, s_i(1/2-\beta)+\beta-1, (r_i+s_i)(1/2-\beta)+2\beta-2\} 
$$

$$
\leq \max_{k-1 \leq r_1, \ldots, r_s \leq k} \tau(D_1) H^{2k-2} \prod_{i \in I} p_i^\max\{1, (1/2-\beta)+\beta-1, (r_i+k)(1/2-\beta)+2\beta-2\},
$$

since for $\beta < 1/2$ the first bound is increasing in $s_i$. Recalling that $i \in I$ iff $r_i = k$, the above is maximized if we choose $I_0$ so that $\prod_{i \in I_0} p_i$ is as close as possible to $H$ without exceeding it and for $i \not\in I_0$ we have $r_i = k-1$, so that we get 

$$
\ll \tau(D_1) H^{2k-2} \prod_{i \in I_0} p_i^\max\{1, (1/2-\beta)+\beta-1, (2k(1/2-\beta)+2\beta-2)\} \prod_{i \not\in I_0} p_i^\max\{1, (1/2-\beta)+\beta-1, (2k(1/2-\beta)+2\beta-2)\}
$$

$$
\ll \tau(D_1) H^{2k-2} H^{\max\{1, (1/2-\beta)+\beta-1, (2k(1/2-\beta)+2\beta-2)\}} (D_1/H)^{\max\{1, (1/2-\beta)+\beta-1, (2k(1/2-\beta)+2\beta-2)\}}.
$$

\[ \square \]

We also require the following Lang-Weil type bound for the number of points on the variety.

**Lemma 37.** For $a \neq 0 \pmod{p}$ we have

$$
\sum_{N(b_1, \ldots, b_k) = aN(c_1, \ldots, c_k) \pmod{p^n}} 1 = p^{2k-3}(1 + O(p^{-1/2})).
$$

**Proof.** By Lemma [30] it suffices to show that for $q = p^f$ we have for all $n \geq 1$

$$
\sum_{N(b_1, \ldots, b_k) = aN(c_1, \ldots, c_k) \in \mathbb{F}_q^n} 1 = (q^n)^{2k-3}(1 + O((q^n)^{-1/2})).
$$

Th form $N(b_1, \ldots, b_k)$ splits into linear factors of $\mathbb{F}_q$, so that after a linear change of variables we are counting solutions to

\begin{equation}
(9.6) \quad b_1' \cdots b_{k-1}' L(b_1', \ldots, b_{k-1}') = ac_1' \cdots c_{k-1}' L(c_1', \ldots, c_{k-1}')
\end{equation}

for

$$
L(b_1', \ldots, b_{k-1}') = \alpha_1 b_1' + \ldots + \alpha_{k-1} b_{k-1}'
$$

with some $\alpha_j \neq 0$. The number of points $(b_1', \ldots, b_{k-1}', c_1', \ldots, c_{k-1}')$ where $b_2' \cdots b_{k-1}' = 0$ or $c_2' \cdots c_{k-1}' = 0$ is clearly $\ll (q^n)^{2k-4}$, so we may assume that $b_2' \cdots b_{k-1}' \neq 0$ and $c_2' \cdots c_{k-1}' \neq 0$. Fix $b_2', \ldots, b_{k-1}'$ and $c_2', \ldots, c_{k-1}'$ consider the affine curve defined by the equation \((9.6)\) in variables $b_1', c_1'$. If the curve is non-singular, then by the Weil bound the number of points $(b_1', c_1')$ is $q^n + O((q^n)^{-1/2})$. Thus, it suffices to show that the curve is non-singular for all but $\ll (q^n)^{2k-4-1/2}$ of the variables $b_2', \ldots, b_{k-1}', c_2', \ldots, c_{k-1}'$. By the product rule we see that a point is singular only if

$$
b_2' \cdots b_{k-1}' (\alpha_1 b_1' + L(b_1', \ldots, b_{k-1}')) = c_2' \cdots c_{k-1}' (\alpha_1 c_1' + L(c_1', \ldots, c_{k-1}')) = 0.
$$

