The Supersingular ℓ-Isogeny Path and Endomorphism Ring Problems: Tighter Unconditional Reductions

Maher Mamah

Department of Combinatorics & Optimization, University of Waterloo, ON, Canada

Abstract. In this paper we show that the supersingular ℓ-isogeny path problem and the endomorphism ring problem are unconditionally equivalent under polynomial-time reductions given access to a factoring oracle. We show access to such oracle is sufficient to solve the Quaternion Path Problem of [KLPT14], removing the heuristic and GRH-based assumptions of previous works. Using Shor's factorization algorithm, this implies unconditional quantum polynomial-time equivalences between the ℓ-Isogeny Path and Endomorphism Ring problems. Our result strengthens a previous work of Wesolowski's reduction by removing the Generalized Riemann Hypothesis assumption and achieving polynomial-time equivalences with significantly lower polynomial degree, whereas the original reduction incurs a polynomial overhead of impractically high degree.

1. Introduction

Isogeny-based cryptography has attracted significant attention due to its potential resistance to quantum attacks. While polynomial-time quantum algorithms break widely used systems such as RSA and elliptic curve cryptography, the best known classical and quantum attacks on several isogeny-based schemes still require exponential time. The security of many such constructions relies on the presumed hardness of the *supersingular isogeny path problem*, which asks for an isogeny between two given supersingular elliptic curves.

As the scope of isogeny-based primitives has expanded, new computational problems have emerged. One particularly important example is the *endomorphism ring problem* (EndRing), which asks to compute the endomorphism ring of a supersingular elliptic curve. This problem has been studied since the work of Kohel [Koh96] and underlies the security of several cryptosystems, including [GPS16, CPV20, CGL09, DFK$^+$20b, BBC$^+$23].

In 2014, Kohel, Lauter, Petit, and Tignol [KLPT14] introduced a heuristic algorithm for an analogue of the isogeny path problem in quaternion algebras, known as the *quaternion path problem*. Building on this, Eisenträger et al. [EHL$^+$18] established a polynomial-time equivalence between the isogeny path and endomorphism ring problems under heuristic assumptions. Later, Wesolowski [Wes21] proved this equivalence unconditionally assuming the Generalized Riemann Hypothesis.

A closer examination of the approaches in [KLPT14, Wes21] reveals that integer factorization constitutes a fundamental bottleneck in solving the quaternion path problem. In this work, we show that this difficulty is intrinsic to the problem itself and requires new techniques to overcome.

Our methods are inspired by [Wes21] but differ in several key aspects. In particular, while previous approaches rely on variants of the Generalized Riemann Hypothesis and incur large polynomial overheads. We show that by working in the quantum setting one is able not only to remove the dependence on GRH but also obtain conceptually simpler and substantially tighter results.

Isogeny-based problems Fix primes p and ℓ (typically $\ell = 2$ or 3). Supersingular elliptic curves over $\mathbb{F}_{p^2}$ form the vertices of the supersingular ℓ-isogeny graph, whose edges correspond to ℓ-isogenies. This graph is $(\ell+1)$-regular, has about $p/12$ vertices, and is Ramanujan, hence exhibits rapid mixing. Many cryptographic constructions are based on the presumed hardness of navigating this graph.

The most restrictive formulation is the following.

Problem 1 (ℓ-IsogenyPath). Given supersingular elliptic curves $E, E'/\mathbb{F}_{p^2}$, find an isogeny $\varphi : E \to E'$ of degree ℓ^e for some $e \geq 0$.

This may be viewed as finding a path between E and E' in the ℓ-isogeny graph. Several cryptographic schemes (e.g. [CGL09, GPS20]) rely on the hardness of this problem. A relaxed variant removes the restriction on the degree.

Problem 2 (IsogenyPath). Given supersingular elliptic curves $E, E'/\mathbb{F}_{p^2}$, find an efficiently computable isogeny $\varphi : E \to E'$.

While Problem 2 is formally weaker, it is not a priori clear whether it is easier. Understanding the relationship between these problems leads naturally to endomorphism computations.

Indeed, endomorphism rings of supersingular elliptic curves over $\mathbb{F}_{p^2}$ are isomorphic to maximal orders in the quaternion algebra $B_{p,\infty}$. This motivates the following problems.

Problem 3 (EndRing). Given a supersingular elliptic curve $E/\mathbb{F}_{p^2}$, compute a generating set for (E) in efficient representation.

Problem 4 (MaxOrder). Given p and $E/\mathbb{F}_{p^2}$ supersingular, compute a maximal order $\mathcal{O} \subset B_{p,\infty}$ such that $\mathcal{O} \simeq (E)$.

It was shown in [PW24] that IsogenyPath and EndRing are unconditionally equivalent under polynomial-time reductions. However, these reductions are not tight (see [ABF+25]) and do not address the restricted ℓ-IsogenyPath problem, which remains open [HLMW25].

These relationships are most naturally studied in the quaternionic setting, where isogenies correspond to ideals between maximal orders, leading to the following problem.

Problem 5 (QuaternionPath). Given maximal orders $\mathcal{O}_1, \mathcal{O}_2 \subset B_{p,\infty}$ and a set $\mathcal{N} \subset \mathbb{Z}_{>0}$, find a left $\mathcal{O}_1$-ideal I with $\mathrm{Nrd}(I) \in \mathcal{N}$ and right order $\mathcal{O}_R(I) \simeq \mathcal{O}_2$. If $\mathcal{N} = \{\ell^e\}$, we write ℓ-QuaternionPath.

Contributions & Organization of the paper. Section 2 reviews background material on elliptic curves, the supersingular setting, and quadratic forms. Section 2 develops the analytic tools required for representing integers and primes by quaternary quadratic forms and plays a foundational role in the remainder of the paper.

In Section 3, we present algorithms for sampling vectors $\mathbf{v} \in \mathbb{Z}^4$ such that a special quaternary quadratic form $Q(\mathbf{v})$ is prime, equivalently sampling ideals of prime norm within a fixed ideal class of a maximal order. Section 4 introduces the main algorithm for solving norm form equations of the type

$$af(s,t) + bg(x,y) = n,$$

where f and g are definite quadratic forms.

Section 5 contains the principal result of the paper, giving an algorithmic solution to the QuaternionPath problem and its variant B-PSQuaternionPath. Finally, we establish the equivalence between ℓ-IsogenyPath and MaxOrder in Section 6, between EndRing and MaxOrder in Section 7, and between IsogenyPath and ℓ-IsogenyPath in Section 8.

A Note on previous reductions. From a practical standpoint, existing reductions exhibit significant limitations. As previously mentioned, the conditional reductions of [Wes21] as stated and proved, while polynomial-time in theory, are far from tight due to the constants not being tracked explicitly at certain steps. More precisely, given an algorithm that solves the EndRing problem in time $T(n)$, where n denotes the input size, the corresponding reduction in [Wes21] yields an algorithm for ℓ-IsogenyPath running in time $T(n) + q(n)$, where $q(n)$ is a polynomial overhead with $\deg(q) > 12320$. Consequently, despite being polynomial-time asymptotically, such reductions incur an overhead that renders them impractical for cryptographic parameter sizes. We refer the interested reader to 5 for more details regarding the claimed degree overhead.

Factoring oracle. We assume access to a factoring oracle $\mathscr{O}$ for factoring large integers, which is required in Section 4 onward. Factoring appears unavoidable in addressing the QuaternionPath problem, as previous approaches rely on Cornacchia's algorithm for solving equations of the form $x^2+dy^2 = m$, which requires the factorization of m and becomes inefficient when m is highly composite. However, it might be the case here that we need to factor highly composite integers. For the purpose the author in a separate work [Mah25], introduced an algorithm for solving $f(s,t) = m$ for arbitrary definite quadratic forms f that remains efficient when $\text{disc}(f) = \text{polylog}(m)$, as is the case here. Given our interest in post-quantum cryptography, Shor's algorithm naturally plays the role of our necessary oracle.

1.1. Background on Elliptic curves.

We refer the reader to [Sil86] for a detailed discussion on this topic. An elliptic curve E over a field K of characteristic $p > 3$ is defined by the equation $y^2 = x^3 + Ax + B$ for $A, B \in K^\times$ with $4A^3 + 27B^2 \neq 0$. The K-rational points of E are $(x, y) \in K$ satisfying the aforementioned equation along with an identity called the point at infinity ∞.
An isogeny $\varphi : E_1 \to E_2$ defined over K is a non-constant map which is also a group homomorphism that takes ∞_1 to ∞_2. The degree of an isogeny is the degree of φ as a rational map. Every isogeny of degree $n > 1$ can be factored into a composition of isogenies of prime degrees such that the product of the degrees equals n. If $deg(\varphi) = d$ then there exist a unique isogeny called the dual isogeny denoted $\hat{\varphi} : E_2 \to E_1$ such that $\varphi\hat{\varphi} = [d]$ where $[d] : E_2 \to E_1$ is the multiplication-by-d map. An isomorphism is an isogeny $\iota : E_1 \to E_2$ of degree 1. We say that E_1 and E_2 are isomorphic over k (an extension of K) if there is an isomorphism between them that is defined over k. The j invariant of E as defined by the above equation is $j(E) = \frac{265 \cdot 27 \cdot A^3}{4A^3+27B^2}$. An isogeny with a degree that is coprime to p is uniquely determined by its kernel. Given the kernel, the isogeny can be computed in polynomial time with respect to the degree of the isogeny and $\log(p)$ using Vélu's formula [Vel71].

Endomorphism ring and Supersingular versus Ordinary curves An isogeny from E into itself is called an endomorphism of E. If E is defined over $\mathbb{F}_q$ then an endomorphism of E will be defined over a finite extension of $\mathbb{F}_q$. The set of endomorphisms of E with the zero map form a ring under pointwise addition and composition which is called the endomorphism ring of E and denoted by $\text{End}(E)$. If $\text{End}(E)$ is isomorphic to an order in a imaginary quadratic field then the curve is said to be *ordinary*. Otherwise if $\text{End}(E)$ is isomorphic to a maximal order in a Quaternion algebra in $B_{p,\infty}$ (defined in section 1.2) then E is said to be *supersingular*. Every supersingular elliptic curve over a field of characteristic p has an isomorphic curve that is defined over $\mathbb{F}_{p^2}$ because the j-invariant of such a curve is in $\mathbb{F}_{p^2}$.
Up to k-isomorphism, all supersingular elliptic curves are defined over $\mathbb{F}_{p^2}$, with approximatley $\lfloor \frac{p}{12} \rfloor$ such curves. Let ℓ be a prime different from p. The supersingular ℓ-isogeny graph (for p) consists of vertices corresponding to these supersingular elliptic curves (up to isomorphism), with an edge connecting E_1 to E_2 for each l-isogeny from E_1 to E_2. This graph is regular with degree $\ell + 1$ since each curve E possesses $\ell + 1$ subgroups H of order ℓ, each generating an isogeny with kernel H. The l-isogeny graph is Ramanujan, meaning that random walks on this graph converge rapidly to the uniform distribution, and any two curves in the graph are linked by an isogeny of degree ℓ^m with $m = O(\log p)$ (see [Koh96], Theorem 79).

1.2. Quaternion Algebras and the Deuring Correspondence

For a detailed account on the arithmetic of quaternion algebras we refer the reader to [Voi21]. For $a, b \in \mathbb{Q}^\times$, let $B(a, b)$ denote the quaternion algebra over $\mathbb{Q}$, with basis $1, i, j, ij$, i.e.

$$B(a, b) = \mathbb{Q} + \mathbb{Q}i + \mathbb{Q}j + \mathbb{Q}ij,$$

such that $i^2 = a$, $j^2 = b$ and $ij = -ji$. The quaternion algebra B has a *canonical involution* that sends $\alpha = x_1 + x_2i + x_3j + x_4ij$ to $\bar{\alpha} = x_1 - x_2i - x_3j - x_4ij$, and we define both the reduced trace

and norm of an element α in B by

$$\mathrm{Trd}(\alpha) = \alpha + \bar{\alpha} = 2x_1, \quad \mathrm{Nrd}(\alpha) = \alpha\bar{\alpha} = x_1^2 - ax_2^2 - bx_3^2 + abx_4^2.$$

We say Λ is a lattice in B if $\Lambda = \mathbb{Z}x_1 + \mathbb{Z}x_2 + \mathbb{Z}x_3 + \mathbb{Z}x_4$ where the x_i's form a basis for the vector space B over $\mathbb{Q}$.
If $I \subset B$ is a lattice, then the reduced norm of I, $\mathrm{Nrd}(I) = \gcd(\mathrm{Nrd}(\alpha)|\alpha \in \Lambda)$. We associate to Λ the normalised quadratic map

$$q_\Lambda : \Lambda \to \mathbb{Z}, \quad \alpha \mapsto \frac{\mathrm{Nrd}(\alpha)}{\mathrm{Nrd}(\Lambda)}$$

and notice that $\frac{\mathrm{Nrd}(\alpha)}{\mathrm{Nrd}(\Lambda)} \in \mathbb{Z}$ as $\mathrm{Nrd}(\Lambda)|\mathrm{Nrd}(\alpha)$. The quaternion algebra B is an inner product space with respect to the bilinear form

$$\langle x, y\rangle = \frac{1}{2}(\mathrm{Nrd}(x+y) - \mathrm{Nrd}(x) - \mathrm{Nrd}(y)),$$

and the basis $\{1, i, j, ij\}$ is an orthogonal basis with respect to this inner product.
An order $\mathcal{O}$ in B is a full rank lattice that is also a subring. It is maximal if it is not contained in any other order. For an lattice Λ we define the *left order* and *right order* of Λ to be

$$\mathcal{O}_L(\Lambda) = \{\alpha \in B|\ \alpha\Lambda \subseteq \Lambda\}, \quad \mathcal{O}_R(\Lambda) = \{\alpha \in B|\ \Lambda\alpha \subseteq \Lambda\}\}.$$

If $\mathcal{O}$ is a maximal order and I is a left $\mathcal{O}$-ideal, then $\mathcal{O}_R(I)$ is also a maximal order. Given two maximal order $\mathcal{O}$ and $\mathcal{O}'$, then there exist a lattice I, called a connecting ideal, such that $\mathcal{O}_L(I) = \mathcal{O}$ and $\mathcal{O}_R(I) = \mathcal{O}'$. Additionally, we say left $\mathcal{O}$-ideals I and J are equivalent if there exists an $\alpha \in B^\times$ such that $I = \alpha J$. The set of classes of this equivalenece relation is called the *left*-ideal class set of $\mathcal{O}$ and is denoted by $\mathrm{Cls}(\mathcal{O})$.
For a maximal order $\mathcal{O}$ in $B_{p,\infty}$, the discriminant is $\mathrm{disc}(\mathcal{O}) = p^2$. Moreover, for any left $\mathcal{O}$-ideal I, we have $\#(\mathcal{O}/I) = \mathrm{Nrd}(I)^2$, and the associted normalized quadratic map q_I has determinant p^2.
Let p be a prime and $B_{p,\infty}$ be the unique quaternion algebra ramified exactly at at p and ∞. The following lemma from [EHL$^+$18] gives the structure of $B_{p,\infty}$.

Lemma 1.1 ([EHL$^+$18], proposition 1)**.** *Let $p > 2$, then $B_{p,\infty} = \left(\frac{-q,-p}{\mathbb{Q}}\right)$ where*

$$q = \begin{cases} 1 & \text{if } p \equiv 3 \pmod 4, \\ 2 & \text{if } p \equiv 5 \pmod 8, \\ q_p & \text{if } p \equiv 1 \pmod 8. \end{cases}$$

where q_p is the smallest prime such that $q_p \equiv 3 \mod 4$ and $\left(\frac{p}{q_p}\right) = -1$. Assuming GRH, we have that $q_p = O(\log^2 p)$ and hence can be computed in polynomial time in $\log p$.

Remark 1. Notice that this is the only case were the generalised Riemann hypothesis is assumed which is not directly related to the problems at hand, but rather to the very structure of the quaternion algebra we're working in. The problem of the least quadratic non-residue has been studied extensively in the past century especially by Erdős and Linnik. It's worthy to note that GRH presents an efficiently computable result and unconditional results hold yet they are a bit far from practical. For instance the following lemma due to Linnik [Lin41] tells us that not a lot of primes p have an exceptionally large least quadratic non-residue q_p.

Theorem 1.2 (Linnik)**.** *Let $\varepsilon > 0$, then there exists a constant C_ε such that for all N there are at most C_ε primes $p \leq N$ with $q_p > N^\varepsilon$.*

In the remainder of this article, we will assume that q_p has been found and its $O(\log^2 p)$. Addtionally we assume that the algebra $B_{p,\infty}$ is identified by the pair $(-q, -p)$ as defined in Lemma 1.1, along with the corresponding basis $(1, i, j, ij)$. We now identify a special maximal order, which we call an extremal order that has been first studied by [KLPT14], which will be crucial for our applications.

Lemma 1.3. *[EHL+18, Proposition 1] For any $p > 2$, the quaternion algebra $B_{p,\infty}$ contains the maximal order*

$$\mathcal{O}_0 = \begin{cases} \langle 1, i, \frac{i+ij}{2}, \frac{1+j}{2} \rangle & \text{if } p \equiv 3 \mod 4 \\ \langle 1, i, \frac{2-i+ij}{4}, \frac{-1+i+j}{2} \rangle & \text{if } p \equiv 5 \mod 8 \\ \langle \frac{1+i}{2}, \frac{j+ij}{2}, \frac{i+cij}{q}, ij \rangle & \text{if } p \equiv 1 \mod 8 \end{cases}$$

where in the last case c is a constant such that $q | c^2 p + 1$. Assuming the Generalized Riemann Hypothesis (GRH), the maximal order O_0 contains the suborder $R + Rj$ with an index of $O((\log p)^2)$, where R is the ring of integers of $\mathbb{Q}(i)$. If ω is a reduced generator of R, then the reduced norm

$$\mathrm{Nrd}(s + t\omega + xj + y\omega j) = f(s, t) + pf(x, y),$$

where f is a principal, primitive, positive definite, integral binary quadratic form with discriminant $\mathrm{disc}(\mathbb{Q}(i))$.

This norm form will be an extremely important tool in our applications, which will allow us to compute elements in the order of some desired norm.

1.3. Deuring Correspondence.

For a detailed account of on the Deuring correspondence, see chapter 42 of [Voi21]. Let E be a supersingular elliptic curve defined over $\mathbb{F}_{p^2}$, Deuring showed that endomrophism rings of supersingular elliptic curves are in bijection with maximal order in $B_{p,\infty}$ up to conjugation with norms and traces on the endomorphisms side correspond to norms and traces in the quaternion algebra.

Deuring's correspondence establishes that, for a supersingular elliptic curve E defined over $\mathbb{F}_{p^2}$, every isogeny $\varphi : E \to E'$ of degree n can be associated with a left ideal I in $\mathrm{End}(E)$ whose norm is n. According to Kohel's work [Koh96, Section 5.3], every nonzero left ideal of $\mathrm{End}(E)$ arises from such an isogeny.

