Secondary terms in the first moment of $|\mathrm{Sel}_2(E)|$

Arul Shankar and Takashi Taniguchi

Abstract

We prove the existence of secondary terms of order $X^{3/4}$, with power saving error terms, in the counting functions of $|\mathrm{Sel}_2(E)|$, the 2-Selmer group of E, for elliptic curves E having height bounded by X. This is the first improvement on the error term of $o(X^{5/6})$, proved in [14], where the primary term of order $X^{5/6}$ for this counting function was obtained.

Contents

1 Introduction

The Poonen–Rains heuristics [37] predict the distribution of the p-Selmer groups of elliptic curves over $\mathbb{Q}$, for all primes p. These heuristics are supported by works of Bhargava and the first named author [14, 15, 13], where it is proven that when the set of all elliptic curves over $\mathbb{Q}$ is ordered by height, the average size of the p-Selmer groups is $p+1$ for $p = 2$, 3, and 5. On the computational side, Balakrishnan–Ho–Kaplan–Spicer–Stein–Weigandt [3] collect and analyze data on ranks, 2-Selmer groups, and other arithmetic invariants of elliptic curves when they are ordered by height. They find a persistently smaller average size of the 2-Selmer group in the data. This is despite the fact that the average rank appears bigger in the data than its predicted value of 0.5. This discrepancy between the Goldfeld [29] and Katz–Sarnak [34] prediction of the average rank and the data was observed, for families of quadratic twists of an elliptic curve, and the family of all elliptic

curves ordered by conductor, in [4]. Thus it is natural to ask whether there exists a secondary term in the counting function of $|\mathrm{Sel}_2(E)|$, which explains the discrepancy between the data and and the theory.

Our main result proves the existence of a secondary term in this counting function. More precisely, let $\mathcal{E}$ denote the family of all elliptic curves over $\mathbb{Q}$. Every elliptic curve in $\mathcal{E}$ can be uniquely represented in the form $E_{AB} : y^2 = x^3 + Ax + B$, where A and B are integers such that $p^4 \mid A$ implies that $p^6 \nmid B$ for all primes p. We define the height $H(E_{AB}) := \max\{4|A|^3, 27B^2\}$, and for a real number X, define the set

$$\mathcal{E}_X^{\pm} := \{E \in \mathcal{E} : H(E) < X,\ \pm\Delta(E) > 0\},$$

where $\Delta(E)$ is the discriminant of E. Then we prove the following result.

Theorem 1 *With notation as above, we have*

$$\sum_{E \in \mathcal{E}_X^{\pm}} |\mathrm{Sel}_2(E)| = 3 \cdot \sum_{E \in \mathcal{E}_X^{\pm}} 1 + C(\mathcal{E})^{\pm} X^{3/4} + O_\epsilon(X^{3/4-\alpha+\epsilon}),$$

for constants $C(\mathcal{E})^{\pm}$ and some $\alpha > 0$. (More precisely, we show that we can take α to be $1/3804$.)

Since the size of $\mathcal{E}_X^{\pm}$ grows like $c^{\pm}X^{5/6} + O(X^{1/2})$ for positive constants $c^{\pm}$, the above result recovers a secondary term with a power saving error term for the sum of $|\mathrm{Sel}_2(E)|$ over elliptic curves E ordered by height. In fact, this result is the first instance of a power saving error term obtained for the counting function of the 2-Selmer groups of elliptic curves.

We express the constants $C^{\pm}$ in Theorem 1 as the limit of the values at $s = 1/2$ of certain Dirichlet series' which converges absolutely only to the right of $\mathrm{Re}(s) = 1$. We prove that these Dirichlet series' have a holomorphic continuation to the right of $\Re(s) = 1/3$ except for a simple pole at $s = 1$ (in particular, their values at $s = 1/2$ are well defined!), and that the limit of these values at $s = 1/2$ exists. However, we are not yet able to find closed form formulas for $C^{\pm}$, or numerically evaluate them, leaving this investigation to future work.