Since $b_2' \cdots b_{k-1}' \neq 0$ and $c_2' \cdots c_{k-1}' \neq 0$, a point is singular only if

$$
\alpha_1 b_1' + L(b_1', \ldots, b_{k-1}') = \alpha_1 c_1' + L(c_1', \ldots, c_{k-1}') = 0.
$$
so that \( b'_1 = L_1(b_2, \ldots, b_{k-1}) \) and \( c'_1 = L_1(c_2, \ldots, c_{k-1}) \) for some linear form \( L_1 \). Substituting these into (9.6) we get

\[
L(b_2, \ldots, b_{k-1})^2 b'_{k-1} = aL(c_2, \ldots, c_{k-1})^2 c'_{k-1},
\]

so that the number of such \( b_2, \ldots, b'_{k-1}, c_2, \ldots, c'_{k-1} \) is \( \ll (q^n)^{2k-5} \).

Similarly as in Section 5 using Lemma 37 with the Chinese remainder theorem we get the following.

**Lemma 38.** Assume that \( D/D_1 \leq Y \) and \( (a, D) \leq Y \). Then we have for some \( \varepsilon_d^{(k)}(a) \ll Y d \tau(d)^{O(1)} \)

\[
\sum_{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv a \mod N(c_1, \ldots, c_{k-1})}^{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv a \mod N(c_1, \ldots, c_{k-1})} \frac{D^{2k-3}}{d} \varepsilon_d^{(k)}(a) + O \left( \frac{D^{2k-3} \tau(D)^{O(1)}}{Y} \right).
\]

**Proof.** By the Chinese remainder theorem we have

\[
\sum_{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv a \mod N(c_1, \ldots, c_{k-1})}^{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv a \mod N(c_1, \ldots, c_{k-1})} 1 = \prod_{p^m \mid D} \sum_{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv a \mod N(c_1, \ldots, c_{k-1})}^{b_1, c_1, \ldots, b_{k-1}, c_{k-1} \equiv a \mod N(c_1, \ldots, c_{k-1})} 1.
\]

The claim follows by using Lemma 37 for \( m = 1 \) and the trivial bound \( \ll (p^m)^{2k-2} \) for \( m \geq 2 \), noting that

\[
(D/D_1)^{2k-2} \leq Y (D/D_1)^{2k-3}.
\]

\[\square\]

10. **The arithmetic information for Theorem 3**

10.1. **Type I sums.** We have the following Type I information, similar to Section 5.

**Proposition 39.** (Type I estimate). Let \( k \geq 3 \) and let \( \eta, \eta' > 0 \) and suppose that \( \eta \) is sufficiently small in terms of \( \eta' \). Let \( D \leq X^{1-1/((2k)\eta')} \). Then for any bounded coefficients \( \alpha(d) \) we have

\[
\sum_{d \sim D, n \sim X/d} \alpha(d)(a_{dn}^{(k)} - b_{dn}) \ll_{\eta} X^{1-\eta}.
\]

**Proposition 40.** (Fundamental lemma of the sieve). Let \( \eta, \eta' > 0 \) and suppose that \( \eta \) is sufficiently small in terms of \( \eta' \). Let \( W := X^{1/\log \log X} \). Let \( D \leq X^{1-1/((2k)\eta')} \). Then for any bounded coefficients \( \alpha(d) \) we have for any \( C > 0 \)

\[
\sum_{d \sim D, n \sim X/d} \alpha(d) \mathbf{1}_{(n, P(W)) = 1}(a_{dn}^{(k)} - b_{dn}) \ll_{\eta, C} X \log^{-C} X.
\]

where \( \eta > 0 \) depends on \( \eta' > 0 \).