To construct an isogeny from a given left ideal, one considers a nonzero integral left ideal $I \subset \mathrm{End}(E)$ and defines $E[I]$ as the scheme-theoretic intersection of the kernels of all endomorphisms $\alpha \in I$, i.e.

$$E[I] = \bigcap_{\alpha \in I} \ker(\alpha).$$

This subgroup then determines an isogeny $\varphi_I : E \to E/E[I]$. We also have for each isogeny $\varphi : E \to E'$, the corresponding ideal which is given by $I_\varphi = \mathrm{Hom}(E', E)\varphi$, and its a left $\mathrm{End}(E)$-ideal and a right $\mathrm{End}(E')$-ideal. Interestingly, $I_{\varphi_I} = I$ and $\varphi_{I_\varphi} = \varphi$. Moreover, if the reduced norm of I is coprime to p, φ is separable, and $E[I] = \{P \in E(\bar{\mathbb{F}}_{p^2}) : \alpha(P) = 0, \ \forall \alpha \in I\}$. Lastly, the translations from ideals to isogenies and isogenies to ideals can be done efficiently in $\log p$ and the maximum of prime powers dividing the norm or the degree of the ideal and isogeny respectively. This is why, in isogeny based cryptography, we're interested in powersmooth isogenies. For a detailed account of that see lemma 5 of [GPS20] and lemma 7.1 of [Wes21].

Remark. Even though its possible to translate an ideal to an isogeny in rigorous polynomial time even if the degree is not powersmooth by going to an 8-dimensional isogeny (see for instance [PR24]), for purposes of efficiency we will stick to the above methods in [GPS20].

1.4. Brandt Graph and Supersingular ℓ-isogeny graph.

Definition 1.4 (Brandt Graph)**.** *Let p be a prime number, and $\mathcal{O}$ a maximal order in $B_{p,\infty}$. Let I, J be two left $\mathcal{O}$-ideals. We say J is an ℓ-neighbor of I if $J \subseteq I$ and $\mathrm{Nrd}(J) = \ell \cdot \mathrm{Nrd}(I)$. The ℓ-Brandt graph is the graph with vertices $\mathrm{Cls}(\mathcal{O})$ and an edge from $[I_i]$ to $[J]$ for each ℓ-neighbor $J \subseteq I_i$, where $(I_i)_{i=1}^{\#\mathrm{Cls}(\mathcal{O})}$ is a list of ideal class representatives.*

The ℓ-Brandt graph is in fact isomorphic to the ℓ-isogeny graph through the Deuring correspondence. One can perform random walks on the ℓ-isogeny graphs by randomly chooing an ℓ-subgroup and computing the corresponding isogeny. Given that, its also possible to do random walks on ℓ-Brandt graphs. Since ℓ-isogeny graphs are Ramanujan, we can say the same thing about ℓ-Brandt graphs and, in fact, we have the following precise result.

Theorem 1.5 ([Wes21]. Theorem 6.2)**.** *Let p be a prime number, and $\mathcal{O}$ a maximal order in $B_{p,\infty}$. Let N_p be the size of the ideal class set of $\mathcal{O}$. Let I be the ideal obtained from a random walk of norm $n = \prod_i \ell_i^{e_i}$. Then, for any ideal class C, we have*

$$\left|\Pr[I \in C] - \frac{1}{N_p}\right| \leq \prod_i \left(\frac{2\sqrt{\ell_i}}{\ell_i + 1}\right)^{e_i}.$$

2. Representation of integers by Quadratic forms

In this section we study representations of integers by quadratic forms. We will study number of triplets satisfying $az + bg(x, y) = n$ with z represented by some form f and g is definite binary form. Additionally, we will need some estimates for the number of vectors $v \in \mathbb{Z}^4$ such that a quaternary form $Q(v)$ is a prime. The bounds in this section, particularly theorem 2.5 and proposition 3.4, will be critical for designing our main algorithm in section 4. The notation and terminology used in this section are standard and can be found in classical references. To avoid repeating well-known material, and since our focus aligns with the applications considered by Wesolowski, we refer the reader to Section 2.1 of [Wes21].

3.1. Number of representations of an integer by a quadratic form. We let Q be a primitive, integral, positive definite quaternary quadratic form and we denote by $r_Q(n) = \#\{x \in \mathbb{Z}^4 | Q(x) = n\}$. We start with a theorem from ([IK04], Theorem 20.9) which gives an asymptotic for $r_Q(n)$, and though the result is more general, we'll state it for 4 variables for our purposes.

Theorem 2.1. *Let $Q(x)$ be an integral positive definite quadratic form of 4 variables. Then for any $n > 0$, the number of integral solutions $x = (x_1, x_2, x_3, x_4)$ to $Q(x) = n$ satisfies*

$$r(n, Q) = \frac{(2\pi)^2 n}{|det(A)|^{1/2}} \mathfrak{S}(n, Q) + O_{\varepsilon,Q}(n^{\frac{3}{4}+\varepsilon}), \tag{1}$$

where $\mathfrak{S}(n, Q)$ is the so-called Singular series defined as

$$\mathfrak{S}(n, Q) = \sum_{c=1}^{\infty} c^{-4} g_c(n, Q)$$

with

$$g_c(n, Q) = \sum\nolimits^{*}_{d \pmod c} \sum\nolimits_{h \pmod c} exp\left(\frac{2\pi i d}{c}(Q(h) - n)\right),$$

A is the Gram matrix of Q and $\varepsilon > 0$, with the implied constants depending on ε and the quadratic form Q.

Note that ε with it's implied constants cause no harm, but to eliminate the dependence of the implied constants on the quadratic form, we will modify the proof of Theorem 2.1 in [IK04]. The next theorem extracts the dependency of the implied constants on Q, explicitly.

Theorem 2.2. *Let $Q(x)$ be an integral positive definite quadratic form of 4 variables. Then for any $n > 0$, the number of integral solutions $x = (x_1, x_2, x_3, x_4)$ to $Q(x) = n$ satisfies*

$$r(n, Q) = \frac{(2\pi)^2 n}{|det(A)|^{1/2}} \mathfrak{S}(n, Q) + O_{\varepsilon}(N^2 n^{\frac{3}{4}+\varepsilon}), \tag{2}$$

where A is the Gram matrix of Q, and N is the level of Q, $\mathfrak{S}(n, Q)$ and $g_c(n, Q)$ are as defined in Theorem 2.1.

Proof. Following the proof of the theorem 20.9 of [IK04], one notices that the only occurrence of the implied constants depending on the Quadratic form Q is coming from ([IK04][Lemma 20.1) which states that

$$G_m\left(\frac{d}{c}\right) = \sum_{h \pmod c} exp\left(\frac{2\pi i d}{c}(Q(h) + h^t m)\right) \ll_Q c^{\frac{r}{2}}.$$

Working out the details, we obtain

$$\begin{aligned}
\left|G_m\left(\frac{d}{c}\right)\right|^2 &= \sum\nolimits^*_{x,y \pmod c} exp\left(\frac{2\pi i d}{c}(Q((x) - Q(y) + (x-y)^t m)\right) \\
&= \sum\nolimits^*_{y,z \pmod c} exp\left(\frac{2\pi i d}{c}(y^t Az - Q(z) + z^t m)\right) \\
&\leq c^r \cdot |\{z(mod\ c) \mid Az \equiv 0\ (mod\ c)\}| \\
&\leq c^r \cdot |\{z(mod\ c) \mid NA^{-1}Az \equiv 0\ (mod\ c)\}| \\
&= c^r \cdot |\{z(mod\ c) \mid Nz \equiv 0\ (mod\ c)\}| \\
&= c^r \cdot gcd(N,c)^r \\
&\leq c^r N^r
\end{aligned}$$

taking square roots both sides, we get the desired result. Now substituting in (20.112) in the proof of [IK04], we get the statement of the theorem. □

Now in order for the estimate to make sense, we need to ensure $\mathfrak{S}(n,Q) > 0$.

Lemma 2.3. *Given an integral positive definite Quaternary quadratic form $Q(x) = x^T Ax$, if*

$$Q(x) \equiv n\ (mod\ 2^7 |det(A)|^3) \tag{3}$$

is solvable, then there exist an absolute constant $\kappa > 0$

$$\mathfrak{S}(n,Q) \geq \frac{\kappa}{\log\log(n)}. \tag{4}$$

Proof. Denote by $\chi(q) = \left(\frac{|det(A)|}{q}\right)$ the legendre symbol. If 3 is solvable, then from ([Iwa12],11.76) we have

$$\mathfrak{S}(n,Q) \asymp \prod_{p|n}\left(1 + \frac{\chi(p)}{p}\right),$$

i.e there exist an absolute constant c_1 such that

$$\mathfrak{S}(n,Q) \geq c_1 \prod_{p|n}\left(1 + \frac{\chi(p)}{p}\right) \geq c_1 \prod_{p|n}\left(1 - \frac{1}{p}\right) = c_1 \frac{\varphi(n)}{n} \geq \frac{\kappa}{\text{loglog}(n)}.$$

where we have used the fact that $\varphi(n) = n\prod_{p|n}\left(1 - \frac{1}{p}\right) \geq \frac{c_2 n}{\log\log(n)}$. This concludes the proof of the lemma. □

Remark. In the case that (3) is solvable, ([Iwa12],11.76) states that $\mathfrak{S}(n)$ actually ranges between $(\log\log 3n)^{-1}$ and and $\log\log 3n$.

For our applications, we'll need the following immediate corollary.

Corollary 2.4. *For every $\varepsilon > 0$ and $n \gg_\varepsilon N^{8+\varepsilon}|det(A)|^{2+\varepsilon}$, where N is the level of the form $Q = x^t Ax$,*

$$r(n,Q) \gg_\varepsilon \frac{1}{|det(A)|^{1/2}} \frac{n}{\log\log n}, \tag{5}$$

provided that the congruence in 3 is solvable.

Remark. We note that Since $A^{-1} = \mathrm{adj}(A)/\det A$ with $\mathrm{adj}(A) \in M_n(\mathbb{Z})$ we have $N \leq \det A$.

3.2. Number of triplet solutions to $az + bg(x, y) = n$. Here we study the number of integral solutions (z, x, y) to $az + bg(x, y) = n$ with the condition that z is represented by some form f. More precisely, let f and g be primitive, positive definite, integral, binary quadratic forms, a, b, n positive integers such that $(a, 2nb) = (b, 2n) = 1$ and let $\Delta_f = \mathrm{disc}(f)$ and $\Delta_g = \mathrm{disc}(g) = f_0^2 D$ be the fundamental and non-fundamental discriminants of f and g respectively. Additionally assume $a|\Delta_g$ We seek a lower bound on the quantity

$$S(n) = \{(x, y, z) \in \mathbb{Z}^3 \mid az + bg(x, y) = n, \quad with \;\; z = f(s, t) \text{ for some } (s, t) \in \mathbb{Z}^2\}. \tag{6}$$

Note for general $(z, x, y) \in \mathbb{Z}^3$ with $z > 0$, it is a classical result on number of lattice points in an ellipse, that

$$\#\{(z, x, y) \in \mathbb{Z}^3 \mid az + bg(x, y) = n, \; z > 0\} \sim \frac{\pi n}{ab|\mathrm{disc}(g)|^{1/2}}.$$

We write $r_f(n) = \#\{(x, y) \in \mathbb{Z}^2 \mid f(x, y) = n\}$ and similarly $r_g(n)$ for g. It is also a known fact that

$$r_f(k) \ll \tau(k) \ll \exp\left(\frac{C \log k}{\log\log k}\right) \tag{7}$$

for some absolute constant C, and similarly for $r_g(k)$. In particular, such an upper bound on the number of representations of an integer k by binary definite form is solely dependent on the structure of k.
Suppose now that there are no local obstructions in the quaternary form equation $Q(s, t, x, y) = af(s, t) + bg(x, y) = n$, such that $\mathfrak{S}(n) > 0$ (as in lemma 2.3). Notice that

$$r_Q(n) = \sum_{az+bm=n} r_f(z) r_g(m),$$

then by using the trivial bound (7) and corollary 2.4, we obtain a simple bound

$$\#S(n) \gg_\varepsilon \frac{n \exp\left(-C \log n/\log\log n\right)}{|\det Q|^{1/2}}$$

for $n \gg_\varepsilon N^{8+\varepsilon}|det(A)|^{2+\varepsilon}$. However, for the sake of applications, we need a more precise lower bound with a logarithmic saving. We can improve this by observing that the number of k with $r_f(k)$ large is quite rare. In fact, the main result of this subsection is the following theorem. Recall $\omega(\cdot)$ is the omega prime divisor counting function.

Theorem 2.5. *Let f and g be primitive, integral, positive definite, reduced binary quadratic forms with $\Delta_f \leq -3$ fundamental, and let a, b, n be positive integers such that $(a, 2nb) = (b, 2n) = 1$ and $\omega(a) = 0$ or 1. Assume the congruence (3) is solvable such that $\mathfrak{S}(n) > 0$, and $n \gg_\varepsilon a^{30+\varepsilon} b^{20+\varepsilon} |\Delta_f|^{10+\varepsilon} |\Delta_g|^{51/4+\varepsilon}$ for any $\varepsilon > 0$. Then,*

$$\#S(n) \gg_\varepsilon \left(\frac{C_{f,g,a}^{-1/\alpha}}{a\, b\, |\Delta_f|^{1+1/\alpha} |\Delta_g|}\right) n (\log n)^{-\gamma},$$

where $\gamma \approx 7.532..$, $\alpha = 1/\ln 2$, and

$$C_{f,g} = 4^{\omega(\Delta_g)} \log^2(a\Delta_g).$$

The proof of this theorem is conceptually straightforward. However, it requires an upper bound for the quantity

$$\sum_{\substack{z \leq X \\ z \equiv k \pmod m}} r_g^2(z), \qquad (k, m) = 1.$$

We record the required estimate in the following proposition. Since the proof is technically involved and relies on Heath–Brown's δ-method, it is deferred to the Appendix in section 10.

Proposition 2.6. *Assuming the hypothesis of theorem 2.5, then for $X \gg_\varepsilon a^{30+\varepsilon}|\Delta_g|^{51/4+\varepsilon}$ we have*

$$\sum_{\substack{z\leq X\\ z\equiv k \bmod a}} r_g^2(n) \ll \frac{4^{\omega(\Delta_g)}\log^2(a\Delta_g)}{a^2|\Delta_g|}X\log X,$$

The following lemma is the key ingredient in the proof of the main theorem.

Lemma 2.7. *Let $T>1$ and $\alpha>0$. Then*

$$\sum_{\substack{az+bm=n\\ r_f(z)>T}} r_f(z)r_g(m) \ll_{\alpha,\varepsilon} T^{-\alpha/2}\,\frac{n}{ab}\left(\frac{4^{\omega(\Delta_g)}\log^2(a\Delta_g)}{|\Delta_g|}\right)^{1/2}(\log n)^{2^{1+\alpha}},$$

for $n \gg_\varepsilon a^{30+\varepsilon}\,b|\Delta_g|^{51/4+\varepsilon}$.

Proof. By Rankin's trick and Cauchy-Schwarz we have

$$\sum_{\substack{z\leq n/a\\ z\equiv na^{-1} \bmod b\\ r_f(z)>T}} r_f(z)r_g\left(\frac{n-az}{b}\right) \leq \Big(\sum_{\substack{z\leq n/a\\ z\equiv na^{-1} \bmod b\\ r_f(z)>T}} r_f(z)^2\Big)^{1/2}\Big(\sum_{\substack{z\leq n/a\\ z\equiv na^{-1} \bmod b}} r_g\left((n-az)/b\right)^2\Big)^{1/2}$$

$$\leq T^{-\alpha/2}\Big(\sum_{\substack{k\leq n/a\\ k\equiv na^{-1} \bmod b}} r_f(k)^{2+\alpha}\Big)^{1/2}\Big(\sum_{\substack{m\leq n/b\\ m\equiv nb^{-1} \bmod a}} r_g(m)^2\Big)^{1/2}. \quad (\star)$$

Using the inequality $r_f(n)\ll\tau(n)$ and proposition 2.6, we have

$$(\star) \ll T^{-\alpha/2}\left(\frac{4^{\omega(\Delta_g)}\log^2(a\Delta_g)}{ab|\Delta_g|}n(\log n)\right)^{1/2}\Big(\sum_{\substack{k\leq n/a\\ k\equiv na^{-1} \bmod b}} \tau(k)^{2+\alpha}\Big)^{1/2},$$

which, by standard analytic estimates is

$$\ll T^{-\alpha/2}\,\frac{n}{ab}\left(\frac{4^{\omega(\Delta_g)}\log^2(a\Delta_g)}{|\Delta_g|}\right)^{1/2}(\log n)^{2^{1+\alpha}},$$

as desired. □

Proof of Theorem 2.5. Notice that we have a natural surjection from the set

$$R_Q(n)=\{(s,t,x,y)\in\mathbb{Z}^4 \mid Q(s,t,x,y)=af(s,t)+bg(x,y)=n\}$$

to

$$S(n)=\{(z,x,y)\in\mathbb{Z}^3 \mid az+bg(x,y)=n,\ z=f(s,t)\}$$

sending $(s,t,x,y)\mapsto(f(s,t),x,y)$, accordingly

$$r_Q(n)=\sum_{\substack{(z,x,y)\in S(n)\\ r_f(z)\leq T}} r_f(z) \;+\; \sum_{\substack{(z,x,y)\in S(n)\\ r_f(z)>T}} r_f(z)$$

$$\ll_{\varepsilon,\alpha} \quad T\#S(n) \quad + \quad T^{-\alpha/2}\,\frac{n}{ab}\left(\frac{4^{\omega(\Delta_g)}\log^2(a\Delta_g)}{|\Delta_g|}\right)^{1/2}(\log n)^{2^{1+\alpha}}$$

for $n \gg_\varepsilon a^{30+\varepsilon}\, b|\Delta_g|^{51/4+\varepsilon}$. Choosing

$$T = a^{-1/\alpha}\,|\Delta_f|^{1/\alpha}\left(4^{\omega(\Delta_g)}\log^2(a\Delta_g)\right)^{1/\alpha}(\log n)^{(2^{2+\alpha})/\alpha}.$$

and using corollary (2.4) with

$$n \gg_\varepsilon N^{8+\varepsilon}|det(A)|^{2+\varepsilon} \gg a^{20+\varepsilon}b^{20+\varepsilon}|\Delta_f|^{10+\varepsilon}|\Delta_g|^{10+\varepsilon},$$

we obtain

$$\#S(n) \gg_{\varepsilon,\alpha} T^{-1}\frac{n}{ab(\Delta_f\Delta_g)^{1/2}} \gg_{\varepsilon,\alpha} \left(\frac{C_{f,g,a}^{-1/\alpha}}{a\,b\,|\Delta_f|^{1+1/\alpha}|\Delta_g|}\right) n(\log n)^{-(2^{2+\alpha})/\alpha}$$

for $n \gg_\varepsilon a^{30+\varepsilon}b^{20+\varepsilon}|\Delta_f|^{10+\varepsilon}|\Delta_g|^{51/4+\varepsilon}$ and any fixed $\alpha > 0$ with $C_{f,g,\alpha} = 4^{\omega(\Delta_g)}\log^2(a\Delta_g)$. The choice $\alpha = 1/\ln 2$ is to minimize the power saving logarithm, which justifies $\gamma \approx 7.532..$.