There has been, in arithmetic statistics and analytic number theory, a long history of studying lower order terms [40, 44, 45, 38, 16, 46, 26, 25, 19, 17]. This is because, beyond their own inherent interest, proving the existence of secondary terms have a number of consequences. First, the implied improvement in the size of the error terms have many applications, for example to the study of associated families of L-functions (see for example [20, 42]). Next, understanding secondary terms is necessary for providing numerical evidence to support conjectures, since these secondary terms have a significant contribution in the height range in which we are able to perform computations. (For example, even in the height range $H(E) \sim 10^{12}$, the secondary term of $X^{3/4}$ is within a factor of 10 of the primary term of $X^{5/6}$.) Finally, secondary terms are of considerable theoretical interest. In the function field case for instance, primary terms are obtained via proving homological stability results. In [30, Problem 5], Venkatesh poses the question of what the topological significance of secondary terms are, and it is speculated in [6] that secondary terms might be related to secondary homological stability in the sense of [27].

As in [14], Theorem 1 is proven by exploiting the connection between the 2-Selmer groups of elliptic curves, and $\mathrm{GL}_2(\mathbb{Z})$-orbits on $V(\mathbb{Z}) = \mathrm{Sym}^4(\mathbb{Z}^2)$, the set of irreducible integral binary quartic forms. The action of $\mathrm{GL}_2(\mathbb{Z})$ on the set of integral binary *cubic* forms has been extensively studied. The symmetric cubed representation of $\mathrm{GL}_2(\mathbb{Q})$ is *prehomogeneous*, i.e., the action over $\overline{\mathbb{Q}}$ has a Zariski open orbit. As a consequence, the ring of polynomial invariants for the action of

$\mathrm{GL}_2(\mathbb{Z})$ on integral binary cubic forms is generated by a single element, namely the discriminant. A famous result of Davenport [23] develops and uses geometry-of-numbers tehniques to determine asymptotics (of size $\sim \zeta(2)X/3$) for the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits on the set of integral irreducible binary cubic forms having discriminant bounded by X.

Landmark work of Shintani [44] recovered Davenport's result, and in addition proved the existence of a secondary term of size $cX^{5/6}$ for a negative constant c. More precisely, combined with his later work [45] on his double zeta function, Shintani proved that the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits on integral irreducible binary cubic forms, having nonzero discriminant bounded by X, can be expressed as a sum of two main terms, growing like X and $X^{5/6}$, along with an error term of size $O_\epsilon(X^{2/3+\epsilon})$. This was accomplished by considering the (Shintani) zeta function constructed from the counting function of binary cubic forms, proving that this zeta function has a meromorphic continuation to $\mathbb{C}$, and analyzing the location and residues of the poles. In fact, general theory developed by Sato–Shintani [39] considers any prehomogeneous vector space of a reductive group whose singular set is an irreducible hypersurface, and proves that the associated zeta function has a meromorphic continuation to $\mathbb{C}$ under a certain condition. Moreover, they prove that the set of poles is contained within the set of zeros of the Bernstien–Sato polynomial associated to the invariant polynomial of this prehomogeneous representation. Thus in those prehomogeneous cases, there are natural guesses for the possible exponents of the lower order terms in the counting functions.

In our case, however, the symmetric fourth power representation of $\mathrm{GL}_2(\mathbb{Z})$ is not prehomogeneous. The ring of polynomial invariants for the action of $\mathrm{GL}_2(\mathbb{Z})$ on $V(\mathbb{Z})$ is generated by two elements, usually denoted I and J. Explicitly, for $f(x,y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$, we have

$$I(f) = 12ae - 3bd + c^2; \quad J(f) = 72ace + 9bcd - 27ad^2 - 27eb^2 - 2c^3.$$

The discriminant polynomial on binary quartic forms can be expressed in terms of I and J: we have $\Delta(f) := \Delta(I(f), J(f)) = (4I(f)^3 - J(f)^2)/27$. Throughout this paper, we order $\mathrm{GL}_2(\mathbb{Z})$-orbits on $V(\mathbb{Z})$ by *height* H, defined as

$$H(f) := H(I(f), J(f)) := \max\{|I(f)|^3, J(f)^2/4\}. \tag{1}$$

The primary term in the counting function of $\mathrm{GL}_2(\mathbb{Z})$-orbits on irreducible binary quartic forms, ordered by height, was obtained in [14, Theorem 1.6]. In this article, we prove the existence of a secondary term, generalizing Shintani's theorem to the setting of binary quartic forms. Before stating this result, we need some additional notation. For a pair $(I,J) \in \mathbb{R}^2$ with $\Delta(I,J) := (4I^3 - J^2)/27 \neq 0$, let $E^{I,J}$ be the elliptic curve over $\mathbb{R}$ given by the equation $y^2 = x^3 - (I/3)x - (J/27)$, define

$$\Omega(E^{I,J}) := \int_{\substack{(x,y)\in E^{I,J}(\mathbb{R}) \\ y>0}} \frac{dx}{y}; \qquad \widetilde{\Omega}(I,J) := \Omega(E^{I,J}) + \Omega(E^{I,-J}). \tag{2}$$