**Proof.** The proofs of these Propositions are analogous to the proofs of Propositions 4 and 5 using Lemma 27 to bound the number of representations as \( N(b_1, \ldots, b_{k-1}) \) at the end of the large sieve bound, similarly as in Section 5.3.3. The only thing we have to check is that the main terms match for \( h = 0 \) after applying Poisson summation to sum over \( a \), which follows by the construction of the weight \( \Omega(b_1, \ldots, b_{k-1}) \). To
see this, we expand the definition of \( \Omega(b_1, \ldots, b_{k-1}) \) to get a finer-than-dyadic split for \( N(b_1, \ldots, b_{k-1}) \), so that the main term corresponding to Section 3.3.1 looks like

\[
\kappa_k \int_{1/2 \leq B \leq 2\sqrt{X} \atop N(b_1, \ldots, b_{k-1}) \equiv 0(c)} \psi_B(N(b_1, \ldots, b_{k-1})) \frac{\int \psi_B(u)du}{\int_{u \in [0, N^{1/2k})} \psi_B(N(u_1, \ldots, u_{k-1}))du} dB.
\]

The conditions \( b_i \in [0, X^{1/(2k)}] \) are handled by using a finer-than-dyadic decomposition (Section 3.1) to each of the variables \( b_i \), so that by Poisson summation (Lemma 7) the above sum is up to a negligible error term

\[
kappa_k \int_{1/2 \leq B \leq 2\sqrt{X} \atop N(b_1, \ldots, b_{k-1}) \equiv 0(c)} \psi_B(N(u_1, \ldots, u_{k-1})) du \frac{\int \psi_B(u)du}{\int_{u \in [0, N^{1/2k})} \psi_B(N(u_1, \ldots, u_{k-1}))du} dB.
\]

which matches the term coming from the comparison sequence \( b_\alpha \) after summing over \( c \) by the definition of the arithmetic factor \( \kappa_k \).

10.2. Type II sums. Our type II information is given by the following.

**Proposition 41.** (Type II estimate). Let \( MN \asymp X \) and for any small \( \eta > 0 \) let

\[
X^{1/(2k) + \eta} \ll N \ll X^{1/k - 2/(k(k+1)) - \eta}
\]

Let \( \alpha(m) \) and \( \beta(n) \) be bounded coefficients. Assume that \( \beta(n) \) is supported on square-free numbers with \( (n, P(W)) = 1 \) and satisfies the Siegel-Walfisz property (4.1) and assume that \( \alpha(m) \) is supported on \( (m, P(W)) = 1 \). Then for any \( C > 0 \)

\[
\sum_{\substack{m \sim M \atop n \sim N}} \alpha(m)\beta(n)(a_{mn}^{(k)} - b_{mn}) \ll_{\eta,C} X \log^{-C} X.
\]

Similarly as in Section 4 we can use Proposition 40 to reduce this to the following.

**Proposition 42.** Let \( W = X^{1/(\log \log X)^2} \). Let \( MN \asymp X \) and for any small \( \eta > 0 \) let

\[
X^{1/(2k) + \eta} \ll N \ll X^{1/k - 2/(k(k+1)) - \eta}
\]

Let \( \alpha(m) \) and \( \beta(n) \) be bounded coefficients. Assume that \( \beta(n) \) is supported on square-free numbers with \( (n, P(W)) = 1 \) and satisfies the Siegel-Walfisz property with main term equal to 0, that is, (4.3). Assume that \( \alpha(m) \) is supported on \( (m, P(W)) = 1 \). Then for any \( C > 0 \)

\[
\sum_{\substack{m \sim M \atop n \sim N}} \alpha(m)\beta(n)a_{mn}^{(k)} \ll_{\eta,C} X \log^{-C} X
\]

and

\[
\sum_{\substack{m \sim M \atop n \sim N}} \alpha(m)\beta(n)b_{mn} \ll_{C} X \log^{-C} X.
\]
Proof. The arguments are essentially the same as in Section 6. Denoting
\[ S(w, z) := \sum_{\text{Re}(\overline{w}z) = N(b_1, \ldots, b_{k-1})} \Omega(b_1, \ldots, b_k) \]
and
\[ U_{11}^{(k)} := \sum_{|z_1|^2, |z_2|^2 \sim N, \Delta \neq 0, (z_1, z_2) = 1} \beta_{z_1} \beta_{z_2} \sum_{b_1, \ldots, b_{k-1}} \Omega(b_1, \ldots, b_{k-1}) \Omega(b_{k+1}, \ldots, b_{2k-1}) \]
we need to show
\[ U_{11}^{(k)} \ll C X N \log^{-C} X. \]
The contribution from the diagonal $\Delta = 0$ is handled using $N \gg X^{1/(2k) + \eta}$ by the same argument as in Section 6.2 using Lemma 27 to bound the number of representations of any give integer by the form $N(b_1, \ldots, b_{k-1})$. The contribution from the diagonal $(z_1, z_2) > 1$ is bounded by essentially the same argument as in Section 6.3.