3.3 Number of representations of primes by Quaternary form. Given a primitive positive definite integral quaternary quadratic form Q of $\det Q = q^2$ where q is prime, in this subsection we seek estimates for the number of representations of primes by this form between two given close bounds. Namely, for x, y with $y < x$, we're interested in obtaining lower bounds for

$$\sum_{\substack{y\le p\le x\\ p \text{ prime}}} r_Q(p). \tag{8}$$

The main idea is to utilize corollary 2.4 since Q is quaternary definite, as we have estimates in this case for $r_Q(n)$ when $n \gg_\varepsilon N^{8+\varepsilon}|\det(A)|^{2+\varepsilon}$. However for the estimate to make sense, we need to show that the singular series doesn't vanish, i.e that the congruence in 3 is solvable. It suffices, by the Chinese remainder theorem, to ensure solvability mod p^k for $k \in \mathbb{Z}^+$ for primes $p|2\det(A)$. In fact, we shall prove a stronger result, but first we introduce the following definition.

Definition 2.8. *A quadratic form is said to locally represent an integer n, if $Q(v) \equiv n \mod p^k$ is solvable for all $k > 0$.*

Lemma 2.9. *Let $Q = x^T Ax$ be a primitive, integral positive-definite quateranry quadratic form, and $\det(A) = q^2$ for some prime q. Then the level of Q is q, and Q locally represents all positive integers.*

Proof. First, we prove that the level of Q is q using a classification of Shimura's (see [Shi06], Section 6.3). Let $\mathfrak{S}_n$ denote the set of matrices $T \in \mathrm{GL}_n(\mathbb{Q})$ that are symmetric, have integer diagonal entries, and whose off-diagonal entries are either integral or half-integral. We call an element of T *reduced* if whenever P is an invertible matrix with integer entries and $P^{-1}T(P^{-1})^T \in \mathfrak{S}_n$, then $\det(P) = \pm 1$. We first claim that the matrix A in the statement of the lemma above is reduced.

If it were not reduced, then there would be a matrix P with $|\det(P)| > 1$ so that $P^{-1}(A)(P^{-1})^T \in \mathfrak{S}_n$. The determinant of this matrix must be d for some integer d, but also, it equals $\det(P^{-1})^2 \cdot q^2$. It follows that $\det(P) = q$. However, this would imply that $P^{-1}A(P^{-1})^T$ is the Gram matrix for a positive-definite Quaternary form with determinant 1. Shimura says at the beginning of [Shi06] that it is classically known that a positive-definite form with determinant 1 must have the number of variables $\equiv 0 \pmod 8$. (For a proof of this, see Corollary 2 on page 53 of [Ser73] or Theorem 6.4 of [Shi06].) Since $4 \not\equiv 0 \pmod 8$, this is a contradiction and so the matrix A must be reduced.
In Theorem 6.4 of [Shi06], a classification is given of reduced forms that depends on four parameters: σ (the signature of Q), δ (either 1 or the discriminant of the ring of integers of a quadratic field), e_0 and e_1. The integer e_0 must divide δ, e_1 must be prime to δ, and we must have that the determinant is $|\delta|e_1^2$ and the level of the form is $|\delta|e_1$. We must have $\delta = 1$ and $e_1^2 = q^2$. Therefore, the level of the form is $|\delta|e_1 = 1 \cdot q = q$.

Now we turn to showing that Q locally represents every positive integer. Criteria to establish this are well-known and quite old in the quadratic forms literature. Assume first that $q > 2$.
Lemma 1 of Gordan Pall's paper [Pall46] is sufficient for our purposes. It shows that the form Q stated above represents every element of $\mathbb{Z}_r$ for any odd prime $r \neq q$. Moreover, since the level of the form is q, its Jordan decomposition has the form $\langle \alpha_1, \alpha_2, q\alpha_3, q\alpha_4 \rangle$ where $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{Z}_q^\times$. The lemma of [Pall46] now implies that Q represents every element of $\mathbb{Z}_q$. For $r = 2$, calculations in the proof of Proposition 15 of [Rouse14] show that since the determinant of Q is $\equiv 1 \pmod 8$, Q is equivalent over $\mathbb{Z}_2$ to $xy + zw$, which represents every element of $\mathbb{Z}_2$.

For $q = 2$, there is a unique Quaternary form of determinant q^2, namely $Q(x, y, z, w) = x^2 + y^2 + z^2 + w^2 + xw + yw + zw$. Lemma 1 from [Pall46]'s paper easily implies that Q represents every element of $\mathbb{Z}_r$ for $r > 2$, and it is simple to check that Q represents the integers $\{1, 3, 5, 7, 2, 6, 10, 14\}$. It follows that Q represents the 8 squareclass in $\mathbb{Z}_2$ and so Q locally represents all positive integers. □

Now that $\mathfrak{S}(n) > 0$ with respect to the form Q, we turn to the problem of finding lower bounds to (8).

Proposition 2.10. *Let $Q = x^t Ax$ be a primitive, positive definite, integral quaternary quadratic form with determinant $\det A = q^2$ for some prime q. Suppose that x and y are sufficiently large, satisfying $x \gg y \gg_\varepsilon q^{12+\varepsilon}$ for arbitrary $\varepsilon > 0$. Then, the sum of representation counts over prime numbers in the interval $[y, x]$ satisfies the lower bound*

$$P_Q(x, y) \coloneqq \sum_{\substack{y \leq p \leq x \\ p \text{ prime}}} r_Q(p) \gg_\varepsilon \frac{x^2}{q \log x \log\log x}.$$

Proof. The fact that the congruence (3) is solvable by lemma 2.9 and $y \gg_\varepsilon p^{12+\varepsilon}$ corollary 2.4 imply

$$P_Q(x, y) \coloneqq \sum_{\substack{y \leq p \leq x \\ p \text{ prime}}} r_Q(p) \gg_\varepsilon \sum_{\substack{y \leq p \leq x \\ p \text{ prime}}} \frac{p}{q \log\log p} \gg_\varepsilon \frac{1}{q \log\log x} \sum_{\substack{y \leq p \leq x \\ p \text{ prime}}} p.$$

A standard application of Euler's summation and bounds for $\pi(x)$ from [Dus10] gives

$$\sum_{\substack{y \leq p \leq x \\ p \text{ prime}}} p = \int_y^x u \, d\pi(u) = x\pi(x) - y\pi(y) - \int_y^x \pi(u) du \gg \frac{x^2}{\log x};$$

the result follows. □

3. Prime Norms and Quadratic Forms Sampling

In this section, we study the following problem: given an ideal I, find another ideal J equivalent to I whose norm is a prime. Alternatively, this is equivalent to finding vectors $v \in \mathbb{Z}^4$ such that the corresponding quaternary form $q_I(v)$ is prime.

3.1. Sampling primes. Let Q be a primitive, integral, positive definite quaternary quadratic form. In this section we discuss the problem of sampling vectors from $\{x \in \mathbb{Z}^4 \mid Q(x) \leq \rho\}$ such that $Q(x)$ is a prime. To do that we need some lemmas.

Definition: *The i^{th} successive minimum λ_i of a quadratic form Q of dimension n is defined as the smallest value $r > 0$ such that there exist i linearly independent lattice vectors $v_1, v_2, ...v_i$ with $Q(v_j) \leq \lambda_i$ for all $j = 1, 2, ...i$.*

Now we develop an algorithm which samples integral vectors within a 4-ellipsoid. For practical reasons, we will specialize our generalization of Lemma 3.3 from [Wes21] using a particular bound so that our ellipsoid contains enough lattice points.

Lemma 3.1. *Let Q be a primitive, positive definite, integral Quaternary quadratic form, and let $\rho \geq 4\mu^2(Q) > 0$ where $\mu(Q)$ is the covering radius of Q. There is an algorithm that samples uniformly random elements from*

$$\{x \in \mathbb{Z}^4 \mid Q(x) \leq \rho\}$$

in polynomial time in $\log \rho$*, and the length of Q.*

Proof. The approach is analogous to the one in [Wes21] we'll just use the restriction on ρ and the result will follow for four variables.
Let $B_Q(r) = \{v \in \mathbb{R}^4 | Q(v) \leq r^2\}$ be the ellipsoid around the origin define by the quadratic form. Let $r = p^{1/2}$, our goal is to sample vectors from $B_Q(r) \cap \mathbb{Z}^4$. We first compute a Minkowski-reduced basis (b_1, b_2, b_3, b_4) of Q with $Q(b_1) \leq Q(b_2) \leq Q(b_3) \leq Q(b_4)$. Now we know that $Q(b_4) \leq \lambda_4 \leq 2\mu \leq r$ where λ_4 is the 4^{th} *successive minimum* (in fact $Q(b_4) = \lambda_4$), i.e we have enough lattice points within the ellipsoid. Let $V = \{v \in \mathbb{R}^4 | Q(v) = min_{\lambda \in \mathbb{Z}^4} f(v + \lambda)\}$ be the Voronoi cell aorund the origin, i.e all points in V are closer to the origin than any other lattice point. For any $v \in \mathbb{R}^4$, we define the closest lattice vector $\lambda(v) \in \mathbb{Z}^4$ to satisfy $v \in V + \lambda(v)$. Computing the closest lattice vector (CLV) is often expensive as it's complexity is exponential in the dimension but since we're working in dimension 4, we can compute the CLV efficiently and its unique for almost all v. We sample according to algorithm 1.

To show there is a good probability that $\lambda(v)$ will land in $B_Q(r)$, we use the fact that $r - \mu > 0$, this

Algorithm 1 QuadSampling(Q, ρ)

Input: Positive definite quaternary quadratic form Q and a bound $\rho \geq 4\mu^2(Q)$
Output: uniform $x \in \mathbb{Z}^4$ such that $Q(x) \leq \rho$

1: **while** $\lambda \notin B_Q(r)$ **do**
2: Sample a uniform $v \in B_Q(r + \mu)$;
3: Solve the closest vector problem $\lambda(v)$;
4: **end while**
5: **return** $\lambda(v)$.

implies

$$\Pr[\lambda(v) \in B_Q(r)] \geq \Pr[v \in B_Q(r - \mu)] = \frac{\mathrm{Vol}(B_Q(r - \mu))}{\mathrm{Vol}(B_Q(r + \mu))} = \frac{(r - \mu)^4}{(r + \mu)^4} \geq \frac{(1 - 2^{-1})^4}{(1 + 2^{-1})^4} \geq 0.012,$$

i.e. the process succeeds after a constant number of trials. Finally, our output vector is uniform where the result follows analogously as in [Wes21].

□

The approach taken in [Wes21] to sample vectors $v \in \mathbb{Z}^4$ such that $Q(v)$ is a prime is by transforming our quaternary quadratic form into a binary form f of discriminant D and use the well-known result on primes and the effective Chebotarev density's theorem by Lagarias and Odlyzko [LO77] (assuming GRH) which states that

$$\pi_f(\rho) = \frac{\delta \rho}{h(D) \log \rho} + O(\rho^{1/2} \log(|D|\rho)),$$

where $\pi_f(\rho) = \#\{v \in \mathbb{Z}^2 | f(v) \leq \rho \text{ is a prime}\}$, $h(D)$ is the class number and δ is either 1 or 1/2. In order to be able to sample vectors which represent primes without relying on GRH we will take a different route through the theory of quaternary forms and use proposition 2.10.

Proposition 3.2. *Let $Q = x^T Ax$ be a primitive, integral positive-definite Quaternary quadratic form, and $\det(A) = q^2$ for some prime q sufficiently large. There is an algorithm that finds an integral vector $x \in \mathbb{Z}^4$ such that $Q(x)$ is a prime number at most $O_\delta(q^{12+\delta})$ for a suitable $\delta > 0$, and runs in expected polynomial time in* $\log q$.

Proof. From Lemma 3.1, one can sample a uniformly random integral vector x such that $Q(x) \leq \rho$, where $4\mu^2(Q) \leq \rho$ and $\mu(Q) \ll \mathrm{Vol}(Q) = q$. Let $\varepsilon > 0$ be given, and suppose that $\rho \gg_\delta q^{12+\delta}$, where $\delta \geq \varepsilon > 0$ is a constant to be adjusted.

A result of Landau [Lan15] gives,

$$\#\{v \in \mathbb{Z}^4 \mid Q(v) \leq r\} = \frac{(2\pi)^2}{|\det Q|^{1/2}\Gamma(3)} r^2 + O(r^{6/5+\varepsilon}).$$

Applying Lemma 3.1, the probability that a sampled vector $v \in \mathbb{Z}^4$ with $Q(v) \ll_\varepsilon q^{12+\varepsilon}$ is very small, provided that δ is suitably chosen (q is sufficiently large). More precisely,

$$\Pr[Q(x) \ll_\varepsilon q^{12+\varepsilon}] = \frac{\#\{v \in \mathbb{Z}^4 \mid Q(v) \leq q^{12+\varepsilon}\}}{\#\{v \in \mathbb{Z}^4 \mid Q(v) \leq \rho\}} \ll \frac{q^{23+\varepsilon}}{\rho^2} \ll \frac{1}{q^{2\delta-\varepsilon}}.$$

Since the sampled vectors will, with overwhelming probability, satisfy

$$q^{12+\varepsilon} \ll_\varepsilon Q(v) \ll_\delta q^{12+\delta},$$

using Proposition 2.10, we obtain

$$\Pr[Q(v) \text{ is prime}] \gg_\delta \frac{1}{\log q \log\log q}.$$

This completes the proof.

□

3.2 Sampling Equivalent ideals of prime norm We now return to the problem of sampling a left $\mathcal{O}$-ideal J in the same class of a given left ideal I.

Algorithm 2 EquivalentPrimeIdeal$_\delta(I)$

Input: A left ideal I in a maximal order $\mathcal{O}$.

Output: An ideal J of prime norm, and an element $\alpha \in I$ such that $J = I\bar{\alpha}/\mathrm{Nrd}(I)$.

1: Find an element $\alpha \in I$ such that $q_I(\alpha)$ is a prime; {Proposition 3.2}.

2: **return** $J = I\bar{\alpha}/\mathrm{Nrd}(I)$, and α.

Theorem 3.3. *For a suitable $\delta > 0$, Algorithm 2 is correct and runs in expected polynomial time in $\log p$, and the output J has reduced norm $Nrd(J) = O_\delta(p^{12+\delta})$.*

Proof. It follows straightforwardly from Proposition 3.2, where $Q = q_I$ is a primitive integral positive definite Quaternary quadratic form of determinant p^2. □

Remark . In [KLPT14] the heuristic algorithm is expected to return an equivalent ideal of norm $\mathrm{Nrd}(J) = O(p^{1/2})$ most of the times, and $\mathrm{Nrd}(J) = O(p)$ in the worst case. In [Wes21] the output ideal J is expected to have $\mathrm{Nrd}(J) = O_\varepsilon(p^{2+\varepsilon})$ assuming the generalized Riemann hypothesis. Even though our result is certainly not as tight as the previous ones, nevertheless the size of the norm is still polynomial in the input size and will be sufficient for our purposes. We conclude this section with a proposition that modifies the last algorithm to output an ideal with certain properties.

Proposition 3.4. *Given a left ideal I, a bound $\rho \asymp_\delta p^{12+\delta}$ for $\delta > 0$, and a prime $l \neq p$, there is an algorithm which returns an ideal J equivalent to I such that $\mathrm{Nrd}(J)$ is a prime between ρ and $\rho^{1+\theta}$ for any $0 < \theta \leq 1/2$, and l is a non-quadratic residue modulo $\mathrm{Nrd}(J)$. The algorithm runs in expected polynomial time in $\log p$.*

Proof. The proof of this proposition follows from Theorem 3.3 together with [Wes21, Proposition 3.8], and proceeds along analogous lines. The condition on θ ensures that the norm of the sampled ideal J is sufficiently large. Finally, the number of primes $z = \mathrm{Nrd}(J)$ in an interval $[X, Y]$ for which ℓ is a quadratic residue modulo z is approximately $\frac{1}{2}(\pi(Y) - \pi(X))$, since such primes occur with density $1/2$ by the Chebotarev Density Theorem. Consequently, this restriction increases the running time by at most a factor of 2. □

4. Solving Diophantine Equations

5.1. The Algorithm. In this section we gather and employ our previous results to solve Diophantine equations of the form

$$af(s,t) + bg(x,y) = n \tag{9}$$

where f and g are positive definite binary quadratic forms, and $a, b, n \in \mathbb{Z}$. Put more precisely, we prove the following theorem.

Theorem 4.1. *For every $\varepsilon > 0$, there exist a constant $c_\varepsilon > 0$ and an algorithm $\mathcal{A}$ such that the following holds. Let f and g be reduced, primitive, positive definite, integral binary quadratic forms where $\mathrm{disc}(f) = \Delta_f$ is fundamental. Let a, b, n be positive integers satisfying $(a, 2bn) = (b, 2n) = 1$ and assume a divides Δ_g with $\omega(a) = 0$ or 1. Then $\mathcal{A}(f, \gamma, b, n)$ returns an integral solution (s, t, x, y) to the equation*

$$Q(s,t,x,y) = af(s,t) + bg(x,y) = n \tag{10}$$

provided that $n \gg_\varepsilon a^{30+\varepsilon} b^{20+\varepsilon} |\Delta_f|^{10+\varepsilon} |\Delta_g|^{51/4+\varepsilon}$ and $Q(s,t,x,y) \equiv n$ is solvable modulo $2^7 \det^3(Q)$. Assuming access to a factoring oracle $\mathscr{O}$, the algorithm runs in expected polynomial time in Δ_f, $2^{\omega(\Delta_g)}$, $\log(ab)$, and $\log(n)$, where $\omega(\cdot)$ denotes the number of distinct prime factors. The output is random with min-entropy $\Omega(\log n)$.

In [Wes21], Wesolowski employs genus theory of quadratic forms and prime ideals in a quadratic extension to solve a similar Diophantine equation, where $a = \det(\gamma)^2$ and $g = f^\gamma$. His approach is motivated by the need to find primes ℓ represented by f that also satisfy the generalized Hardy-Littlewood equation

$$a\ell + bf(x,y) = n.$$

It is well known that Cornacchia's algorithm finds a solution to the Diophantine equation $f(s,t) = z$ in polynomial time in $\log z$ and $\log \Delta_f$, provided that the factorization of z is known and m doesn't have too many prime factors. Since primality testing can be performed efficiently, Wesolowski requires z to be prime.
A mere access to a factoring oracle is not sufficient to solve $f(s,t) = z$ especially when z is highly composite. However, the author recently developed algorithms in another work- see [Mah25] - that find solutions to $f(s,t) = z$ where there is a tradeoff between the size of Δ_f and $2^{\omega(z)}$. Precisely put, if Δ_f is small, i.e. polynomial in $\log n$, then algorithm 1 of [Mah25] finds a solution efficiently.
By leveraging our factoring oracle $\mathscr{O}$ and the fact that $\Delta_f = \mathrm{polylog}(n)$, we circumvent these obstacles, allowing us to solve the problem using the theory of quaternary quadratic forms without relying on the Generalized Riemann Hypothesis (GRH).