We then define the quantities $C_*^{\Delta>0}$ and $C_*^{\Delta<0}$, for $* \in \{5/6, 3/4\}$ to be

$$\begin{aligned} C_{5/6}^{\Delta>0} &:= \int_{\mathbb{R}^2_{H<1,\Delta>0}} dIdJ, & C_{5/6}^{\Delta<0} &:= \int_{\mathbb{R}^2_{H<1,\Delta<0}} dIdJ, \\ C_{3/4}^{\Delta>0} &:= \int_{\mathbb{R}^2_{H<1,\Delta>0}} \widetilde{\Omega}(I,J)dIdJ, & C_{3/4}^{\Delta<0} &:= \int_{\mathbb{R}^2_{H<1,\Delta<0}} \widetilde{\Omega}(I,J)dIdJ, \end{aligned} \tag{3}$$

where

$$\begin{aligned} \mathbb{R}^2_{H<1,\Delta>0} &:= \{(I,J) \in \mathbb{R}^2 \mid H(I,J) < 1, \Delta(I,J) > 0\}, \\ \mathbb{R}^2_{H<1,\Delta<0} &:= \{(I,J) \in \mathbb{R}^2 \mid H(I,J) < 1, \Delta(I,J) < 0\}. \end{aligned} \tag{4}$$

For $i \in \{0,1,2\}$, and any subset S of $V(\mathbb{R})$, let $S^{(i)}$ denote the set of elements in S having $4-2i$ real roots and i pairs of complex conjugate roots in $\mathbb{P}^1_{\mathbb{C}}$. We further partition $S^{(2)}$ into $S^{(2+)} \cup S^{(2-)}$, where $S^{(2+)}$ (resp. $S^{(2+)}$) consists of positive (resp. negative) definite forms. Finally, for $i \in \{0,1,2+,2-\}$, let $h^{(i)}(I,J)$ denote the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits on irreducible elements in $V(\mathbb{Z})^{(i)}$ having invariants I and J. Then we have the following result.

Theorem 2 *For $i \in \{0,1,2+,2-\}$, we have*

$$\sum_{\substack{(I,J)\in\mathbb{Z}^2 \\ H(I,J)<X}} h^{(i)}(I,J) = \frac{2\zeta(2)}{27\sigma_i} C^{\circ}_{5/6} \cdot X^{5/6} + \frac{\zeta(1/2)}{108\sigma_i} C^{\circ}_{3/4} \cdot X^{3/4} + O_\epsilon(X^{2/3+\epsilon}),$$

where $\sigma_0 = \sigma_{2\pm} = 4$, $\sigma_1 = 2$, and we take $\circ$ to be $\Delta > 0$ when $i \in \{0, 2\pm\}$ and $\Delta < 0$ when $i = 1$.

We note that the values of $C^{\circ}_{5/6}$ are easily computed: from [14, (22),(23)], we see that $C^{\Delta>0}_{5/6} = 8/5$ and $C^{\Delta<0}_{5/6} = 32/5$. Thus the primary terms of Theorem 2 agree with [14, Theorem 1.6]. We leave a numerical estimation of the secondary constant to future work, but note that since $\zeta(1/2) < 0$, the secondary term must be negative.

We note that Yukie (in [49]) introduced and analyzed a certain "global zeta integral" for the space of binary forms of general degree d. For $d = 4$, he showed that the zeta integral converges absolutely for $\Re(s) > 5/6$ and has a holomorphic continuation to the region $\Re(s) > 2/3$ except for simple poles at $s = 5/6$ and $s = 3/4$. We expect that Yukie's result is related to a smoothed version of Theorem 2.