Thus, we are reduced to bounding the off-diagonal contributions
\[ U_{11}^{(k)} := \sum_{|z_1|^2, |z_2|^2 \sim N} \beta_{z_1} \beta_{z_2} \sum_{\Delta \neq 0} \Omega(b_1, \ldots, b_{k-1}) \Omega(b_{k+1}, \ldots, b_{2k-1}) \]
\[ F_M \left( \frac{z_2 N(b_1, \ldots, b_{k-1}) - z_1 N(b_{k+1}, \ldots, b_{2k-1})}{\Delta} \right)^2. \]
Let $Y = \log^C X$ and let $\Delta_2$ denote the powerful part of $\Delta$. We separate the contribution from $\Delta_2 > Y$ to get
\[ U_{11}^{(k)} = U_{11}^{(k)} + U_{11}^{(k)} \]
By similar arguments as in Section 6.4.5 for $V_{11}^{(1)}$, in place of (6.11) we get
\[ U_{11}^{(k)} \ll X^{1-1/(2k) + 2\epsilon} N^{3/2} + \frac{X N \log^{O(1)} X}{Y}, \]
which is sufficient since $N \ll X^{1/k - 10\epsilon}$.

For the main contribution $U_{11}^{(k)}$, similarly as in Section 6.4 we introduce smooth weights $\psi_{B_{j}}$ for the variables $b_{j}$ using Section 5.3.4 with $X^{1/(2k) - \eta'} \ll B_{j} \ll X^{1/(2k)}$ and denote
\[ H_j := X^\epsilon |\Delta| B_j^{-1}, \quad j \in \{1, \ldots, k-1, \ldots, k+1, \ldots, 2k-1\}. \]
We may restrict to $B_{j} \gg X^{1/(2k) - \eta'}$, for some $\eta'$ small compared to $\eta$, by estimating trivially the contribution from any $b_{j} \ll X^{1/(2k) - \eta'}$ before the application of Cauchy-Schwarz. This ensures that $H_j$ are all within $X^{\eta'}$ of each other and we set
\[ H := \max_{j \in \{1, \ldots, k-1, \ldots, k+1, \ldots, 2k-1\}} H_j. \]
Similarly in Section 6.3 we introduce a smooth partition for $|z_j|$ and arg $z_j$ and we restrict to $|\theta_1 - \theta_2$ (mod $\pi$)| $> \sqrt{\delta}$ (by bounding the complement as in Section 6.4.4), which implies $|\Delta| > \sqrt{\delta}N$ and therefore $H = X^{O(\eta + \epsilon)} N X^{-1/(2k)}$. Note that by $\Delta_2 \leq Y$ we then have $H < |\Delta_1|$ always. After this the smooth cross-condition $F_M$ may be
Thus, denoting
\[ \Omega(B_1, \ldots, B_{k-1}) \Omega(B_{k+1}, \ldots, B_{2k-1}) \ll X^{1/k}. \]

we need to show that
\[ U_1^{(1)}(B, N, \theta) \ll C \frac{X^2}{N \log^C X}. \]