Proof of theorem 4.1. Most of our work has been done in section 2, we just have to gather our tools. Assume we have access to a factoring oracle $\mathscr{O}$, $n \gg_\varepsilon a^{30+\varepsilon} b^{20+\varepsilon} |\Delta_f|^{10+\varepsilon} |\Delta_g|^{51/4+\varepsilon}$, and $Q(s,t,x,y) \equiv n \pmod{2^7 \det^3 Q}$ is solvable. Using ([Wes21], *proposition* 5.2) we can sample almost uniform (up to small error $O(n^{-1/2})$) integral solutions (z, x, y) to $az + bg(x,y) = n$ with $z > 0$ in polynomial time, and the sample space X has cardinality

$$\#X = \frac{2^{\omega(a)} \pi n}{ab\ \mathrm{Vol}(g)} + O\left(2^{\omega(a)} \left(\frac{n}{b}\right)^{1/2} + a 2^{\omega(a)} \mathrm{Vol}(g)\right) \ll \frac{2^{\omega(a)} \pi n}{ab\ \mathrm{Vol}(g)} \tag{11}$$

Each time we sample a solution (z, x, y) we factor z using $\mathscr{O}$; we use [Mah25, algorithm 1] to efficiently find a solution or decide that $z = f(s,t)$ is not solvable. In the latter case, we sample another triplet solution. We would be done if we can show that a large proportion of the sampled solutions (z, x, y) satisfy $z = f(s,t)$. Thanks to theorem 2.5, we have

$$\#\{(z,x,y) \in \mathbb{Z}^3 \mid az + bg(x,y) = n,\ z = f(s,t)\} \gg_\varepsilon \left(\frac{C_{f,g,a}^{-1/\alpha}}{a\, b\, |\Delta_f|^{1+1/\alpha} |\Delta_g|}\right) n (\log n)^{-\gamma},$$

and $C_{f,g,a} = 4^{\omega(\Delta_g)} \log^2(a\Delta_g)$. Combining this with (11), we obtain

$$\Pr[(z,x,y) \in X \ , z = f(s,t)] \geq c_\varepsilon \ \frac{C_{f,g,a}^{-1}}{2^{\omega(a)}|\Delta_f|^{1+1/\alpha} \log^\gamma n},$$

for a constant c_ε. Hence, we expect the sampling process to succeed after

$$O\left(2^{\omega(a)} C_{f,g,a} |\Delta_f|^{1.7} (\log n)^{7.54}\right)$$

many trials. Since sampled solutions in X behave almost uniformly, we expect the min-entropy $H_{min}(X) = -\log(P_{max}) = \Omega(\log n)$; this concludes the proof. □

Remark. The algorithm's total running time is polynomial provided that $\omega(m) \ll \log\log n$ where $m = \Delta_f$, Δ_g, i.e. those integers should not be highly composite. This is satisfied readily in the next corollary.

5.2. Finding elements in a special maximal order The following immediate corollary allows us to sample elements from the special maximal order $\mathcal{O}_0$ in $B_{p,\infty}$ as defined in the lemma 1.3.

Corollary 4.2. *There exists an algorithm $\mathcal{A}$ such that the following holds. For any integer n and prime p with $(n,p) = 1$ and $n \gg_\varepsilon p^{20+\varepsilon}$ for every $\varepsilon > 0$, $\mathcal{A}$ finds an element $\alpha \in \mathcal{O}_0$ with $\mathrm{Nrd}(\alpha) = n$. Using the oracle $\mathscr{O}$, the algorithm runs in expected polynomial time in terms of $\log(p)$ and $\log(n)$.*

Proof. From lemma 1.3 we know that elements $\alpha = s + t\omega + xj + y\omega j \in R + Rj \subset \mathcal{O}_0$ satisfy

$$\mathrm{Nrd}(\alpha) = f(x,y) + pf(s,t),$$

where $\mathrm{disc}(f) = O(\log^2 p)$ (we assume the least quadratic non-residue has been found and has size polynomial in $\log p$). Notice that $\omega(\Delta_f) = 1$ for all cases in 1.1, hence the algorithm is efficient. However, in order to apply Theorem 4.1 we need $f(x,y) + pf(s,t) \equiv n \bmod 2^7 det^3(A)$ to be solvable. It is enough to show the solvability modulo 2^3, p and $\mathrm{disc}(f)$, then use Hensel's lemma and Chinese remainder theorem, which are both straightforward but require some computations in the case mod 2^3. The condition on ε comes from the fact that we need $n \gg_\varepsilon a^{30+\varepsilon} b^{20+\varepsilon} |\Delta_f|^{10+\varepsilon} |\Delta_g|^{51/4+\varepsilon}$. □

5. Solving the QuaternionPath Problem

For the remaining sections we consider our special maximal order $\mathcal{O}_0$ from lemma 1.3 in the quaternion algebra $B_{p,\infty}$. The main result here is theorem 5.1 that solves the QuaternionPath. Since there exist efficient algorithms to compute connecting ideals between maximal orders, see for instance [KV10], our problem is reduced to the following: given a maximal order $\mathcal{O}$ in $B_{p,\infty}$, a left $\mathcal{O}$-ideal I, and an integer N, find an equivalent ideal J of $\mathrm{Nrd}(J) = N$. In [[KLPT14], section 4.6], its been argued that the general case reduces to the special maximal order $\mathcal{O}_0$ and hence we focus on this one.

6.2. Solving the quaternion path problem.

Algorithm 3 $\textsc{QuaternionPath}_\varepsilon(I, n_1, n_2, \ell)$

Input: A left ideal I in a special maximal order $\mathcal{O}_0$, positive integers n_1, n_2 and a prime ℓ.
Output: An equivalent ideal J of norm $n_1 n_2$ or $n_1 n_2 l$.

Define R, ω, and f as in lemma 1.3;
while β has not been found **do**
  Random walk in the Brandt graph to obtain $I' \subset I$ of norm n_1; {Theorem 1.5}
  Find an element $\rho \in I'$ such that $I'' = I'\rho/\mathrm{Nrd}(I)$ equivalent to I' of prime norm $N \in [c_\varepsilon p^{12+\varepsilon}, c_\varepsilon p^{12.5+\varepsilon}]$ and ℓ is a quadratic non residue modulo N; {Proposition 3.4, *with* $\theta = \frac{1}{2}$}.
  Find an element $\gamma \in \mathcal{O}_0$ such that $\mathrm{Nrd}(\gamma) = N$; {Corollary 4.2}
  Find an element $\beta \in R$ such that $I'' = \mathcal{O}_0 N + \mathcal{O}_0 \beta j$ if it exists;
end while
Find a matrix $\Gamma \in M_{2\times 2}(\mathbb{Z})$ such that $\mathbb{Z}\beta + RN = \{x + y\omega \mid (x, y) \in \mid \Gamma\mathbb{Z}^2\}$;
Find an integral solution (s, t, x, y) of $N^2 f(s, t) + p f^\Gamma(x, y) = n_2 l^e$ for $e \in \{0, 1\}$; {Theorem 4.1}
Set $(x', y') = \Gamma(x, y)$, and $\alpha = (s + t\omega)N + (x' + y'\omega)j$;
Set $\delta = \rho\gamma\alpha/N \in I' \subset I$;
return $J = I\bar{\delta}/\mathrm{Nrd}(I)$

Theorem 5.1. *For every $\varepsilon > 0$ there exists a constant c_ε such that for integers $n_1, n_2 \geq c_\varepsilon p^{1089+\varepsilon}$ and a prime ℓ with $n_2 \ell^e \not\equiv 2, 4 \pmod 8$ for $e \in \{0, 1\}$, Algorithm 3 is correct and runs in expected polynomial time in ℓ, $\log p$, $\log n_i$, and $\log \mathrm{Nrd}(I)$, assuming access to a factoring oracle $\mathscr{O}$.*

Proof. The algorithm has the same approach as in ([Wes21], *Theorem* 6.3), we will just comment on the steps where our variants are implemented.

Step 3. For $\log n_1 \geq \log(c_\varepsilon p^{1089+\varepsilon})$, the output ideal will lie in any class with probability between $1/2N_p$ and $3/2N_p$, i.e. its almost uniform (In fact, from lemma 1 of [GPS20], $\log n_1 \geq (1 + \varepsilon)\log p$ suffices).

Step 4, 5 To utilize proposition 3.4, the norm of the desired ideal should be at least $\rho = c_\varepsilon p^{12+\varepsilon}$, but to have a concrete bounds on $\mathrm{Nrd}(I')$ we take $\theta = 1/2$. Given $N \geq c_\varepsilon p^{12+\varepsilon}$, we can sample an equivalent ideal with the required properties using corollary 3.4.

Step 6 As dealt with in ([KLPT14], *section* 4.3), this step can be solved using linear algebra over $\mathbb{Z}/N\mathbb{Z}$. Additionally, its been proven by [Wes21], that the randomization of the ideal in step 3 and the large entropy of γ in step 5, ensure that this step succeeds since I''/NO_0 and $O_0\gamma/NO_0$ are expected to different from the, at most, two fixed points under the action of $(R/NR)^\times$.

Steps 8,9 Given β, we find the matrix $\Gamma \in M_{2\times 2}(\mathbb{Z})$ with $\det\Gamma = N$ which can be done by computing square roots modulo N, as outlined in ([KLPT14], *section* 4.4). Theorem 4.1 works efficiently in Step 9 since $n \gg_\varepsilon p^{1089+\varepsilon} \gg_\varepsilon a^{30+\varepsilon} b^{20+\varepsilon} |\Delta_f|^{10+\varepsilon} |\Delta_g|^{51/4+\varepsilon}$ and $\omega(\Delta_f),\ \omega(\Delta_g) \ll 1$.

What remains is to show that

$$N^2 f(s, t) + p f^\Gamma(x, y) = n_2 \ell^e$$

is solvable modulo $2^7 \det(A)^3$, where A is the Gram matrix of the quaternary form $N^2 f(s, t) + p f^\Gamma(x, y)$. Since Q is quaternary, it suffices to prove solvability modulo 2^3, N, p, and Δ_f, and then combine the resulting congruences using Hensel's lemma and the Chinese remainder theorem.

For the modulus 2^3, this can be verified by a direct (albeit tedious) computation, using that f is a principal form; the problem reduces to representing 3 and 5 modulo 8. Finally, since ℓ is a quadratic nonresidue modulo N, we choose the parity of the exponent e according to whether $p f^\Gamma(x, y)$ is a quadratic residue modulo N.

Reducing the equation modulo p and using the fact that N is invertible mod p, gives $f(s, t) \equiv N^{-2} n_2 l^e$, where $f(s, t)$ is a principal form of $\Delta_f = -4$, -8, or $-q_p$ depending whether $p \equiv 3 \mod 4$, $p \equiv 5 \mod 8$, or $p \equiv 1 \mod 8$ according to lemma 1.1. Using the fact that in $\mathbb{Z}/p\mathbb{Z}$ every

element is a sum of two squares, and a sum of a quadratic residue and non residue we obtain a solution *modulo p*.
The existence of a solution modulo Δ_f has been established in detail in the analogous result of [Wes21], and we therefore do not repeat the argument here. □

On the polynomial overhead in Wesolowski's reduction. Since this material is necessarily notation-heavy and involves chasing constants through multiple intermediate results, we separate it from the main narrative: readers only interested in the high-level statement may skip the bookkeeping entirely, while readers wishing to verify the claimed overhead can check this remark. Throughout this discussion we assume familiarity with the setup and notation of [Wes21]. The purpose of this remark is purely quantitative: we explain why the polylogarithmic exponent appearing in the running time of Wesolowski's reduction [Wes21] is forced to be very large.

More precisely, the exponent c in the $\Omega(\log^c p)$ running time in [Wes21, Theorem 6.3] ultimately depends on the small exponent implicit in the analytic input used in the reduction. Concretely, Theorem 5.1 of [Wes21] (which is invoked via Theorem 6.3 to solve QuaternionPath) inherits its exponent from the distribution results in Corollary 4.2 / Theorem 4.3 of [Wes21]. These in turn rely on [ABL20, Theorem 2.1] (with minor modifications in [Wes21]).

Let $\delta' > 0$ be the constant coming from theorem 2.1 of [ABL20] upon assuming GRH and suppose we wish to work uniformly in the range

$$c, d \leq x^{\delta'},$$

with a power-saving error of size $\ll x^{1-\delta'}$. In the proof of [ABL20, Thm. 2.1] an auxiliary parameter R is introduced and the main quantity is decomposed as $S_1 + S_2$. The tail term satisfies [ABL20, (4.2)]

$$S_2 \ll c\, x\, (\log x)^{O(1)}\, R^{-1/11}.$$

Writing $R = x^{\delta}$ and using $c \leq x^{\delta'}$ gives

$$S_2 \ll x^{1+\delta'-\delta/11+o(1)}.$$

Hence $S_2 \ll x^{1-\delta'}$ is ensured as soon as $\delta \geq 22\delta'$; in particular we may choose $R := x^{22\delta'}$.

In addition, the proof explicitly reduces the problem to proving the key estimate in the enlarged modulus range

$$Q \leq x^{1/2+4\delta} \qquad \text{([ABL20, p. 17])}.$$

At this point a dispersion-type input is invoked (their citation of [BFI86, Lemma 2]), which is available for Q up to the classical level $Q \leq x^{4/7-\varepsilon}$. Equating $x^{1/2+4\delta} \leq x^{4/7-\varepsilon}$ shows that it suffices to require

$$4\delta < \frac{1}{14}, \qquad \text{hence we may take any fixed } \ \delta < \frac{1}{56}.$$

Combining this with the constraint $\delta \geq 22\delta'$ coming from S_2, we obtain the explicit admissible choice

$$\delta' < \frac{1}{56 \cdot 22} = \frac{1}{1232}.$$

Therefore, under GRH one may take the exponent in the range $c, d \leq x^{\delta'}$ (and the corresponding power saving $\ll x^{1-\delta'}$) to be any fixed constant $\delta' < 1/1232$.
Finally, in the proof of [Wes21, Theorem 4.1], Wesolowski invokes [ABL20, Theorem 2.1] and subsequently reduces the resulting exponent further (by additional factors of 5 and 2; see [Wes21, p. 12]). With our explicit choice this yields

$$\delta' < \frac{1}{12320}.$$

Recalling that the exponent c discussed above (in Wesolowski's framework) is at least the reciprocal of this parameter, we obtain

$$c \ > \ \frac{1}{\delta'} \ > \ 12320,$$

which explains the claimed degree-12320 polynomial overhead.
We stress that this lower bound is very conservative. In practice, the exponent δ' implicit in [ABL20, Theorem 2.1] is expected to be significantly smaller, which further increase the resulting degree c. □

The following corollaries give solutions to the two important variants of the QuaternionPath problems, namely ℓ-QuaternionPath and BS-QuaternionPath.

Corollary 5.2. *For every $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ and an algorithm $\mathcal{A}$ which on input a left $\mathcal{O}$-ideal I, returns an equivalent ideal J whose norm is either a prime power ℓ^e or $(C_\varepsilon \log p)$-powersmooth. Given a factoring oracle $\mathscr{O}$, the algorithm runs in expected polynomial time in $\log p$ and $\log \mathrm{Nrd}(I)$.*

Proof. Notice first that the maximal order is not necessarily the special extremal order in 1.1, but as stated at the beginning of the section, theorem 9 of [KLPT14] generalizes the special case $\mathcal{O}_0$ to arbitrary maximal orders in $B_{p,\infty}$. For the ℓ-QuaternionPath, we can take n_1, n_2 large enough powers of ℓ and the result follows.

For BS-QuaternionPath, as outlined in [Wes21], its enough to find two powersmooth integers n_1 and n_2 to apply Theorem 5.1. Let C_ε be a constant to be adjusted, and let

$$n_1 = n_2 = \prod_{\substack{\ell \text{ prime} \\ \log p < \ell < C_\varepsilon \log p}} \ell$$

Using Theorem 6.9 from [Dus10], which states that

$$\pi(x) \geq \frac{x}{\ln x}\left(1 + \frac{1}{\ln x}\right) \quad \forall x > 599, \tag{12}$$

we obtain that

$$\log(n_i) = \sum_{\substack{\ell \text{ prime} \\ \log p < \ell < C_\varepsilon \log p}} \log(\ell) \geq (\log\log p)\log p\left(\frac{C_\varepsilon}{\log(C_\varepsilon \log p)} - \frac{1}{\log\log p}\right) \geq (C_\varepsilon/2 - 1)\log p,$$

by adjusting C_ε so that $C_\varepsilon/2 - 1 > \max(599, 350 + \log c_\varepsilon,)$ ensures that $n_i > c_\varepsilon p^{350+\varepsilon}$ where c_ε is the constant from theorem 5.1; this finishes the proof. □

We finish this section with an algorithm which on input a left $\mathcal{O}_0$-ideal of prime power norm, outputs the corresponding isogeny φ_I.

Algorithm 4 Ideal to prime power isogeny

Input: A left $\mathcal{O}_0$-ideal of norm ℓ^e with ℓ prime, $\ell \neq p$ and $\ell \nmid I$.
Output The corresponding isogeny φ_I of degree ℓ^e.

1: Find the constant C_ε from Theorem 5.2;
2: **for** $i = 1, 2, ...e$ **do**
3: Set $I_i = I + \mathcal{O}_0\ell^i$;
4: Find J_i equivalent to I_i of $(C_\varepsilon \log p)$-powersmooth norm; {Theorem 5.2}
5: Translate J_i into an isogeny ψ_i; {[GPS20], Lemma 5}
6: Set E_i to be the target curve of ψ_i;
7: Compute the ℓ-isogeny φ_i from E_{i-1} to E_i using Velu's formulae; [Vel71]
8: **end for**
9: **return** $\varphi_e \circ \cdots \varphi_1$.

Lemma 5.3. *Algorithm 4 is correct an runs in expected polynomial time in $\log p$, ℓ and e, assuming access to a factoring oracle $\mathscr{O}$.*

Proof. This version appeared in ([Wes21]-Algorithm 4), using Theorem 5.2 instead makes the result unconditional. □

6. Equivalence of MaxOrder and ℓ-IsogenyPath

In a recent paper by Le Merdy and Wesolowski [HLMW25], they demonstrate the unconditional equivalence of IsogenyPath, MaxOrder, and EndRing without explicitly relying on the ℓ-IsogenyPath. It is immediate that ℓ-IsogenyPath implies IsogenyPath, which, according to their paper, further implies MaxOrder. However, the reverse direction leading to ℓ-IsogenyPath remains conditional on the Generalized Riemann Hypothesis (GRH). We now begin by proving this equivalence.

7.1. **Reducing ℓ-IsogenyPath to MaxOrder.** We now show the first direction in our reduction, namely if we have two supersingular elliptic curves E_1 and E_2 with $End(E_1) \cong \mathcal{O}_1$ and $End(E_2) \cong \mathcal{O}_2$ then we can find an isogeny path from E_1 to E_2 in the ℓ-isogeny graph.

Theorem 6.1. *Given a factoring oracle $\mathscr{O}$, algorithm 5 is correct and runs in expected polynomial time in $\log p$, ℓ and the length of outputs $\mathcal{O}_1$, $\mathcal{O}_2$ of $\mathcal{A}_{MaxOrder}$.*

Proof. The idea is to make use of our special maximal order $\mathcal{O}_0$, by connecting it to each of $\mathcal{O}_1$ and $\mathcal{O}_2$ and then use the QuaternionPath algorithm. This reduction is similar to Algorithm 7 in [EHL+18]. However instead of reducing to EndRing as in [EHL+18] we reduce to MaxOrder, and we use our unconditional result instead of [KLPT14]. □

Algorithm 5 Reducing ℓ-IsogenyPath to MaxOrder.

Input: Two supersingular elliptic curves E_1, E_2 over $\mathbb{F}_{p^2}$. We assume we have access to an algorithm $\mathcal{A}_{MaxOrder}$ which solves the MaxOrder problem and outputs $\mathcal{O}_1 \cong End(E_1)$, $\mathcal{O}_2 \cong End(E_2)$ in $B_{p,\infty}$.
Output An ℓ-isogeny $\varphi : E_1 \to E_2$.