Outline of the proofs

As described previously, asymptotics for the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits on integral irreducible binary cubic forms with bounded discriminant were first obtained by Davenport [23] using geometry-of-numbers techniques. Using zeta function methods, Shintani [44, 45] recovered these asymptotics, and also proved the existence of secondary main terms in these counting functions. Using a "slicing technique", Bhargava–Shankar–Tsimerman [16], obtained secondary terms for the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits on integral binary cubic forms, reproving Shintani's result. Combining sieving techniques developed by Belabas–Bhargava–Pomerance [5] with these two counting methods, Taniguchi–Thorne [46] and Bhargava–Shankar–Tsimerman [16] independently and simultaneously obtained secondary terms for the counting functions of cubic fields.

In our case, since the representation $V(\mathbb{Q}) = \mathrm{Sym}^4(\mathbb{Q}^2)$ of $\mathrm{GL}_2(\mathbb{Q})$ is not prehomogeneous, Shintani's methods are not available to us. However, the slicing technique generalizes in a straightforward manner and allows us to prove Theorem 2, which we do in the largely self contained §4. Unfortunately, the sieving techniques used in the cubic case fail for us at the very first step. This is why, previous to our work, even power saving error terms were not known for the counting function of $|\mathrm{Sel}_2(E)|$. The primary reason is that to prove Theorem 1, it is necessary for us to count $\mathrm{GL}_2(\mathbb{Z})$-orbits on $V(\mathbb{Z})$, with height bounded by X, where each orbit f is weighted by $\phi(f)$, where $\phi : V(\mathbb{Z}) \to \mathbb{R}$ is a $\mathrm{GL}_2(\mathbb{Z})$-invariant function. It is possible to write ϕ as a product over p of some functions $\phi_p : V(\mathbb{Z}_p) \to \mathbb{R}$. However, the functions ϕ_p are not defined modulo p^k for any integer k. This is quite in contrast to the cubic case, where the function analogous to ϕ_p is the characteristic function of the set of binary cubic forms corresponding to maximal orders over $\mathbb{Z}_p$. This set (and hence its characteristic function) is defined modulo p^2. Resolving this difficulty of ϕ_p not being a

periodic function requires many new ideas and tools, on both algebraic and analytic aspects of the arithmetic of binary quartic forms.

To begin, we develop approximation techniques in Section 5, in order to analyze these functions ϕ_p. We prove that they can be written as an infinite sum of functions $\phi_p^{(k)}$ for $k \geq 0$, where p is an odd prime, $\phi_p^{(k)}$ is defined modulo p^{2k} and supported on the set of binary quartic forms whose discriminants are divisible by p^{2k}. We prove something analogous (though slightly weaker) for $p = 2$ as well. We can then write

$$\phi(f) = \prod_p \phi_p(f) = \prod_p \sum_{k \geq 0} \phi_p^{(k)}(f) = \sum_{n \geq 1} \phi(n; f), \tag{5}$$

where $\phi(n; \cdot) : V(\mathbb{Z}) \to \mathbb{R}$ is given by $\phi(n; f) = \prod_{p^k \| n} \phi_p^{(k)}(f)$. Our results on $\phi_p^{(k)}$ imply that $\phi(n; \cdot)$ is defined modulo n^2, and its support is on a sparse set: namely the set of elements in $V(\mathbb{Z})$ whose discriminants are divisible by n^2 (up to absolutely bounded powers of 2).

To sum $\phi(n; \cdot)$ over $\mathrm{GL}_2(\mathbb{Z})$-orbits on $V(\mathbb{Z})$ with height bounded by X when n is small (i.e., $n \ll X^{1/12+\delta}$ for small positive δ), we use equidistribution methods. More specifically, we combine nontrivial bounds on orbital exponential sums on $V(\mathbb{Z}/p^2\mathbb{Z})$ with twisted Poisson summation. This allows us to carry out the analogous sum of $\phi(n; \cdot)$, when the $H(f) < X$ condition is replaced with a smooth approximation. We then prove that this approximation is good enough to carry out the weighted sharp sum.