By Poisson summation (Lemma 7) we obtain
\[
U_1^{(1)}(B, N, \theta) = X^{1/k} \sum_{z_1, z_2} \beta_{z_1, z_2}^{(1)} \sum_{b_1, \ldots, b_{2k-1}} \psi_{B_1}(b_1) \cdots \psi_{B_{2k-1}}(b_{2k-1}) \psi(N(b_1, \ldots, b_{2k-1}) = \alpha) \ll C \frac{X^2}{N \log^C X},
\]

where
\[
\beta_{z_1, z_2}^{(1)} = \beta_{z_1, z_2}^{(1)} \beta_{z_1, z_2}^{(2)} \beta_{z_1, z_2}^{(3)} \psi_{B_1}(b_1) \cdots \psi_{B_{2k-1}}(b_{2k-1})
\]

we have
\[ U_1^{(1)}(B, N, \theta) \ll C \frac{X^2}{N \log^C X}. \]
By $1 \leq H \leq N$ we have
\[
N(1 + H^{2k(1/2 - \beta) + 2\beta - 3}) \ll N + N^{(2k-1)(1/2-\beta)+2\beta-2} H^{1/2-\beta},
\]
\[
N^{k(1/2 - \beta) + \beta - 1} (1 + H^{k(1/2 - \beta) + \beta - 1}) \ll N^{k(1/2 - \beta) + \beta - 1} + N^{(2k-1)(1/2-\beta)+2\beta-2} H^{1/2-\beta}
\]
and we get (using $N^\beta = N X^{-1/(2k)+O(\eta')}$)
\[
\hat{U}_1^{(1)} (B, N, \theta) \ll \varepsilon X^{O(\varepsilon + \eta')} N^{1/k} N^{2+\eta} \left( N + N^{k(1/2 - \beta) + \beta - 1} + N^{(2k-1)(1/2-\beta)+2\beta-2} H^{1/2-\beta} \right)
\]
\[
\ll \varepsilon X^{O(\varepsilon + \eta')} X^{1/k} \left( N^{2+\eta} + N^{3k/2-(k-1)\beta} + N^{2k-3/2-(2k-1)\beta} H^{1/2-\beta} \right)
\]
\[
\ll \varepsilon X^{O(\varepsilon + \eta')} \left( X^{1/k} N^{2+\eta} + X^{1/k+(k-1)/(2k)} N^{k/2-1} + X^{1/k+(2k-1)/(2k)} N^{k/2-1} \right)
\]
\[
\ll \varepsilon X^{O(\varepsilon + \eta')} \left( X^{1/k} N^{2+\eta} + X^{1/2+1/(2k)} N^{k/2-1} + X^{1+1/(2k)} N^{-1/2} H^{1/2-\beta} \right),
\]
which is $\ll X^{1-\eta} N$ since $1/(2k) + \eta \leq \alpha \leq 1/k - 2/(k(k+1)) - \eta$ and since $\eta'$ is small compared to $\eta$.

\section{11. The sieve argument for Theorem 3}

Theorem 3 is a corollary of the following quantitative version, which approaches an asymptotic formula for large $k$. We could apply Vaughan’s identity if we had Type II information with the range $X^{1/(2k)} < N < X^{1/k}$, so that by Proposition 41 we are only missing a width of $2/(k(k+1))$ from this.

\textbf{Theorem 43.} For some small $\eta > 0$ we have
\[
S(\mathcal{A}^{(k)}, 2\sqrt{X}) = (1 + O((\log k)^{-\eta \log k})) S(\mathcal{B}, 2\sqrt{X}).
\]

\textbf{Proof.} Using the same argument as in Proposition 25 we get the following.

\textbf{Proposition 44.} Let $D := X^{1/(2k)-\eta}$,
\[
\gamma := \frac{1}{2k} - \frac{2}{k(k+1)} - 2\eta,
\]
and let $U \leq D$. Let $W = X^{1/(\log \log x)^2}$. Then for any bounded coefficients $\alpha(m)$ supported on $(m, P(W)) = 1$ we have
\[
\sum_{m \sim U} \alpha(m) S(\mathcal{A}^{(k)}_m, X^\gamma) = \sum_{m \sim U} \alpha(m) S(\mathcal{B}_m, X^\gamma) + O_{\eta,C}(X/\log^C X).
\]