Find $\mathcal{O}_1$ and $\mathcal{O}_2$ in $B_{p,\infty}$ using $\mathcal{A}_{\text{MaxOrder}}$;
Find C_ε from corollary 5.2 for $\varepsilon > 0$;
Set $e = \lceil \frac{1089 \log p + \log(C_\varepsilon)}{\log \ell} \rceil$;
Find the special order and curve $(\mathcal{O}_0, E_0)$ from lemmas 1.3 and [EHL+18, Proposition 3];
for $i = 1, 2$ **do**
    Compute ideal $I_i = I(\mathcal{O}_0, \mathcal{O}_i)$ connecting $\mathcal{O}_0$ to $\mathcal{O}_i$; {[KV10]}
    Compute J_i equivalent to I_i using $\text{QuaternionPath}_\varepsilon(I_i, \ell^e, \ell^e, \ell)$; {Corollary 5.2}
    Translate J_i into the corresponding isogeny φ_i; {Algorithm 4}
end for
return $\varphi = \varphi_2 \circ \hat{\varphi}_1$.

Algorithm 6 Reducing MaxOrder to ℓ-IsogenyPath

Input: A supersingular elliptic curve E defined over $\mathbb{F}_{p^2}$. We assume there is an algorithm $\mathcal{A}_{\ell\text{-IsogenyPath}}$ which solves the ℓ-IsogenyPath problem.
Output: A basis for $\mathcal{O} \cong \text{End}(E)$ in $B_{p,\infty}$.

Compute $(\mathcal{O}_0, E_0)$ the special maximal order and curve from lemmas 1.3 and [EHL+18, Proposition 3];
Compute C_ε from Theorem 5.2 for $\varepsilon > 0$;
Find an isogeny $\varphi = \varphi_e \circ \cdots \circ \varphi_1 : E_0 \to E$ with $\deg(\varphi_i) = \ell$ using $\mathcal{A}_{l\text{-IsogenyPath}}(E_0, E)$
Set $\psi_0 = id_{E_0}$
for $i = 1, 2, \ldots, e$ **do**
    Compute the ideal I_i corresponding to $\varphi_i \circ \psi_i$; {[GPS20], Lemma 5}
    Compute J_i of $(C_\varepsilon \log p)$-powersmooth norm in the same class as I_i; {Corollary 5.2}
    Translate J_i into it's corresponing isogeny ψ_i; {[Wes21], Lemma 7.1}
end for
Compute $O_R(J_e)$; {[Ron92], Theorem 3.2}
Set $\mathcal{O} = O_R(J_e)$
return a basis for $\mathcal{O}$.

7.2. **Reducing MaxOrder to IsogenyPath.** These reduction seem to be folklore, however for the sake of completeness, we will give the full algorithm. Given a supersingular elliptic curve E, we prove that an isogeny from E_0 (with known endomorphism ring) to E allows one to retrieve the endomorphism ring of E.

Theorem 6.2. *Given $\mathscr{O}$, algorithm 6 is correct and runs in expected polynomial time in $\log p$, and output size of the $\mathcal{A}_{l\text{-IsogenyPath}}$, in addition to the complexity time to perform $\mathcal{A}_{\ell\text{-IsogenyPath}}$.*

Proof. A version of this algorithm has first appeared in [[DMPS19]-6.2], and then [Wes21]. Writing $\varphi_i : E_{i-1} \to E_i$ with $E_e = E$, we see that each step of the loop J_i is equivalent to the ideal induced by the isogeny $\varphi_i \circ \cdots \circ \varphi_0$. Since the running time of the algorithm in ([Wes21], Lemma 7.1) is polynomial in $\max(l_i^{e_i})$, i.e. polynomial in p, we resort to Theorem 5.2 instead of 5.1 in constructing our equivalent ideal J_i, and this proves the correctness of the algorithm. The running time follows from our previous results in each step of our algorithm. In particular, if we're given two arbitrary supersingular elliptic curves E_1 and E_2, then one can recover both maximal orders $\mathcal{O}_1 \cong End(E_1)$ and $\mathcal{O}_2 \cong End(E_2)$. This can be done by constructing an ℓ-isogeny path from our special curve E_0 to E_1 and E_2, and applying Algorithm 5. □

7. Equivalence Of MaxOrder And Endomorphism Ring

In this section we prove that the Maximal Order and Endomorphism ring problems are equivalent. We start by proving the easier direction.

8.1. **Reducing EndRing to MaxOrder.** Given a supersingular elliptic curve E, if we have an algorithm which solves MaxOrder, we show to recover the endomorphism ring of E.

Theorem 7.1. *Given $\mathscr{O}$, algorithm 7 is correct and runs in expected polynomial time in $\log p$ and the length of output of $\mathcal{A}_{\text{MaxOrder}}$.*

Proof. This version of the algorithm appeared in [[EHL$^+$18]-Algorithm 4], we just use our unconditional result instead of [KLPT14]. The endomorphisms' representation is certainly not the most natural, yet we can still evaluate points P on our elliptic curve E in polynomial time. □

Algorithm 7 Reducing EndRing to MaxOrder

Input: A supersingular elliptic curve E defined over $\mathbb{F}_{p^2}$, with $p \neq \ell$. We suppose we have an algorithm $\mathcal{A}_{\text{MaxOrder}}$ which solves the MaxOrder problem.
Output: Four endomorphisms of E that generate $End(E)$.

1: Use Algorithm $\mathcal{A}_{\text{MaxOrder}}$ to obtain a maximal order $\mathcal{O} \cong End(E)$.
2: Compute C_ε from Theorem 5.2 for $\varepsilon > 0$;
3: Compute $(\mathcal{O}_0, \text{End}(E_0))$ the special maximal order as in lemma 1.3; {[EHL$^+$18], proposition 3}
4: Find an ideal I connecting $\mathcal{O}_0$ and $\mathcal{O}$; [KV10]
5: Compute an ideal J of $(C_\varepsilon \log p)$-powersmooth norm N in the same class as I; {Theorem 5.2}
6: Translate J into its corresponding isogeny φ; {[GPS20], Lemma 5}
7: Set $1, \phi_2, \phi_3, \phi_4$ to be the generators of $\text{End}(E_0)$;
8: Set $(1, \omega_2, \omega_3, \omega_4)$ and $(1, \alpha_2, \alpha_3, \alpha_4)$ to be the generators of $\mathcal{O}$ and $\mathcal{O}_0$ respectively.
9: Find integers $(c_{ij})_{i,j=1}^4$ such that $\text{Nrd}(J)\omega_i = \sum_{j=1}^4 c_{ij}\alpha_j$
10: **return** $(N, \varphi, (c_{ij})_{i,j=1}^4)$ representing the endomorphisms $\frac{1}{N}\sum_{j=1}^4 c_{ij}\varphi\phi_j\hat{\varphi}$.

8.2. **Reducing MaxOrder to EndRing:** This is last direction in our reductions, namely that MaxOrder reduces to EndRing. The approach in [EHL$^+$18] succeeds under heuristic assumptions that some numbers appearing in the algorithm behave like random numbers and hence are easy to factor. The approach followed in [Wes21] was to force such numbers to be primes, to avoid factorisation, and make use of a sampling algorithm (essentially 3.4 with GRH) along with an explicit parameterisation of solutions of quadratic forms. Given the oracle $\mathscr{O}$, we will follow the approach of [EHL$^+$18].

Algorithm 8 Reducing MaxOrder to EndRing

Input: A supersingular elliptic curve E over $\mathbb{F}_{p^2}$. We assume there is an algorithm $\mathcal{A}_{\text{EndRing}}$ which solves the EndRing problem.

Output: A maximal order $\mathcal{O} \subset B_{p,\infty}$ such that $\mathcal{O} \cong End(E)$.

1: Use $\mathcal{A}_{\text{EndRing}}$ to output four endomorphism $1, \alpha_2, \alpha_3, \alpha_4$ generating $End(E)$;
2: Compute the Gram matrix G_0 of the sequence $(1, \alpha_2, \alpha_3, \alpha_4)$
3: Find a rational invertible linear transformation mapping $(1, \alpha_2, \alpha_3, \alpha_4)$ to some $(1, \beta_2, \beta_3, \beta_2\beta_3)$ where $1, \beta_2, \beta_3, \beta_2\beta_3$ are orthogonal basis for $B_{p,\infty}$ over $\mathbb{Q}$;
4: Factorise the numerator and denomenantor of $\text{Nrd}(\beta_2)$ and $\text{Nrd}(\beta_3)$ using $\mathscr{O}$;
5: Compute $a, b, c \in \mathbb{Q}$ such $\text{Nrd}(\iota) = q$, where $\iota = a\beta_2 + b\beta_3 + c\beta_2\beta_3$; [Sim05]
6: Find a rational invertible linear transformation mapping $(1, \beta_2, \beta_3, \beta_2\beta_3)$ to $(1, \iota, \delta, \iota\delta)$ for some $\delta \in B_{p,\infty}$, with $(1, \iota, \delta, \iota\delta)$ orthogonal basis for $B_{p,\infty}$ over $\mathbb{Q}$;
7: Factorise the numerator and denominator of $\text{Nrd}(\delta)$ using $\mathscr{O}$;
8: Compute $a, b \in \mathbb{Q}$ such that $\text{Nrd}(\delta)(a^2 + qb^2) = p$; [Mah25]
9: Find a a rational invertible linear transformation sending $(1, \iota, \delta, \iota\delta)$ to $(1, \iota, \lambda, \iota\lambda)$;
10: Invert and compose all linear transformation to write $(1, \alpha_2, \alpha_3, \alpha_4)$ in terms of $(1, \iota, \lambda, \iota\lambda)$;
11: Compute a basis for $\mathcal{O}$ in $B_{p,\infty}$;
12: **return** the basis of $\mathcal{O}$.

Theorem 7.2. *Given a factoring oracle $\mathscr{O}$, algorithm 8 is correct and runs in expected polynomial time in* $\log p$ *and the length of output of* $\mathcal{A}_{\text{EndRing}}$.

Proof. A version of this algorithm appeared in [EHL+18, algorithm 6], where some heuristics were needed in order to ensure steps 4 and 7 go through. With our factoring oracle this is trivially true.

8. Equivalence of ℓ-IsogenyPath and IsogenyPath

Most of the work has been done in the previous sections, the main result here is just the following theorem where we show the equivalence of IsogenyPath and ℓ-IsogenyPath problems. Reduction to ℓ-IsogenyPath is straightforward, so we focus on the other direction.

Theorem 8.1. *Given access to a factoring oracle $\mathscr{O}$ and an oracle $\mathcal{A}$ for* IsogenyPath*, there exists an algorithm that solves ℓ-*IsogenyPath *in polynomial time.*

Proof. Let E and E' be two instances of supersingular elliptic curves. We know ℓ-IsogenyPath$(E,\ E')$ reduces to MaxOrder in polynomial time via algorithm 5; which further reduces to EndRing by algorithm 8. In [PW24], it been shown that EndRing reduces to IsogenyPath unconditionally in polynomial time, hence given an oracle for IsogenyPath, we obtain a solution for the ℓ-IsogenyPath. □

9. Acknowledgments

The author expresses his sincere gratefulness to Dr Kirsten Eisentraeger for her assistance and moivation on starting this work. Moreover, many thanks to Dr Robert Vaughan, Valentine Blomer and Jeremy Rouse for their feedback on section 2. Lastly, the author would like to thank Benjamin Wesolowski for reporting a previous mistake in Theorem 4.1.

References

[ABL20] Edgar Assing, Valentin Blomer, and Junxian Li. *Uniform Titchmarsh divisor problems.* Preprint, arXiv:2005.13915, 2020. arXiv:2005.13915.

[ABF+25] Aardal, M. A., Basso, A., De Feo, L., Patranabis, S., & Wesolowski, B. (2025). *A complete security proof of SQIsign*. In Y. Tauman Kalai & S. F. Kamara (Eds.), *Advances in Cryptology – CRYPTO 2025* (Lecture Notes in Computer Science, Vol. 16005, pp. 190–222). Springer.

[BBC+23] Andrea Basso, Giacomo Borin, Wouter Castryck, Maria Corte-Real Santos, Riccardo Invernizzi, Antonin Leroux, Luciano Maino, Frederik Vercauteren, and Benjamin Wesolowski. *PRISM: Simple and compact identification and signatures from large prime degree isogenies*.https://eprint.iacr.org/2025/135.pdf

[BD08] T. D. Browning and R. Dietmann. *On the representation of integers by quadratic forms. Proc. London Math. Soc.* **96** (2008), no. 2, 389–416.

[BFI86] E. Bombieri, J. B. Friedlander, and H. Iwaniec, Primes in arithmetic progressions to large moduli, Acta Mathematica, 156 (1986), 203–251

[BG06] Valentin Blomer and Andrew Granville. *Estimates for representation numbers of quadratic forms. Duke Math. J.* **135** (2006), no. 2, 261–302.

[Brown07] T. D. Browning. *Density of integer solutions to diagonal quadratic forms. Monatsh. Math.* **152** (2007), 13–38.

[CD22] Wouter Castryck and Thomas Decru. *An efficient key recovery attack on SIDH (preliminary version)*. Cryptology ePrint Archive, 2022.

[CGL09] Denis X. Charles, Kristin E. Lauter, and Eyal Z. Goren. *Cryptographic hash functions from expander graphs. J. Cryptology* **22** (2009), no. 1, 93–113.

[CLM+18] Wouter Castryck, Tanja Lange, Chloé Martindale, Lorenz Panny, and Joost Renes. *CSIDH: an efficient post-quantum commutative group action*. In *Advances in Cryptology – ASIACRYPT 2018*, LNCS 11274, Springer, 2018, pp. 395–427.

[CPV20] Wouter Castryck, Lorenz Panny, and Frederik Vercauteren. *Rational isogenies from irrational endomorphisms*. In *Advances in Cryptology – EUROCRYPT 2020*, LNCS 12106, Springer, 2020, pp. 523–548.

[Coh13] Henri Cohen. *A Course in Computational Algebraic Number Theory*. Graduate Texts in Mathematics, vol. 138, Springer, 2013.

[Cor08] G. Cornacchia. *Su di un metodo per la risoluzione in numeri interi dell'equazione* $\sum_{h=0}^{n} c_h x^{n-h} y^h$. *Giorn. Mat. Battaglini* **46** (1908), 33–90.

[DFK+20b] Luca De Feo, David Kohel, Pierre Leroux, Christophe Petit, and Benjamin Wesolowski. *SQISign: Compact post-quantum signatures from quaternions and isogenies*. In *Advances in Cryptology – EUROCRYPT 2020*, Lecture Notes in Computer Science, vol. 12106, Springer, 2020, pp. 523–548.

[DMPS19] Luca De Feo, Simon Masson, Christophe Petit, and Antonio Sanso. *Verifiable delay functions from supersingular isogenies and pairings*. In *Advances in Cryptology – ASIACRYPT 2019*, LNCS 11921, Springer, 2019, pp. 248–277.

[Dub14] C. Dubey and T. Holenstein. *Sampling a uniform solution of a quadratic equation modulo a prime power. Leibniz Int. Proc. Inform.*, 2014.

[Dus10] Pierre Dusart. *Estimates of some functions over primes without R.H.* Preprint, 2010.

[EHL+18] Kirsten Eisenträger, Sean Hallgren, Kristin E. Lauter, Travis Morrison, and Christophe Petit. *Supersingular isogeny graphs and endomorphism rings: reductions and solutions*. In *EUROCRYPT 2018*.

[GPS16] Steven D. Galbraith, Christophe Petit, Barak Shani, and Yuval Ti. *On the security of supersingular isogeny cryptosystems*. In *ASIACRYPT 2016*, LNCS 10031, Springer, pp. 63–91.

[GPS17] Steven D. Galbraith, Christophe Petit, and Javier Silva. *Identification protocols and signature schemes based on supersingular isogeny problems*. In *ASIACRYPT 2017*, LNCS 10624, Springer, pp. 3–33.

[GPS20] Steven D. Galbraith, Christophe Petit, and Javier Silva. *Identification protocols and signature schemes based on supersingular isogeny problems. J. Cryptology* **33** (2020), no. 1, 130–175.

[HB96] D. R. Heath-Brown. *A new form of the circle method, and its application to quadratic forms. J. Reine Angew. Math.* **481** (1996), 149–206.

[HLMW25] Arthur Herlédan Le Merdy and Benjamin Wesolowski. *Unconditional foundations for supersingular isogeny-based cryptography.* In Benny Applebaum and Huijia (Rachel) Lin (eds.), *Theory of Cryptography (TCC 2025), Proceedings, Part III*, Lecture Notes in Computer Science, vol. 16270, Springer, 2025, pp. 266–297. doi:10.1007/978-3-032-12296-4_9.

[IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic Number Theory.* AMS Colloquium Publications, vol. 53, AMS, 2004.

[Iwa12] Henryk Iwaniec. *Topics in Classical Automorphic Forms.* AMS, 2012.

[JDF11] David Jao and Luca De Feo. *Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies.* In *PQCrypto 2011*, LNCS 7071, Springer, pp. 19–34.

[KLPT14] David Kohel, Kristin E. Lauter, Christophe Petit, and Jean-Pierre Tignol. *On the quaternion–isogeny path problem. LMS J. Comput. Math.* **17** (2014), 418–432.

[Koh96] David Kohel. *Endomorphism rings of elliptic curves over finite fields.* Ph.D. thesis, University of California, Berkeley, 1996.

[K93] Y. Kitaoka. *Arithmetic of Quadratic Forms.* Cambridge Tracts in Mathematics, vol. 106, Cambridge Univ. Press, 1993.

[KV10] Markus Kirschmer and John Voight. *Algorithmic enumeration of ideal classes for quaternion orders. SIAM J. Comput.* **39** (2010), no. 5, 1714–1747.

[Lan15] Edmund Landau. *Zur analytischen Zahlentheorie der definiten quadratischen Formen. Berliner Sitzungsberichte* **31** (1915), 458–476.

[Lin41] U. V. Linnik. *The large sieve. Dokl. Akad. Nauk SSSR* **30** (1941), 292–294.

[Lin17] S. Lindqvist. *Weak approximation results for quadratic forms in four variables.* Preprint, arXiv:2405.05592.

[LO77] J. C. Lagarias and A. M. Odlyzko. *Effective versions of the Chebotarev density theorem.* In *Algebraic Number Fields*, Academic Press, 1977, pp. 409–464.

[Mah25] Maher Mamah. *Enhanced algorithms for the representation of integers by binary quadratic forms: reduction to subset sum.* Cryptology ePrint Archive, 2025.

[Mat12] L. Matthiesen. *Linear correlations amongst numbers represented by positive definite binary quadratic forms. Acta Arith.* **154** (2012), 235–306.

[MM22] Luciano Maino and Chloé Martindale. *An attack on SIDH with arbitrary starting curve.* Cryptology ePrint Archive, 2022.

[Pall46] Gordon Pall. *The completion of a problem of Kloosterman. Amer. J. Math.* **68** (1946), 47–58.

[Piz80] Arnold Pizer. *An algorithm for computing modular forms on* $\Gamma_0(n)$. *J. Algebra* **64** (1980), no. 2, 340–390.

[PW24] Aurel Page and Benjamin Wesolowski. *The supersingular endomorphism ring and one endomorphism problems are equivalent.* In *EUROCRYPT 2024*, Springer, 2024.