Finally, it remains to show that the contribution from $n \gg X^{1/12+\delta}$ is negligible. To do this, it is necessary for us to prove uniformity (or tail) estimates on the number of $\mathrm{PGL}_2(\mathbb{Z})$-orbits on integral binary quartic forms of bounded height with discriminant divisible by n^2, on average over n. Moreover, unlike in the cubic case, we need these estimates for all n, not just squarefree n. In Section 6, we prove the following result, giving us a sufficiently strong uniformity estimate:

Theorem 3 *Let $\mathcal{W}_n \subset V(\mathbb{Z})$ denotes the set of irreducible integral binary quartic forms whose discriminants are divisible by n^2, and whose associated Galois group is S_4 or A_4. For positive real numbers X and M, we have*

$$\sum_{n \geq M} \#\Big\{ f \in \frac{\mathcal{W}_n}{\mathrm{PGL}_2(\mathbb{Z})} : H(f) < X \Big\} \ll_\epsilon \frac{X^{5/6+\epsilon}}{M} + X^{11/15+\epsilon}. \tag{6}$$

It is of course crucial for us that $11/15 = 3/4 - 1/60 < 3/4$. The primary term in the above uniformity estimate should be optimal (up to a factor of X^ϵ). Thus, we expect (6) to be optimal (up to a factor of X^ϵ) in the range $M \ll X^{1/10}$.

Theorem 3 is used to bound the contribution from $n \gg X^{1/12+\delta}$. Combining this with the methods and results described above to handle smaller n yields Theorem 1.

This paper is organized as follows. We begin by introducing some notation and stating versions of our main results for congruence families of elliptic curves and binary quartic forms in §2. In §3, we collect an assortment of parametrization results regarding elements in the 2-Selmer groups of elliptic curves and regarding quartic fields. We prove our quartic analogue of Shintani's theorem in §4, using the "slicing method" developed in [16]. We construct periodic approximations to the relevent local functions in §5. Then in §6, we obtain density estimates required to prove that sum describing our secondary term constant converges. We also obtain bounds on the Fourier transforms of various subsets of $V(\mathbb{Z}/n\mathbb{Z})$; this will be necessary to obtain equidistribution results. Our main uniformity estimate (Theorem 3) is proven in §7, and this is combined with the previous

results to prove all the main theorems in §8. Finally, we compute some local densities appearing in our secondary terms in the appendix.

Acknowledgments

AS was supported by an NSERC discovery grant, and is also supported by a Simons fellowship. TT was supported by JSPS KAKENHI Grant Number 22H01115. It is a pleasure to thank Manjul Bhargava, Dick Gross, Wei Ho, Ananth Shankar, and Jacob Tsimerman for many helpful conversations. The authors would like to thank Hiroyuki Ochiai for many useful comments.

2 Statements of results for congruence families

Versions of our main results, Theorems 1 and 2 also hold if we restrict to families of elliptic curves and binary quartic forms, respectively, satisfying congruence conditions modulo a fixed finite integer. In this section, we state these results, and also describe certain infinite sets of congruence conditions that we may impose on these families.

Let p be a prime number, and let $\phi : V(\mathbb{Z}_p) \to \mathbb{R}$ (resp. $\phi : \mathbb{Z}_p^2 \to \mathbb{R}$) be a function. For a positive integer k, we say that ϕ is *periodic with period* p^k if ϕ can be written as a composition of functions

$$V(\mathbb{Z}_p) \to V(\mathbb{Z}_p)/p^k V(\mathbb{Z}_p) = V(\mathbb{Z}/p^k\mathbb{Z}) \xrightarrow{\overline{\phi}} \mathbb{R} \quad (\text{resp. } \mathbb{Z}_p^2 \to \mathbb{Z}_p^2/p^k\mathbb{Z}_p^2 = (\mathbb{Z}/p^k\mathbb{Z})^2 \xrightarrow{\overline{\phi}} \mathbb{R}),$$

where the first map is reduction modulo p^k. We say that a function $\phi : V(\mathbb{Z}_p)\backslash\{\Delta = 0\} \to \mathbb{R}$ (resp. $\phi : \mathbb{Z}_p^2\backslash\{\Delta = 0\} \to \mathbb{R}$) is *well approximated by periodic functions* if for $k \geq 0$, there exist functions $\phi^{(k)} : V(\mathbb{Z}_p) \to \mathbb{R}$ (resp. $\phi : \mathbb{Z}_p^2 \to \mathbb{R}$) bounded by 1 in absolute value, with $\phi^{(0)}$ being identically 1, satisfying the following properties:

(a) For every element f in $V(\mathbb{Z}_p)\backslash\{\Delta = 0\}$ (resp. $\mathbb{Z}_p^2\backslash\{\Delta = 0\}$), we have

$$\phi(f) = \sum_{k=0}^{\infty} \phi^{(k)}(f) = 1 + \sum_{k=1}^{\infty} \phi^{(k)}(f),$$

where the convergence is absolute. In fact, Condition (c) implies that the sum is finite for each f.