Denote $Z = X^\gamma$. When we iterate Buchsteb’s identity, we can insert the conditions
\[
p_1 \cdots p_{j-1} p_j^2 \leq 4X
\]
with negligible error terms \( \ll X/\log^2 X \) similarly as in Section 7. Thus, after 2\( J \) applications of Buchstab’s identity, we get
\[
S(\mathcal{C}, 2X^{1/2}) = \sum_{0 \leq j \leq 2J - 1} (-1)^j \sum_{\substack{Z \leq p_j < \ldots < p_1 < 2\sqrt{X} \\ p_1 - p_{i-1} p_i^2 \leq 4X, i \leq j}} S(\mathcal{C}_{q_j}, Z) + \sum_{\substack{Z \leq p_{2j} < \ldots < p_1 < 2\sqrt{X} \\ p_1 - p_{i-1} p_i^2 \leq 4X, i \leq 2J}} S(\mathcal{C}_{q_{2j}, p_{2j}}) + O(X/\log^2 X),
\]
where we have denoted \( q_j := p_1 \cdots p_j \). By induction we see that the conditions \( p_1 \cdots p_{i-1} p_i^2 \leq 4X \) imply that
\[
p_1 \cdots p_i \leq 4X^{1-2^{-i}}.
\]
Indeed, this is true for \( i = 1 \) and by the induction hypothesis we have
\[
p_1 \cdots p_i = (p_1^2 \cdots p_i^2)^{1/2} = (p_1 \cdots p_{i-1})^{1/2}(p_1 \cdots p_{i-2} p_i^2)^{1/2} \leq (4X^{1-2^{i-1}})^{1/2}(4X)^{1/2} \leq 4X^{1-2^{-i}}.
\]
Hence, we can give an asymptotic formula for the terms \( j \leq 2J - 1 \) by Proposition 11 if for some small \( \eta' > 0 \) we take
\[
J := \eta' \log k,
\]
and our task is reduced to estimating
\[
S_{2J}(\mathcal{C}) := \sum_{\substack{Z \leq p_{2j} < \ldots < p_1 < 2\sqrt{X} \\ p_1 - p_{i-1} p_i^2 \leq 4X, i \leq 2J}} S(\mathcal{C}_{q_{2j}, p_{2j}}).
\]
Here we could iterate Buchstab’s identity provided that at each stage \( p_1 \cdots p_{i-1} p_i^2 \leq D \), but this is not necessary for us. Note also that this sum is too large to be dropped completely, since by Lemma 7.2 and Stirling’s approximation
\[
\frac{S_{2J}(\mathcal{B})}{S(\mathcal{B}, 2\sqrt{X})} = (1 + o(1)) \int \omega(\beta) \frac{d\beta_1 \cdots d\beta_{2J}}{\beta_1 \cdots \beta_{2J-1} \beta_{2J}} \geq \int_{\beta_1 = 1/2}^{\gamma} \cdots \int_{\beta_{i-1} = 1/2}^{\gamma} \frac{d\beta_1 \cdots d\beta_{2J}}{\beta_1 \cdots \beta_{2J-1} \beta_{2J}} \geq \frac{1}{(2J)!} \int_{\gamma < \beta_{2j}, \ldots, \beta_i < 1/2J+1} \frac{d\beta_1 \cdots d\beta_{2J}}{\beta_1 \cdots \beta_{2J}} \geq \frac{1}{(2J)!} \log \left( \frac{1}{(2J + 1)\gamma} \right)^{2J} \geq \frac{1}{(2J)!} (1 - \varepsilon)^{-2J} (\log k)^{2J} \geq \frac{1}{\sqrt{2J}} \left( \frac{e \log k}{(1 - \varepsilon)2J} \right)^{2J} \geq \frac{1}{\sqrt{2J}} \left( \frac{e \log 2}{(1 - \varepsilon)} \right)^{2J} \to \infty \quad \text{as} \quad k \to \infty,
\]
since the largest \( J \) we can take is \( J = \left[ \frac{\log k}{2\log 2} \right] + 1 \). The problem is that the primes \( p_i \) range over too long intervals and thus we seek to replace them by variables with shorter ranges. To do so we will use the so-called reversal of roles to ”break” any prime variable
\[
p_i \geq V := 2X/D = 2X^{1/2k+\eta} =: X^\beta.
\]
By breaking a prime variable we mean that we will be able to replace the sum by a similar sum where the prime variable is replaced by a product of variables each smaller than $V$. To see how the argument goes, suppose we want to estimate a sum of the form