[PR24] Aurel Page and Damien Robert. *Introducing Clapoti(s): evaluating the isogeny class group action in polynomial time.* Cryptology ePrint Archive, 2023.

[Rob22] Damien Robert. *Breaking SIDH in polynomial time.* Cryptology ePrint Archive, 2022.

[Ron92] Lajos Rónyai. *Algorithmic properties of maximal orders in simple algebras over* $\mathbb{Q}$. *Computational Complexity* **2** (1992), no. 3, 225–243.

[Rouse14] Jeremy Rouse. *Quadratic forms representing all odd positive integers. Amer. J. Math.* **136** (2014), no. 6, 1693–1745.

[Rouse18] Jeremy Rouse. *Integers represented by positive-definite quadratic forms and Petersson inner products. Acta Arith.* (2018).

[Ser73] Jean-Pierre Serre. *A Course in Arithmetic.* Graduate Texts in Mathematics, vol. 7, Springer, 1973.

[Sil86] Joseph H. Silverman. *The Arithmetic of Elliptic Curves.* Graduate Texts in Mathematics, vol. 106, Springer, 1986.

[Sim05] David Simon. *Quadratic equations in dimensions 4, 5 and more.* Preprint, 2005.

[Shi06] Goro Shimura. *Classification, construction, and similitudes of quadratic forms. Amer. J. Math.* **128** (2006), no. 6, 1521–1552.

[Tar29] W. Tartakowsky. *Die Gesamtheit der Zahlen, die durch eine positive quadratische Form $f(x_1, \ldots, x_s)$ darstellbar sind. Izv. Akad. Nauk SSSR* **7** (1929), 111–122, 165–195.

[Tit86] E. C. Titchmarsh. *The Theory of the Riemann Zeta-Function.* 2nd ed., edited by D. R. Heath-Brown, Oxford University Press, Oxford, 1986.

[Vel71] J. Vélu. *Isogénies entre courbes elliptiques. C. R. Acad. Sci. Paris* **273** (1971), 238–241.

[Voi21] John Voight. *Quaternion Algebras.* Graduate Texts in Mathematics, vol. 288, Springer, 2021.

[Wes21] Benjamin Wesolowski. *The supersingular isogeny path and endomorphism ring problems are equivalent.* In *FOCS 2022*, IEEE, pp. 1100–1111.

[Wes22] Benjamin Wesolowski. *Orientations and the supersingular endomorphism ring problem.* In *EUROCRYPT 2022*, LNCS 13277, Springer.

10. Appendix

Here, we we give the proof of proposition 2.6, which provides upper bounds for

$$\sum_{\substack{z\leq X\\ z\equiv k \bmod m}} r_g^2(n) \quad \text{with } (k,m)=1,$$

where $g(x,y)$ is a positive definite binary quadratic form. The main results are theorem 10.1 and corollary 10.2. We will also assume the content of the original theorem that m is a power of a single prime, i.e $\omega(m)=1$ (as the other case $\omega(m)=0$ is trivial). Also, for simplicity we will assume $g(x,y)=ax^2+bxy+cy^2$ is also Gaussian reduced. i.e. $|b|\leq a\leq c$ (since $r_g(n)=r_{\tilde{g}}$ where $\tilde{g}$ is a reduced form). The approach replicates Heath-brown δ-method [HB96] by counting number of weighted integral solutions to an indefinite quaternary quadratic form $F(x)=0$. To make this precise, notice that

$$\begin{aligned} r_g^2(n) &= \{(x,y,s,t)\in\mathbb{Z}^4 \mid g(x,y)=g(s,t)=n\} \\ &= \{(x,y,s,t)\in\mathbb{Z}^4 \mid F(x,y,s,t)=g(x,y)-g(s,t)=0,\ g(x,y)=n\}. \end{aligned}$$

Hence, we have $\sum_{z\leq X} r_g^2(n) \ll \{\mathbf{x}\in\mathbb{Z}^4 \mid F(\mathbf{x})=0,\ \max_{1\leq i\leq 4}|a_ix_i^2|\leq X\}$, where a_i are the coefficients of the non-cross terms. In particular, for a reduced form, we know the $\max_i a_i = c \ll |\Delta_g|$.

Now, accounting for the congruence condition $k \pmod m$, let $\mathcal{S}:=\mathcal{S}_{g,m,k}=\{v\in(\mathbb{Z}/m\mathbb{Z})^2 \mid g(v)\equiv k \bmod m\}$. Then, note that if $v\in\mathbb{Z}^2$ satisfies $g(v)=n\equiv k \bmod m$, then $v\in\mathcal{S}$. Hence, it suffices to estimate the size of $\{v,w\in\mathbb{Z}^2 \mid F(v,w)=0, \text{ for some } (v,w)\in\mathcal{S}\times\mathcal{S}\}$, and then sum over all representatives in $\mathcal{S}\times\mathcal{S}$. Therefore. our goal for this section is to estimate, or find upper bounds, for

$$N_F(P,m,\mathbf{k}) = \#\{\mathbf{x}\in\mathbb{Z}^4 \mid F(\mathbf{x})=0,\ \mathbf{x}\equiv\mathbf{k} \bmod m,\ \max_{1\leq i\leq 4}|a_ix_i^2|\leq P\},$$

where F is a quaternary indefinite quadratic form of non-zero square determinant Δ_F (notice $\det F = (\det g)^2$). We have decided to use P instead of X for the sake of consistency with the literature.
It is a standard fact that the size of the former set grows roughly as $\sim c_{m,F}P^2\log P$, where the $\log P$ factor comes from the $\mathbb{Q}$-lines on the quadric surface. Our primary references are the works of Lindqvist [Lin17], Browning and Dietmann [BD08], and also Browning [Brown07]. Lindqvist's work essentially follows Heath-Brown's approach, with the addition of a congruence condition.

We will try to work a little generally and treat the problem of counting integral solution to $F(\mathbf{x})=0$ with $\mathbf{x}\equiv\mathbf{k} \bmod m$ inside a box of width P as $P\to\infty$. To proceed, we count solutions using a smooth weight function $w:\mathbb{R}^4\to[0,\infty)$ with compact support.

Let F be a fixed quadratic form, and define

$$N_{F,w}(P,m,\mathbf{k}) = \sum_{\substack{\mathbf{x}\in\mathbb{Z}^4\\ \mathbf{x}\equiv\mathbf{k}\,(\mathrm{mod}\,m)\\ F(\mathbf{x})=0}} w(P^{-1}\mathbf{x}).$$

As in [HB96], we assume that $\nabla F\neq 0$ on the closure of $\mathrm{supp}(w)$. This ensures, in particular, that the support of w does not include the origin. We will work under this assumption for the rest of this section without further mention. One can think of the weight w as an approximation of the characteristic function of the region of interest. The main work is now dedicated to proving theorem 10.1.
We will start by introducing the machinery we will need. We will follow the steps of section 2 from [BD08] and that of [HB96]. Let $R\in\mathrm{SO}_n(\mathbb{R})$ be an orthogonal matrix that diagonalizes M, and fix it throughout. Then there exist $\lambda_1,\ldots,\lambda_n\in\mathbb{R}$ such that

$$R^tMR = \mathrm{Diag}(\lambda_1,\cdots,\lambda_4),$$

and $\prod_i \lambda_i = \Delta_F$. In particular, we have $Q(\mathbf{RU}) = \lambda_1 U_1^2 + \cdots + \lambda_n U_n^2$, and $\max_i |\lambda_i| \leq 4||F||$ where $||F|| = c$ is the height of the form F.
Let $w : \mathbb{R}^n \to \mathbb{R}_{>0}$ be an infinitely differentiable function with compact support $\mathrm{supp}(w)$, and assume that $\nabla F \neq 0$ throughout the closure of this support. Then the associated singular integral is defined by

$$\sigma_\infty(w; F) := \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \int_{|F(\mathbf{x})| \leq \varepsilon} w(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{13}$$

This limit exists and is positive if w is positive at some real solution $\mathbf{x}$ of $F(\mathbf{x}) = 0$, according to the first part of [HB96, Theorem 3]. Intuitively, $\sigma_\infty(w; P_Q)$ measures the real density of solutions, weighted by w.

We first specify the weight function to be used. Define the function $w_0 : \mathbb{R} \to \mathbb{R}_{\geq 0}$ by

$$w_0(x) := \begin{cases} e^{-(1-x^2)^{-1}}, & \text{if } |x| < 1, \\ 0, & \text{if } |x| \geq 1. \end{cases}$$

This function is smooth (infinitely differentiable) with compact support in the interval $[-1, 1]$. Now define the product function

$$w_1(\mathbf{x}) := w_0(2x_1 - 2) w_0(x_2) \cdots w_0(x_4),$$

on $\mathbb{R}^n$. The function w_1 is also smooth, with support contained in $\left[\frac{1}{2}, \frac{3}{2}\right] \times [-1, 1]^3$.

Let $\mathbf{R} \in \mathrm{SO}_n(\mathbb{R})$ be the orthogonal matrix chosen to diagonalize the quadratic form F as before. Then we define the weight function associated to F as

$$w_F(\mathbf{x}) := \widetilde{w}(\mathbf{R}^T \mathbf{x}), \tag{14}$$

where

$$\widetilde{w}(\mathbf{x}) := w_1(|\lambda_1|^{1/2} x_1, \ldots, |\lambda_4|^{1/2} x_4),$$

and $\lambda_1, \ldots, \lambda_4$ are the eigenvalues of the matrix corresponding to F, and assume without loss of generality, that λ_1 is the largest enigenvalue.

Lastly, we write the diagonalized form as

$$F_{\mathrm{sgn}}(\mathbf{X}) := \sigma_1 X_1^2 + \sigma_2 X_2^2 + \cdots + \sigma_4 X_4^2, \tag{15}$$

where $\sigma_i := \lambda_i / |\lambda_i|$. Following the lead of [BD08], we write (13) as

$$\begin{aligned} \sigma_\infty(w_F; F) &= \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \int_{|F(\mathbf{x})| \leq \varepsilon} \widetilde{w}(\mathbf{R}^T \mathbf{x})\, d\mathbf{x} \\ &= \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \int_{|F(\mathbf{Ru})| \leq \varepsilon} w_1(|\lambda_1|^{1/2} u_1, \ldots, |\lambda_4|^{1/2} u_4)\, d\mathbf{u} \\ &= \frac{1}{|\Delta_F|^{1/2}} \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \int_{|F_{\mathrm{sgn}}(\mathbf{v})| \leq \varepsilon} w_1(\mathbf{v})\, d\mathbf{v} \\ &= \frac{\sigma_\infty(w_1; F_{\mathrm{sgn}})}{|\Delta_F|^{1/2}}. \end{aligned}$$

where

$$\sigma_\infty = \sigma_\infty(w_1; F_{\mathrm{sgn}}) \asymp 1, \tag{16}$$

and F_{sgn} is defined in 15. We now record our main theorem for $N_{F,w}(P, m, \mathbf{k})$ which is essentially an upper bound inspired by [Lin17].

Theorem 10.1. *Let F be a non-singular quadratic form in four variables with associated matrix M, and let $m \in \mathbb{N}$ with $\omega(m) = O(1)$ and $\mathbf{k} \in \mathbb{Z}^4$ satisfy $F(\mathbf{k}) \equiv 0 \pmod m$. Let w be a smooth weight function. Then, assuming $\det M$ is a square and $P \gg_\varepsilon m^{2+\varepsilon}|\Delta_F|^{8+\varepsilon}$, we have*

$$N_{F,w_F}(P, m, \mathbf{k}) \ll_\varepsilon \frac{\sigma_\infty \sigma^*(F, m, \mathbf{k})}{|\Delta_F|^{1/2} m^4} P^2 \log P + \rho(F, w_F, m, \mathbf{k}) P^2 + m|\Delta_F|^{15/4+\varepsilon} P^{5/3+\varepsilon}.$$

The coefficient σ_∞ satisfies 16, $\sigma^*(F, m, \mathbf{k})$ is defined in 25 and $\rho(F, w_F, m, \boldsymbol{k})$ is a constant independent of P defined in 32.

We now give the statement of proposition 2.6 in the form of a corollary of theorem 10.1.

Corollary 10.2. *Assume $(k, m) = 1$, g is positive definite Gaussian-reduced binary quadratic form and $\omega(m) = 1$, then for $X \gg_\varepsilon m^{30+\varepsilon}|\Delta_g|^{51/4+\varepsilon}$*

$$\sum_{\substack{z \le X \\ z \equiv k \bmod m}} r_g^2(n) \ll \frac{4^{\omega(\Delta_g)} \log^2(m\Delta_g)}{m^2 |\Delta_g|} X \log X,$$

Remark. The assumption $\omega(m) = 1$ is not essential; it is imposed only to simplify the proof and to suit the intended applications.

The starting point of the proof is the application of Heath-Brown's δ-method [HB96], via Theorem 10.3. Even though the original formulation is in terms of a quadratic form of degree d, the homogeneity of the form is in fact not necessary as noted in [Lin17].

Theorem 10.3. *Let F be a polynomial in four variables. For any $P > 1$ there is a positive constant c_P and an infinitely differentiable function $h(x, y)$ defined on the set $(0, \infty) \times \mathbb{R}$, such that*

$$\sum_{\substack{\mathbf{x} \in \mathbb{Z}^4 \\ F(\mathbf{x}) = 0}} w(P^{-1}\mathbf{x}) = c_P P^{-2} \sum_{\mathbf{c} \in \mathbb{Z}^4} \sum_{q=1}^{\infty} q^{-4} S_q(\mathbf{c}) I_q(\mathbf{c}),$$

where

$$S_q(\mathbf{c}) = \sideset{}{^*}\sum_{a \bmod q} \sum_{\mathbf{b} \bmod q} e_q\big(aF(\mathbf{b}) + \mathbf{c} \cdot \mathbf{b}\big) \tag{17}$$

and

$$I_q(\mathbf{c}) = \int_{\mathbb{R}^4} w(P^{-1}\mathbf{x})\, h\Big(\tfrac{q}{P}, \tfrac{F(\mathbf{x})}{P^2}\Big)\, e_q(-\mathbf{c} \cdot \mathbf{x})\, d\mathbf{x}. \tag{18}$$

The constant c_P satisfies $c_P = 1 + O_N(P^{-N})$ for any $N > 0$.

Following the steps of [Lin17], we apply theorem 10.3 with $F(m\mathbf{x}+\mathbf{k})$ and $w(P^{-1}(m\mathbf{x}+\mathbf{k})$ instead of $F(\mathbf{x})$ and $w(P^{-1}\mathbf{x})$ respectively to obtain

$$N_{F,w}(P, m, \mathbf{k}) = c_P P^{-2} \sum_{\mathbf{c} \in \mathbb{Z}^4} \sum_{q=1}^{\infty} q^{-4} S_{q,m,\mathbf{k}}(\mathbf{c}) I_{q,m,\mathbf{k}}(\mathbf{c}) \tag{19}$$

with

$$S_{q,m,\mathbf{k}}(\mathbf{c}) = \sideset{}{^*}\sum_{a \bmod q} \sum_{\mathbf{b} \bmod q} e_q\big(aF(m\mathbf{b}+\mathbf{k}) + \mathbf{c} \cdot \mathbf{b}\big) \tag{20}$$

and

$$I_{q,m,\mathbf{k}}(\mathbf{c}) = \int_{\mathbb{R}^4} w\big(P^{-1}(m\mathbf{x}+\mathbf{k})\big)\, h\big(P^{-1}q,\, P^{-2}F(m\mathbf{x}+\mathbf{k})\big)\, e_q(-\mathbf{c} \cdot \mathbf{x})\, d\mathbf{x}.$$

Substituting $\mathbf{y} = m\mathbf{x} + \mathbf{k}$ in the above integral, we get

$$I_{q,m,\mathbf{k}}(\mathbf{c}) = m^{-4} e_{mq}(\mathbf{c} \cdot \mathbf{k})\, I_q(\mathbf{c}/m), \tag{21}$$

where $I_q(\mathbf{c})$ is defined as in (18). We now record some properties of the quantity $I_q(\mathbf{c})$ mainly adapted from [BD08] and [Brown07].

Lemma 10.4 ([BD08], Lemma 10)**.** *We have*

$$I_q(0) = \frac{\sigma_\infty P^4}{|\Delta_F|^{1/2}} + O_N\left(\frac{q^N P^{4-N}}{|\Delta_F|^{1/2}}\right)$$

for any $N \geq 0$.

Lemma 10.5 ([BD08], Lemma 8)**.** *Let* $\boldsymbol{c} \in \mathbb{Z}^4$ *with* $\boldsymbol{c} \neq 0$. *Then for any* $N \geq 0$, *we have*

$$I_q(\boldsymbol{c}) \ll_N \frac{P^5}{q|\Delta_F|^{1/2}} \frac{||F||^{N/2}}{|\boldsymbol{c}|^N}.$$

Lemma 10.6 ([BD08], Lemma 9)**.** *Let* $\boldsymbol{c} \in \mathbb{Z}^4$ *with* $\boldsymbol{c} \neq 0$. *Then,*

$$I_q(\boldsymbol{c}) \ll_\epsilon \frac{||F||^{1/2+\epsilon}}{|\Delta_F|^{1/2}} P^{3+\epsilon} q |\boldsymbol{c}|^{-1+\varepsilon}.$$

Lemma 10.7. *It holds that*

$$\frac{\partial^k I_q(0)}{\partial q^k} \ll P^4 |\Delta_F|^{-1/2} q^{-k}, \qquad k \in \{0, 1\}.$$

Proof. Following the notation of [BD08], as introduced in Section 4.2, we have

$$I_q(\mathbf{c}) \;=\; \frac{P^4}{|\Delta_F|^{1/2}} I_r^*(\mathbf{v}; F_{\mathrm{sgn}}, w_1),$$

where

$$I_r^*(\mathbf{v}; G, w) := \int_{\mathbb{R}^n} w(\mathbf{x})\, h\big(r, P_G(\mathbf{x})\big)\, e_r(-\mathbf{v}\cdot\mathbf{x})\, d\mathbf{x},$$

with $r = P^{-1}q$ and $\mathbf{v} = \mathbf{D}\mathbf{R}^T\mathbf{c}$. Moreover, by [Brown07, page 18], one has

$$\frac{\partial}{\partial r} I_r^*(0, F_{\mathrm{sgn}}, w_1) \;\ll\; r^{-k},$$

and hence the lemma follows. □

The main work left is to establish upper bounds for the quantities related to $S_{q,m,\mathbf{k}}(\mathbf{c})$.

11.1 The Sum $S_{q,m,\mathbf{k}}(\mathbf{c})$

The main results concerning the sum $S_{q,m,\mathbf{k}}(\mathbf{c})$ can already be found in [Lin17]; our goal here is simply to make the dependence on the implied constants explicit. It is also worth noting that much of our analysis will involve the local factor

$$\sigma_p(m, \mathbf{k}) := \sum_{n=0}^{\infty} \frac{S_{p^n,m,\mathbf{k}}(0)}{p^{4n}},$$

for a prime p. This quantity is analogous to the familiar p-adic density arising in the circle method, namely

$$\sigma_p := \sum_{n=0}^{\infty} \frac{S_{p^n}(0)}{p^{4n}} = \lim_{\nu\to\infty} p^{-3\nu} \#\big\{\mathbf{x} \in (\mathbb{Z}/p^\nu\mathbb{Z})^4 : F(\mathbf{x}) \equiv 0 \pmod{p^\nu}\big\}. \tag{22}$$

Moreover, one can show in the same way that

$$\sigma_p(m, \mathbf{k}) = \lim_{\nu\to\infty} p^{-3\nu} \#\big\{\mathbf{x} \in (\mathbb{Z}/p^\nu\mathbb{Z})^4 : F(m\mathbf{x}+\mathbf{k}) \equiv 0 \pmod{p^\nu}\big\}, \tag{23}$$

so in particular $\sigma_p(m, \mathbf{k}) = \sigma_p$ whenever $(p, m) = 1$, and clearly $\sigma_p(m, \mathbf{k}) \leq \sigma_p$. We begin by establishing a general upper bound for $S_{q,m,\mathbf{k}}(\mathbf{c})$.

Lemma 10.8. *We have*

$$S_{q,m,\mathbf{k}}(\mathbf{c}) \ll q^3 \gcd^2(m^2\Delta_F, q).$$

Proof. By Cauchy-Schwartz inequality, we have

$$|S_{q,m,\mathbf{k}}(\mathbf{c})|^2 \leq \varphi(q) \sideset{}{^*}\sum_{a \bmod q} \sum_{\mathbf{u},\mathbf{v} \bmod q} e_q\big(a(F(m\mathbf{v}+\mathbf{k}) - F(m\mathbf{u}+\mathbf{k})) + \mathbf{c}\cdot(\mathbf{u}-\mathbf{v})\big). \tag{24}$$

Letting $\mathbf{u}-\mathbf{v} = \mathbf{w}$ and using the fact that $F(m\mathbf{v}+\mathbf{k}) = m^2F(\mathbf{u}) + m\nabla F(k)\cdot\mathbf{u} + F(\mathbf{k})$, we have

$$\begin{aligned} e_q\big(a(F(m\mathbf{v}+\mathbf{k}) &- F(m\mathbf{u}+\mathbf{k})) + \mathbf{c}\cdot(\mathbf{u}-\mathbf{v})\big) \\ &= e_q(a(m^2(F(\mathbf{u}) - F(\mathbf{v})) + m\nabla F(\mathbf{k})\cdot\mathbf{w}) + \mathbf{c}\cdot\mathbf{w}) \\ &= e_q(a(m^2(F(\mathbf{w}) + \mathbf{v}\cdot\nabla F(\mathbf{w})) + m\nabla F\cdot\mathbf{w}) + \mathbf{c}\cdot\mathbf{w}) \\ &= e_q(a(m^2F(\mathbf{w}) + m\nabla F(\mathbf{k})\cdot\mathbf{w}) + \mathbf{c}\cdot\mathbf{w})\ e_q(am^2\mathbf{v}\nabla F(\mathbf{w})). \end{aligned}$$

Now summation over $\mathbf{v}$ in (24) will produce a contribution of zero, unless $q|m^2\nabla F(\mathbf{w})$

$$\begin{aligned} &\varphi(q) \sideset{}{^*}\sum_{a \bmod q} \sum_{\mathbf{w} \bmod q} e_q(a(m^2F(\mathbf{w}) + m\nabla F(\mathbf{k})\cdot\mathbf{w}) + \mathbf{c}\cdot\mathbf{w}) \sum_{\mathbf{v} \bmod q} e_q(a\mathbf{v}\cdot m^2\nabla F(\mathbf{w})) \\ &= \varphi(q)q^4 \sideset{}{^*}\sum_{a \bmod q} \sum_{\substack{\mathbf{w} \bmod q \\ q|m^2\nabla F(\mathbf{w})}} e_q(a(m^2F(\mathbf{w}) + m\nabla F(\mathbf{k})\cdot\mathbf{w}) + \mathbf{c}\cdot\mathbf{w}) \\ &\leq \varphi^2(q)q^4 \#\{\mathbf{w} \bmod q \mid m^2\nabla F(\mathbf{w}) \equiv 0 \bmod q\}. \end{aligned}$$

Recall that M is the matrix associated to our quadratic form F. Let L be the level of F, i.e the smallest integer such that $LM^{-1} \in M_4(\mathbb{Z})$. Using the fact that $\nabla F(\mathbf{w}) = 2M\mathbf{w}$ and $L|\Delta_F$ we have

$$\begin{aligned} \#\{\mathbf{w} \bmod q \mid m^2\nabla F(\mathbf{w}) \equiv 0 \bmod q\} &\leq \#\{\mathbf{w} \bmod q \mid 2m^2LM^{-1}M\mathbf{w}) \equiv 0 \bmod q\} \\ &= \#\{\mathbf{w} \bmod q \mid 2m^2L\mathbf{w} \equiv 0 \bmod q\} \\ &= \gcd^4(2m^2L, q) \\ &\leq \gcd^4(2m^2\Delta_F, q). \end{aligned}$$

Substituting into (24) and taking square roots yields the result. □

With this upper bound at hand, we have the following easy corollary.

Corollary 10.9. *We have*

$$\sum_{q \leq X} |S_{q,m,\mathbf{k}}(\boldsymbol{c})| \ll m^4|\Delta_F|^2X^4.$$

We now give a similar result to lemma 3.1 in [Lin17], making the dependence on the implied constants explicit.

Proposition 10.10. *Let Δ_F be a square, then*

$$\sum_{q \leq X} q^{-4}S_{q,m,\mathbf{k}}(0) = \sigma^*(F,m,\mathbf{k})\log X + \sigma_1(F,m,\mathbf{k}) + O_\varepsilon(m^4|\Delta_F|^{9/4}X^{-1/2+\varepsilon}),$$

where

$$\sigma^*(F,m,\mathbf{k}) = \left(\prod_{p\nmid m}(1-p^{-1})\sigma_p\right)\left(\prod_{p|m}(1-p^{-1})\sigma_p(m,\mathbf{k})\right), \tag{25}$$

and

$$\sigma_1(F,m,\mathbf{k}) = \sigma^*(F,m,\mathbf{k}) \times \left(\gamma + \sum_p \frac{d}{ds}\log\left((1-p^{3-s})\sum_{n=0}^{\infty} S_{p^n,m,\mathbf{k}}(0)\,p^{-ns}\right)\Big|_{s=4}\right) \tag{26}$$

with $\gamma = 0.5772..$ the Euler-Mascheroni constant.

Proof. We'll be closely following [HB96] section 11 and [Brown07] section 3. First we define the Dirichlet series

$$\zeta(s;\mathbf{c}) = \sum_{q=0}^{\infty} q^{-s} S_{q,m,\mathbf{k}}(\mathbf{c})$$

as in [HB96] which is absolutely convergent for $re(s) > 4$ by lemma 10.8. We also notice that $S_{q,m,\mathbf{k}}(0)$ is multiplicative and we can write

$$\zeta(s;0) = \prod_p D_p(s)$$

where $D_p(s) = \sum_{n=0}^{\infty} p^{-ns} S_{p^n,m,\mathbf{k}}(0)$. We now investigate the factors $D_p(s)$ more carefully, for which we must distinguish between whether p is a divisor of $2\Delta_F$ or not. If $p|2\Delta_F$ then by lemma 10.8 we have

$$|D_p(s)| \ll \sum_{n=0}^{\infty} p^{-re(s)n} p^{3n} p^{2\nu_p(m^2\Delta_F)} \ll p^{2\nu_p(m^2\Delta_F)}$$

and hence

$$\prod_{p|2\Delta_F} D_p(s) \ll m^4 |\Delta_F|^2.$$

For $p \nmid 2\Delta_F$, we first have $S_{p^n,m,\mathbf{k}}(0) = S_{p^k}(0)$. Hence by lemma 26 of [HB96] and lemma 10.8 again, we have

$$D_p(s) = 1 + \chi_{\Delta_F}(p) p^{2-s}(p-1) + O_\delta(p^{-1-2\delta}) \tag{27}$$

for $\mathrm{re}(s) \geq 7/2 + \delta$. Upon combining both products over primes and recalling that Δ_F is a square, i.e. $L(s, \chi_{\Delta_F}) = \zeta(s) \prod_{p|\Delta_F}(1 - p^{-s})$ with $\zeta(s)$ the Riemann zeta function, we obtain

$$\zeta(s;0) = \zeta(s-3) E(s,0) \tag{28}$$

where $|E(s,0)| \ll_\delta m^4 |\Delta_F|^{2+\varepsilon}$. I particular, $\zeta(s,0)$ has analytic continuation to the half-plane $\mathrm{re}(s) > 7/2$. We now argue along the lines of [HB96] lemma 30. Let Y be half an odd integer; it follows from an application of Perron's formula ([Tit86], lemma 3.12) together with corollary 10.9 that

$$\sum_{q \leq Y} S_{q,m,\mathbf{k}}(0) = \frac{1}{2\pi i} \int_{6-iT}^{6+iT} \frac{\zeta(s;0) Y^s}{s} ds + O\left(\frac{m^4 |\Delta_F|^2 Y^6}{T}\right) \tag{29}$$

for any $T \geq 1$. We let $\alpha = 7/2 + \varepsilon$ for $\varepsilon > 0$ and we proceed to move the line of integration back to α and pick up the doubel pole at $s = 4$. Now (28) yields

$$\zeta(\sigma + it; 0) \ll_\varepsilon m^4 |\Delta_F|^{2+\varepsilon} \zeta(\sigma - 3 + it) \tag{30}$$

for $\sigma = \mathrm{re}(s) \in [\alpha, 6]$. Using the fact that

$$|\zeta(\sigma + it)| \ll_\sigma \begin{cases} |t|^{1-\sigma+\varepsilon}, & \frac{1}{2} \leq \sigma \leq 1, \quad |t| \geq 1, \\ 1, & \sigma > 1, \end{cases}$$

we obtain the upper bound

$$\zeta(\sigma + it; 0) \ll_\varepsilon \begin{cases} m^4 |\Delta_F|^{2+\varepsilon} |t|^{4-\sigma+\varepsilon}, & \sigma \in [\alpha, 4] \\ m^4 |\Delta_F|^{2+\varepsilon}, & \sigma \in (4, 6] \end{cases}.$$

In particular,

$$\int_{\alpha \pm iT}^{6 \pm iT} \frac{\zeta(s;0)\, Y^s}{s}\, ds \;\; \ll_\varepsilon \;\; m^4 |\Delta_Q|^{2+\varepsilon} \left(\frac{Y^{7/2+\varepsilon} T^\varepsilon}{T^{1/2}} \;+\; \frac{Y^6}{T} \right).$$

For the contribution from the vertical lines, we employ the mean value estimate,

$$\int_0^U |\zeta(\sigma+it)|^2 dt \ll_\sigma U$$

that is valid for $\sigma \in (1/2, 1)$. It follows from this and application of 30 and cauchy-schwarz inequality that

$$\begin{aligned}\int_{\alpha-iT}^{\alpha+iT} \frac{\zeta(s;0)Y^s}{s} ds &\ll_\varepsilon m^4|\Delta_F|^{2+\varepsilon} Y^{7/2+\varepsilon} \int_0^T \frac{|\zeta(1/2+\varepsilon+it)|}{|1+t|} dt \\ &\ll m^4|\Delta_F|^{2+\varepsilon} Y^{7/2+\varepsilon} T^\varepsilon.\end{aligned}$$

As argued in [HB96], $\zeta(s+4;0)Y^s/s$ has a double pole at $s=0$, since Δ_F is a square, with the residue taking the form

$$\sigma^*(F,m,\mathbf{k}) \log Y + \sigma_1(F,m,\mathbf{k})$$

where

$$\sigma^*(F,m,\mathbf{k}) = \left(\prod_{p\nmid m}(1-p^{-1})\sigma_p\right)\left(\prod_{p|m}(1-p^{-1})\sigma_p(m,\mathbf{k})\right)$$

and $\sigma_1(F,m,\mathbf{k})$ is a constant solely depending on F. Putting everything together in (29) and choosing $T = Y^{5/2}$, we obtain

$$\sum_{q\le Y} S_{q,m,\mathbf{k}}(0) \;=\; \sigma^*(F,m,\mathbf{k})\, Y^4 \log Y \;+\; \sigma_1(F,m,\mathbf{k})\, Y^4 \;+\; O_\varepsilon\Big(m^4|\Delta_F|^{2+\varepsilon} Y^{7/2+\varepsilon}\Big).$$

Applying partial summation yields the desired result. It remains to analyze the constant $\sigma_1(F,m,\mathbf{k})$ in more detail, for which we now examine the residue of the double pole more carefully. We begin with

$$\zeta(s;0) = \prod_p \left(\sum_{k=0}^\infty S_{p^k}(0)\, p^{-ks}\right) = \zeta(s-3)\prod_p \left((1-p^{3-s})\sum_{k=0}^\infty S_{p^k}(0)\, p^{-ks}\right),$$

where

$$\zeta(s-3) = \frac{1}{s-4} + \gamma + O(s-4).$$

The Euler product converges absolutely in a neighborhood of $s=4$, and each local factor equals $\sigma_p(m,\mathbf{k})$ at $s=4$. For any analytic function f near $s=4$, we have the expansion

$$f(s) = f(4)\left(1 + \left(\frac{d}{ds}\log f\right)(4)\,(s-4) + \cdots\right).$$

Applying this to each factor gives

$$(1-p^{3-s})\sum_{k=0}^\infty S_{p^k,m,\mathbf{k}}(0)\, p^{-ks} = \sigma_p(m,\mathbf{k})\left(1 + \left(\frac{d}{ds}\log\Big((1-p^{3-s})\sum_{k=0}^\infty S_{p^k,m,\mathbf{k}}(0)\, p^{-ks}\Big)\Big|_{s=4}\right)(s-4) + \cdots\right).$$

Hence

$$\zeta(s,0) = \sigma^*(F,m,\mathbf{k})\left(\frac{1}{s-4} + \gamma + \sum_p \frac{d}{ds}\log\Big((1-p^{3-s})\sum_{k=0}^\infty S_{p^k,m,\mathbf{k}}(0)\, p^{-ks}\Big)\Big|_{s=4} + \cdots\right).$$

By multiplying by $X^{s-4}/(s-4)$, it follows that the residue from the double pole is

$$\sigma^*(F,m,\mathbf{k})\left(\log X + \gamma + \sum_p \frac{d}{ds}\log\Big((1-p^{3-s})\sum_{k=0}^\infty S_{p^k,m,\mathbf{k}}(0)\, p^{-ks}\Big)\Big|_{s=4}\right),$$

and consequently

$$\sigma_1(F,m,\mathbf{k}) \;=\; \sigma^*(F,m,\mathbf{k})\left(\gamma+\sum_p \frac{d}{ds}\log\Big((1-p^{3-s})\sum_{k=0}^{\infty} S_{p^k,m,\mathbf{k}}(0)\,p^{-ks}\Big)\Big|_{s=4}\right).$$

□

The last ingredient needed to activate the circle method in (19) is an upper bound for

$$S(X,\mathbf{c},m,\mathbf{k}) \;=\; \sum_{q\le X} q^{-4} S_{q,m,\mathbf{k}}(\mathbf{c})\, e_{mq}(\mathbf{c}\cdot\mathbf{k}), \tag{31}$$

for $\mathbf{c}\neq 0$ where the factor $e_{mq}(\mathbf{c}\cdot\mathbf{k})$ arises from (21). The estimate can be thought of being analogous to lemma 30 of [HB96] for $\mathbf{c}\neq 0$. This estimate was established in Lemma 3.4 of [Lin17], so we do not repeat the argument here. Our purpose is only to make the dependence of the implied constant explicit. The idea is that [Lin17] was able to decompose the sum in (31) into $S(X,\mathbf{c},m,\mathbf{k}) = S_1 + S_2$ with

$$S_1 \;=\; \sum_{\substack{v|m^\infty\\ v\le X^{1/3}}}\sum_{u\le X/v} e_{mv}\Big(\overline{u}^{(mv)}\,\mathbf{c}\cdot\mathbf{k}\Big)\,S_u(\mathbf{c})\,S_{v,m,\mathbf{k}}(\overline{u}^{(v)}\mathbf{c}),$$

and

$$S_2 \;=\; \sum_{\substack{v|m^\infty\\ X^{1/3}<v\le X}}\sum_{u\le X/v} e_{mv}\Big(\overline{u}^{(mv)}\,\mathbf{c}\cdot\mathbf{k}\Big)\,S_u(\mathbf{c})\,S_{v,m,\mathbf{k}}\big(\overline{u}^{(v)}\mathbf{c}\big),$$

where $a|b^\infty$ means that all the prime factors of a are prime factors of b and $\bar{a}^{(b)}$ is the inverse of a modulo b.

The second sum can be bounded above, using lemma 10.8, by

$$|S_2| \ll \sum_{\substack{v|m^\infty\\ X^{1/3}<v\le X}}\sum_{u\le X/v} m^4|\Delta_F|^4u^3v^3 \ll m^4|\Delta_F|^4\sum_{\substack{v|m^\infty\\ v\le X}}\sum_{v\le X^{2/3}} X^3 \ll_\varepsilon m^4|\Delta_F|^4X^{11/3+\varepsilon}.$$

We now handle the sum S_1 in the same as in [Lin17], for that purpose we'll need a lemma. We denote by Δ the conductor of the Dirichlet character $\chi_\Delta = \big(\frac{\Delta_F}{\cdot}\big)$ just as in [Lin17].

Lemma 10.11 ([Lin17], Lemma 3.3)**.** *Let χ be a Dirichlet character of conductor mv for some $v \mid m^\infty$. Then for any $\varepsilon>0$, we have*

$$\sum_{u\le X}\chi(u)S_u(\mathbf{c}) \;=\; \eta(\mathbf{c})\,\xi(\chi)\,\sigma'(F,m)\,\frac{X^4}{4} \;+\; O_\varepsilon\big(X^{7/2+\varepsilon}|\mathbf{c}|^\varepsilon\big),$$

where

$$\sigma'(F,m) \;=\; \left(\prod_{p\nmid m}(1-p^{-1})\,\sigma'_p\right)\left(\prod_{p|m}(1-p^{-1})\right),$$

and

$$\sigma'_p \;=\; \sum_{t\ge 0} p^{-4t}\,S_{p^t}(0)\,\chi_\Delta(p)^t.$$

Here $\eta(\mathbf{c})=1$ if $M^{-1}(\mathbf{c})=0$ and 0 otherwise, and $\xi(\chi)=1$ if $\chi=\chi_0\chi_\Delta$ and 0 otherwise. Recall that χ_0 is the principal character modulo m.

We apply the above lemma for S_1 as in [Lin17], bearing in mind that in our case $\xi(\chi)=0$ since $m=\ell^k$ is assumed to be a power of a prime ℓ which excludes the case that χ splits as $\chi_0\chi_\Delta$ (since the conductor of χ_0 is different from mv given that χ_Δ is trivial). We obtain

$$S_1 = \sum_{\substack{v|m^\infty\\ v\le X^{1/3}}}\sum_{\chi \,(\mathrm{mod}\, mv)} a_\chi(v,\mathbf{c},m,\mathbf{k})\left(O_\varepsilon\Big(\big(\tfrac{X}{v}\big)^{7/2+\varepsilon}|\mathbf{c}|^\varepsilon\Big)\right),$$

where

$$a_\chi(v,\mathbf{c},m,\mathbf{k}) = \frac{1}{\varphi(mv)} \sum_{x \pmod{mv}} \chi(x)\, e_{mv}(x\mathbf{c}\cdot\mathbf{k})\, S_v(x\mathbf{c};m,\mathbf{k}).$$

A trivial bound for $a_\chi(v,\mathbf{c},m,\mathbf{k})$ using lemma 10.8

$$a_\chi(v,\mathbf{c},m,\mathbf{k}) \ll m^4|\Delta_F|^2 v^3$$

implies that S_1 can be bounded above by

$$\sum_{v|m^\infty,\ v\leq X^{1/3}} O_\varepsilon(m^5|\Delta_F|^2 v^{1/2-\varepsilon} X^{7/2+\varepsilon}|\mathbf{c}|^\varepsilon) \ll_\varepsilon m^5|\Delta_F|^2 X^{11/3+\varepsilon}|\mathbf{c}|^\varepsilon,$$

where we have used the fact that the number of values v we're summing over is bounded above by $\log^{\omega(m)}(X) \ll_\varepsilon X^\varepsilon$ since we assumed from the beginning that $\omega(m)=1$.