$$
\sum_{m \sim M} \alpha(m) \sum_{p_1, p_2 \geq V} S(C_{mp_1p_2}, z).
$$

for some function $z = z(mp_1p_2)$ (we will have $z(n) = P^-(n)$, the smallest prime factor). We apply inclusion-exclusion (Buchstab’s identity) twice, first to the prime $p_1$ and then to the prime $p_2$ to get

$$
\sum_{m \sim M} \alpha(m) \sum_{n_1, p_2 \geq V} S(C_{mn_1p_2}, z) - \sum_{m \sim M} \alpha(m) \sum_{p_1, n_2 \geq V} S(C_{mp_1n_1n_2}, z)
$$

$$
+ \sum_{m \sim M} \alpha(m) \sum_{p_1, p_2, n_2 \geq V} S(C_{mp_1n_1p_2n_2}, z).
$$

In the first sum we have an asymptotic formula by Proposition 44 after re-arranging the variables, since $n_1 > V$ implies that the rest of the variables are $\leq 2X/V = D$. Similarly, in the second sum we have an asymptotic formula by Proposition 44, since $n_2 > V$. The third sum is similar to the original sum but we have replaced the large primes $p_1, p_2 > V$ by products of smaller variables and we can iterate the process for the new variables $n_i$ and any unbroken $p_i$ until all but possibly one of the variables $n_i$ satisfy $n_i \leq V$. We will then show that these remaining sums may be discarded and the error terms are sufficiently small by Lemma 26.

To make the above sketch rigorous, we begin by writing

$$
S_{2J}(C) = \sum_{0 \leq j \leq 2J} R_j(C),
$$

where

$$
R_j(C) := \sum_{Z \leq p_2, \ldots, p_J < 2\sqrt{X}} \sum_{p_1 \ldots p_{J-i} \leq 2X, i \leq 2J, p_{i+1} \leq V < p_i} S(C_{q_{2J}, p_{2J}})
$$

$$
\frac{1}{j! (2J - j)!} \sum_{Z \leq p_2, \ldots, p_J < 2\sqrt{X}} \sum_{p_1 \ldots p_{J-i} \leq 2X, i \leq 2J, p_{i+1} \leq V < p_i} S(C_{q_{2J}, p_{\min}}) + O(X/\log^{10} X),
$$

by using crude estimates and Type I information to bound the part where $p_i = p_j$ for some $i \neq j$. For $j \in \{0, 1\}$ we cannot apply reversal of roles, so we leave these as is. By
Lemma 26 the contribution from $j = 1$ is
\[
\ll \frac{1}{(2J-1)!} \int_{[\gamma,\beta]^{2J-1}} \frac{d\beta_1 \cdots d\beta_k}{\prod_{1 \leq i < j \leq k} (\beta - \beta_i)^{\gamma}} \ll \frac{k^2}{(2J-1)!} \left( \frac{\beta}{\gamma} \right)^{2J-1} \ll \frac{k^2}{(2J-1)!} \left( \frac{10}{k+1} \right)^{2J-1}
\]
for $k$ sufficiently large. The contribution from $j = 0$ is bounded similarly. Here and below the implied constants do not depend on $k$.

For $j \geq 2$ we iterate the reversal of roles to get terms with asymptotic formula and the remaining sums
\[
\Delta_j = \frac{1}{j! (J-j)!} \sum_{k_1, \ldots, k_j \geq 0} \frac{1}{k_1! \cdots k_j!} \int_{U_{2J+K}} \frac{d\beta_1 \cdots \beta_k}{\prod_{1 \leq i < j \leq k} (\beta - \beta_i)^{\gamma}}
\]
where $U_{2J+K} = [\gamma, \beta]^{2J-j} \times [\gamma, \beta]^{J-1} \times [\gamma, 1]^{K+1}$.