Combining both estimates from S_1 and S_2 we obtain the following lemma, which is an improvement upon lemma 3.4 of [Lin17].

Corollary 10.12. *For any $\varepsilon>0$ and $\mathbf{c}\neq 0$ it holds that*

$$S(X,\mathbf{c},m,\mathbf{k}) \ll_\varepsilon m^5|\Delta_F|^4 X^{11/3+\varepsilon}|\mathbf{c}|^\varepsilon.$$

Derivation of the Upper Bound. Here we gather all the previous tools to prove theorem 10.1 in accordance with theorem 10.3. We proceed as in [HB96] section 12.

We begin our analysis by noting $I_q(\mathbf{c})$ vanishes for $q \gg P$ and that lemmas 10.5 and 10.8 imply

$$\begin{aligned}\sum_{|\mathbf{c}|\geq P}\sum_{0<q\ll P} q^{-4} S_{q,m,\mathbf{k}}(\mathbf{c}) e_{mq}(\mathbf{c}\cdot\mathbf{k}) I_q(\mathbf{c}/m) &\ll_\varepsilon \sum_{|\mathbf{c}|\geq P}\sum_{0\leq q\ll P} q^{-2}|\Delta_F|^{3/2}P^5||F||^{N/2} \\ &\ll_\varepsilon ||F||^{9/2+\varepsilon}|\Delta_F|^{3/2}P^3,\end{aligned}$$

upon taking $N=4+\varepsilon$ so that the sum in $\mathbf{c}$ converges.
We now look at the sum over $|\mathbf{c}|\leq P$ and $c\neq 0$. Here lemmas 10.6 and corollary 10.12 (since the main term in lemma 10.11 vanishes) along with partial summation implies that

$$\begin{aligned}\sum_{R<q\leq 2R} q^{-4} S_{q,m,\mathbf{k}}(\mathbf{c}) e_{mq}(\mathbf{c}\cdot\mathbf{k}) I_q(\mathbf{c}/m) &\ll A(R)I_R(\mathbf{c}/m) - A(R)I_R(\mathbf{c}/m) \\ &\ll A(R)I_R(\mathbf{c}/m) \\ &\ll_\varepsilon m^4\Delta_F^2 R^{-1/3+\varepsilon}\left|\frac{||F||^{1/2+\varepsilon}m^{1-\varepsilon}}{|\Delta_F|^{1/2}|\mathbf{c}|^{1+\varepsilon}}P^3R\right| \\ &\ll_\varepsilon m^{5-\varepsilon}|\Delta_F|^{3/2}||F||^{1/2+\varepsilon}P^{3+\varepsilon}R^{2/3+\varepsilon},\end{aligned}$$

for any $R>0$, whence

$$\sum_{q=1}^{\infty} q^{-4} S_{q,m,\mathbf{k}}(\mathbf{c}) e_{mq}(\mathbf{c}\cdot\mathbf{k}) I_q(\mathbf{c}) \ll_\varepsilon m^5||F||^{1/2}|\Delta_F|^{3/2+\varepsilon}P^{11/3+\varepsilon}.$$

Extending the sum over $0<|c|<P$, using our former upper bound for $|\mathbf{c}|\geq P$ and plugging in theorem 10.3 we obtain

$$P^{-2}m^{-4}\sum_{|\mathbf{c}|\geq P}\sum_{q=1}^{\infty} q^{-4}\, S_{q,m,\mathbf{k}}(\mathbf{c})\, e_{mq}\big(\mathbf{c}\cdot\mathbf{k}\big)\, I_q(\mathbf{c}/m) \ll_\varepsilon m||F||^{2+\varepsilon}|\Delta_F|^{3/2+\varepsilon}P^{5/3+\varepsilon}.$$

There remains the case $\mathbf{c}=0$. Note that $I_q(\mathbf{c})$ vanishes for $q\gg P$. Using Lemma 10.4, we have

$$I_{q,m,k}(0) \leq m^{-4}I_q(0) = \frac{\sigma_\infty}{m^4|\Delta_f|^{1/2}}\,P^4 + O_N\left(\frac{q^N\,P^{4-N}}{m^4|\Delta_f|^{1/2}}\right) \ll \frac{\sigma_\infty\,P^4}{m^4|\Delta_f|^{1/2}}.$$

Therefore,

$$\sum_{q\ll P} q^{-4}\, S_{q,m,\mathbf{k}}(0)\, I_{q,m,\mathbf{k}}(0) \ll m^{-4}P^{-2}\frac{\sigma_\infty P^4}{|\Delta_f|^{1/2}}\sum_{q\ll P} q^{-4} S_{q,m,\mathbf{k}}(0).$$

By Proposition 10.10, this is

$$\ll \frac{\sigma_\infty P^2}{m^4|\Delta_f|^{1/2}}\Big(\sigma^*(F,m,\mathbf{k})\log P + \sigma_1(F,m,\mathbf{k}) + m^4|\Delta_f|^{9/4}P^{-1/2+\varepsilon}\Big).$$

Combining this with the contribution from $\mathbf{c}\neq 0$ in theorem 10.3, we finally obtain

$$N_F(\omega,m,\mathbf{k}) \ll \frac{\sigma_\infty\sigma^*(F,m,\mathbf{k})}{m^4|\Delta_f|^{1/2}}\,P^2\log P + \frac{\sigma_\infty\sigma_1(F,m,\mathbf{k})}{m^4|\Delta_f|^{1/2}}\,P^2 + m\,\|F\|^{9/4}|\Delta_f|^{3/2}\,P^{5/3+\varepsilon}.$$

where

$$\rho(F,w_F,m,\mathbf{k}) = \frac{\sigma_\infty\sigma_1(F,m,\mathbf{k})}{m^4|\Delta_f|^{1/2}} \tag{32}$$

with $\sigma_1(F,m,\mathbf{k})$ defined in 26. We finally note that since $F = g\oplus g$ then $||F|| = ||g|| \ll \Delta_g = |\Delta_F|^{1/2}$ since g is assumed to be gaussian reduced. This complete the proof of theorem 10.1. □

To prove corollary 10.12, we need bounds on $\sigma^*(F,m,\mathbf{k})$, $\sigma_1(F,m,\mathbf{k})$ and to replace the weighted counting function $N_F(w,m,\mathbf{k})$ by the sum $\sum_{\substack{n\leq X\\ n\equiv k \bmod m}} r_g^2(n)$.
We start by giving a bound for $\sigma^*(F,m,\mathbf{k})$.

Lemma 10.13. *Let $\sigma^*(F,m,\mathbf{k})$ be as defined in 25, then*

$$\sigma^*(F,m,\mathbf{k}) \ll 4^{\omega(\Delta_F)},$$

where $\omega(\cdot)$ is the prime divisor counting function.

Proof. Note that by definition we have

$$\sigma^*(F,m,\mathbf{k}) = \left(\prod_{p\nmid m}(1-p^{-1})\sigma_p\right)\left(\prod_{p|m}(1-p^{-1})\sigma_p(m,\mathbf{k})\right),$$

and $\sigma_p(m,\mathbf{k})\leq\sigma_p$ using 22 and 23. We start by bounding σ_p over primes $p\nmid\Delta_F$, for which we have by 27 and $s=4$ that

$$\sigma_p = 1 + p^{-1} + O(p^{-2}).$$

Hence

$$\prod_{p\nmid\Delta_F}(1-p^{-1})\sigma_p = \prod_{p\nmid\Delta_F}(1+O(p^{-2})) \ll 1,$$

where the final inequality follows from the Riemann zeta function with $s=2$ (i.e. converges absolutely).
For primes $p|\Delta_F$, the same argument as in the last section of [Brown07] over the p-adics gives

$$\prod_{p|\Delta_F}(1-p^{-1})\sigma_p \ll 4^{\omega(\Delta_F)};$$

the lemma then follows. □

We now turn to bounding $\sigma_1(F,m,\mathbf{k})$. For that we need a slightly stronger form of lemma 10.8.

Lemma 10.14. *For any odd $q = p^t$ with p prime satisfying $p|m$, we have*

$$S_{q,m,\mathbf{k}}(\mathbf{0}) \ll q^3(q, m^4\Delta_g),$$

where $\Delta_g = \sqrt{|\Delta_F|}$ and $F(x, y, s, t) = g(x, y) - g(s, t)$.

Proof. We recall (20). Using the decomposition $F(x, y, s, t) = g(x, y) - g(s, t)$ and the special shift $\mathbf{k} = (\mathbf{k}^*, \mathbf{k}^*)$ with $\mathbf{k}^* = (k_1, k_2)$, we obtain

$$S_{q,m,\mathbf{k}}(\mathbf{0}) = \sideset{}{^*}\sum_{a \bmod q} \left(\sum_{\mathbf{b}\in(\mathbb{Z}/q\mathbb{Z})^2} e_q\big(a\, g(m\mathbf{b} + \mathbf{k}^*)\big) \right)^2.$$

Let G be the Gram matrix of g. Then

$$g(m\mathbf{b} + \mathbf{k}^*) = m^2 g(\mathbf{b}) + m\, \mathbf{b}^T G\mathbf{k}^* + g(\mathbf{k}^*),$$

and the constant term may be discarded for bounds. Hence

$$S_{q,m,\mathbf{k}}(\mathbf{0}) = \sideset{}{^*}\sum_{a \bmod q} \left(\sum_{\mathbf{b} \bmod q} e_q\big(am^2 g(\mathbf{b}) + am\, (\mathbf{b}^T G) \cdot \mathbf{k}^*\big) \right)^2.$$

Write

$$L_q := \sum_{\mathbf{b} \bmod q} e_q\big(\mathbf{b}^T A\mathbf{b} + \boldsymbol{\ell} \cdot \mathbf{b}\big), \qquad A = m^2 G,\ \boldsymbol{\ell} = mG\mathbf{k}^*.$$

Since $q = p^\nu$ with p odd, then there exists a unitary $U \in GL_2(\mathbb{Z}_p)$ such that $U^T GU = \mathrm{diag}(d_1, d_2)$ with $d_1 d_2 = \Delta_g$ and $d_i = p^{t_i} u_i$ with u_i units in $\mathbb{Z}_p$. For such a matrix U over $\mathbb{Z}_p$ we have the projection $\bar{U} \in GL_2(\mathbb{Z}/q\mathbb{Z})$. Making the change of variables $\mathbf{b} \mapsto \bar{U}\mathbf{b}$ yields

$$L_q = \prod_{i=1}^{2} \left(\sum_{b \bmod q} e_q(am^2 d_i b^2) \right)^2.$$

By the classical Gauss sum bound,

$$\sum_{b \bmod q} e_q(Ab^2) \ll q^{1/2}(q, A)^{1/2},$$

and therefore

$$|L_q| \ll q^2(q, m^2 d_1)(q, m^2 d_2).$$

Using the key identity of prime-power q we have $(q, m^2 d_1)(q, m^2 d_2) = (q, m^4 d_1 d_2)$. Summing over $a \in (\mathbb{Z}/q\mathbb{Z})^\times$ gives

$$S_{q,m,\mathbf{k}}(\mathbf{0}) \ll q^3(q, m^4\Delta_g),$$

as claimed.

□

Lemma 10.15. *Let $\sigma_1(F, m, \mathbf{k})$ be as in 26, then*

$$\sigma_1(F, m, \mathbf{k}) \ll 4^{\omega(\Delta_F)} \log^2(m\Delta_F).$$

Proof. We have

$$\sigma_1(F, m, \mathbf{k}) = \sigma^*(F, m, \mathbf{k}) \times \left(\gamma + \sum_p \frac{d}{ds} \log\left((1 - p^{3-s}) \sum_{k=0}^{\infty} S_{p^k}(0)\, p^{-ks} \right) \Big|_{s=4} \right).$$

For $p \nmid \Delta_F$,

$$S_{p^k}(0) = p^{3k}(1 - p^{-1}) \quad (k \geq 1),$$

and hence

$$(1 - p^{3-s}) \sum_{k=0}^{\infty} S_{p^k}(0)\, p^{-ks} = 1 - p^{2-s}.$$

Therefore,

$$\frac{d}{ds} \log(1 - p^{2-s})\Big|_{s=4} = \frac{p^{-2}}{1 - p^{-2}} \log p.$$

This shows that the sum over primes $p \nmid \Delta_F$ converges absolutely to a constant. It remains to bound the contribution from primes dividing Δ_F. For those primes we compute the derivative and obtain

$$\gamma + \sum_p \frac{d}{ds} \log\Big((1 - p^{3-s}) \sum_{n=0}^{\infty} S_{p^n,m,\mathbf{k}}(0)\, p^{-ns}\Big)\Big|_{s=4} = \gamma + \sum_p (\log p) \left(\frac{1}{p-1} - \frac{\sum_{n=0}^{\infty} n\, S_{p^n,m,\mathbf{k}}(0)\, p^{-4n}}{\sigma_p(m,\mathbf{k})} \right).$$

The first sum can be bounded above by a constant depending on $\omega(\Delta_F)$, for the second term we use 10.14. Set

$$t := v_p(m^4 \Delta_g).$$

We split the sum

$$\sum_{k=0}^{\infty} k\, S_{p^k,m,\mathbf{k}}(\mathbf{0})\, p^{-4k}$$

into two ranges.

Range I: $0 \leq k \leq t$. Since $(p^k, m^4\Delta_g) = p^k$ for $k \leq t$, we have

$$S_{p^k,m,\mathbf{k}}(\mathbf{0}) \ll p^{3k}\, p^k = p^{4k}.$$

Hence

$$\sum_{k=0}^{t} k\, S_{p^k,m,\mathbf{k}}(\mathbf{0})\, p^{-4k} \ll \sum_{k=0}^{t} k \ll t^2.$$

Range II: $k > t$. In this range $(p^k, m^4\Delta_g) = p^t$, so

$$S_{p^k,m,\mathbf{k}}(\mathbf{0}) \ll p^{3k}\, p^t.$$

Therefore

$$\sum_{k>t} k\, S_{p^k,m,\mathbf{k}}(\mathbf{0})\, p^{-4k} \ll p^t \sum_{k>t} k\, p^{-k} \ll p^t\, p^{-t} \ll 1.$$

Combining the two ranges, we obtain

$$\sum_{k=0}^{\infty} k\, S_{p^k,m,\mathbf{k}}(\mathbf{0})\, p^{-4k} \ll t^2 + 1 \ll t^2,$$

where $t = v_p(m^4\Delta_g)$. A lower bound of $\sigma_p(m,\mathbf{k})$. By noting that the derivative is multiplied by $\sigma^*(F,m,\mathbf{k})$ the term $\sigma_p(m,\mathbf{k})$ in the denominator gets canceled and we obtain

$$\begin{aligned}
\sigma_1(F,m,\mathbf{k}) &\ll \sigma^*(F,m,\mathbf{k}) + \sum_{p|\Delta_F} \left(\prod_{q \neq p} \sigma_q(m,\mathbf{k}) \right) (\log p) v_p(m^4\Delta_g) \\
&\ll 4^{\omega(\Delta_F)} + 4^{\omega(\Delta_F)} \log^2(m\Delta_g) \\
&\ll 4^{\omega(\Delta_F)} \log^2(m\Delta_F),
\end{aligned}$$

where we have used the fact that $\sum_{p|A} \log(A) v_p^2(A) \leq max_p(v_p)^2 \log A \ll \log^2 A$. □

We now have all the tools to prove corollary 10.2.

Proof of corollary 10.2. Using Theorem 10.1 together with the bounds from Lemmas 10.13 and 10.15, we obtain

$$N_{F,w}(P,m,\mathbf{k}) \ll_\varepsilon \frac{4^{\omega(\Delta_F)}\log^2(m\Delta_F)}{m^4|\Delta_F|^{1/2}}\,P^2\log P \;+\; m\|F\|^{9/2}|\Delta_F|^{3/2+\varepsilon}P^{5/3+\varepsilon}.$$

Once

$$P \gg_\varepsilon m^{15+\varepsilon}|\Delta_F|^{51/4+\varepsilon},$$

the main term dominates.

The final step is to pass from the weighted counting function to the unweighted one. There are several ways to do this; here we follow the argument of [Brown07, Theorem 2]. In particular, we define the weight functions $w_F^{(i)}$ as in (14), but with w_1 replaced by

$$w^{(i)}(\mathbf{x}) := w_0(2x_i-2)\prod_{k\neq i} w_0(x_k), \qquad 1\le i\le 4.$$

This yields the dyadic bound

$$\begin{aligned} N_F(P,m,\mathbf{k}) &:= \#\Big\{\mathbf{x}\in\mathbb{Z}^4 : F(\mathbf{x})=0,\ \mathbf{x}\equiv\mathbf{k} \pmod m,\ \max_{1\le i\le 4}|a_ix_i^2|\le P\Big\} \\ &\ll 1+\sum_{j=0}^{\infty}\sum_{i=1}^{4} N_{F,w^{(i)}}\left(\frac{P}{2^j},m,\mathbf{k}\right) \\ &\ll \frac{4^{\omega(\Delta_F)}\log^2(m\Delta_F)}{m^4|\Delta_F|^{1/2}}\,P^2\log P. \end{aligned}$$

To conclude the main problem of this section, namely the proof of Proposition 2.6, recall from the introduction of the appendix that each integer n representable by $g(x,y)$ with $n\equiv k \pmod m$ corresponds to a pair (x,y) lying in the set

$$\mathcal{S} := \mathcal{S}_{g,m,k} = \big\{\,\mathbf{v}\in(\mathbb{Z}/m\mathbb{Z})^2 : g(\mathbf{v})\equiv k \pmod m\,\big\}.$$

Therefore, it suffices to obtain a bound for $N_F(P,m,\mathbf{k})$ and then sum over all representatives in $\mathcal{S}\times\mathcal{S}$.

The cardinality of $\mathcal{S}$ is well understood. For example, Lemma 6.3 of [Mat12] implies

$$|\mathcal{S}| \ll m,$$

upon combining with the Chinese remainder theorem. Taking $P = X^{1/2}$, we deduce

$$\sum_{\substack{z\le X\\ z\equiv k \pmod m}} r_g^2(z) \;\ll\; |\mathcal{S}|^2\,N_F(X^{1/2},m,\mathbf{k}) \;\ll\; \frac{4^{\omega(\Delta_g)}\log^2(m\Delta_g)}{m^2|\Delta_g|}\,X\log X,$$

provided that

$$X \gg m^{30+\varepsilon}|\Delta_g|^{51/4+\varepsilon}.$$

This completes the proof of Corollary 10.2. □