We obtain
\[
\Delta_j \ll \frac{1}{j! (J-j)!} \left( \frac{10}{k+1} \right)^{2J-1} \sum_{k_1, \ldots, k_j \geq 0} \frac{(\log k)^{k_1 + \cdots + k_j}}{k_1! \cdots k_j!}
\ll \frac{1}{j! (J-j)!} \left( \frac{10}{k+1} \right)^{2J-1}
\]
Summing over $j$ we get a total deficiency
\[
\sum_{j \leq 2J} \Delta_j \ll \frac{k^2}{(2J)!} \left( \frac{10}{k+1} \right)^{2J-1} \sum_{j \leq 2J} \frac{J!}{j! (J-j)!} k^j \ll \frac{k^2}{(2J)!} \left( \frac{10}{k+1} \right)^{2J-1} (1+ k)^{2J} \ll (\log k)^{-\eta \log k}
\]
for some $\eta > 0$, by using $J = \eta' \log k$. This completes the proof of Theorem 43.

Note that we could get a stronger bound by iterating reversal of roles to the primes $p'_{i_j} > V$ as well, which would improve the deficiency to $\ll k^{-\eta \log k}$.
References

[1] E. Bombieri. On exponential sums in finite fields. *Amer. J. Math.*, 88:71–105, 1966.
[2] P. Deligne. *Cohomologie étale*, volume 569 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1977. Séminaire de géométrie algébrique du Bois-Marie SGA 4\(\frac{1}{2}\).
[3] E. Fouvry and H. Iwaniec. Gaussian primes. *Acta Arith.*, 79(3):249–287, 1997.
[4] E. Fouvry and N. Katz. A general stratification theorem for exponential sums, and applications. *J. Reine Angew. Math.*, 540:115–166, 2001.
[5] J. Friedlander and H. Iwaniec. Asymptotic sieve for primes. *Ann. of Math. (2)*, 148(3):1041–1065, 1998.
[6] J. Friedlander and H. Iwaniec. The polynomial \(X^2 + Y^4\) captures its primes. *Ann. of Math. (2)*, 148(3):945–1040, 1998.
[7] G. Harman. *Prime-detecting sieves*, volume 33 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2007.
[8] R. Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer, 1977.
[9] D. R. Heath-Brown. Primes represented by \(x^3 + 2y^3\). *Acta Math.*, 186(1):1–84, 2001.
[10] D. R. Heath-Brown and C. Jia. The largest prime factor of the integers in an interval. II. *J. Reine Angew. Math.*, 498:35–59, 1998.
[11] D. R. Heath-Brown and X. Li. Prime values of \(a^2 + p^4\). *Invent. Math.*, 208(2):441–499, 2017.
[12] D. R. Heath-Brown and B. Z. Moroz. Primes represented by binary cubic forms. *Proc. London Math. Soc. (3)*, 84(2):257–288, 2002.
[13] E. Kowalski. Exponential sums over finite fields: elementary methods, https://people.math.ethz.ch/~kowalski/exponential-sums-elementary.pdf.
[14] S. Lang. *Algebraic Number Theory*. Graduate Texts in Mathematics. Springer, 1994.
[15] X. Li. Prime values of a sparse polynomial sequence. *to appear in Duke Math. J.*, 2021.
[16] J. Maynard. Primes represented by incomplete norm forms. *Forum Math. Pi*, 8:e3, 2020.
[17] D. H. J. Polymath. New equidistribution estimates of Zhang type. *Algebra Number Theory*, 8(9):2067–2199, 2014.
[18] K. Pratt. Primes from sums of two squares and missing digits. *Proc. Lond. Math. Soc. (3)*, 120(6):770–830, 2020.
[19] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 163 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, third edition, 2015. Translated from the 2008 French edition by Patrick D. F. Ion.
[20] A. Weil. On some exponential sums. *Proc. Nat. Acad. Sci. U.S.A.*, 34:204–207, 1948.
[21] A. Weil. *Sur les courbes algébriques et les variétés qui s’en déduisent*, volume 7 of *Publ. Inst. Math. Univ. Strasbourg*. Hermann et Cie., Paris, 1948.