(b) The function $\phi^{(k)}$ is periodic with period p^{2k}.

(c) The support of $\phi^{(k)}$ is contained within the set of elements $f \in V(\mathbb{Z}_p)$ with $p^{2k} \mid \Delta(f)$ (resp. $(I, J) \in \mathbb{Z}_p^2$ with $p^{2k} \mid \Delta(I, J)$).

If instead of Property (c), we only have that the functions $\phi^{(k)}$ are supported on elements f in $V(\mathbb{Z}_p)$ or $\mathbb{Z}_p^2$ with $p^{2k-O(1)} \mid \Delta(f)$, then we say that ϕ is *almost well approximated by periodic functions*. A bounded function $\phi : \mathbb{Z}^2 \to \mathbb{R}$ is said to be *large and locally well approximated* if for every p there exist bounded functions $\phi_p : \mathbb{Z}_p^2\backslash\{\Delta = 0\} \to \mathbb{R}$ such that $\phi(I, J) = \prod_p \phi_p(I, J)$ for every $(I, J) \in \mathbb{Z}^2\backslash\{\Delta = 0\}$, the functions ϕ_p are all almost well approximated, and for all large enough p, the functions ϕ_p are well approximated.

We will need an analogous notion not only for functions from $V(\mathbb{Z}) \to \mathbb{R}$ but for $\mathrm{PGL}_2(\mathbb{Z})$-*invariant* functions. Here, for a ring R, the action of $\mathrm{PGL}_2(R)$ on $V(R)$ comes from descending the following twisted action of $\mathrm{GL}_2(R)$ on $V(R)$: for $\gamma \in \mathrm{GL}_2(R)$ and $f \in V(R)$, define

$$\gamma \cdot f(x,y) := f((x,y)\cdot\gamma)/(\det(\gamma)^2). \tag{7}$$

(Note that over $\mathbb{Z}$, the twisted action and the ordinary action coincide.) We say that a function $\psi : V(\mathbb{Z}/p^{2k}\mathbb{Z}) \to \mathbb{R}$ is *strongly invariant* if ψ is $\mathrm{PGL}_2(\mathbb{Z}/p^{2k}\mathbb{Z})$-invariant, and $\psi(t^2 f) = \psi(f)$ for $t \in (\mathbb{Z}/p^{2k}\mathbb{Z})^\times$. Then a $\mathrm{PGL}_2(\mathbb{Z})$-invariant function $\phi : V(\mathbb{Z}) \to \mathbb{R}$ is said to be *large and locally well approximated* if there exist almost well approximated functions $\phi_p : V(\mathbb{Z}_p)\backslash\{\Delta = 0\} \to \mathbb{R}$ for each p, such that $\phi(f) = \prod_p \phi_p(f)$ for every $f \in V(\mathbb{Z})\backslash\{\Delta = 0\}$, and for all large enough p, the functions ϕ_p are well approximated via strongly invariant periodic functions $\phi_p^{(k)}$.

Remark 4 A periodic function $\phi : V(\mathbb{Z}_p) \to \mathbb{R}$ is always almost well approximated by periodic functions. Thus, every periodic $\mathrm{PGL}_2(\mathbb{Z})$-invariant function $\phi : V(\mathbb{Z}) \to \mathbb{R}$ (i.e., a lift to $V(\mathbb{Z})$ of a $\mathrm{PGL}_2(\mathbb{Z}/n\mathbb{Z})$-invariant function $V(\mathbb{Z}/n\mathbb{Z}) \to \mathbb{R}$ for some integer n) is large and locally well approximated. Moreover, it is easy to check that the characteristic functions of many other natural sets in $V(\mathbb{Z})$ are large and well approximated. For example, the set of f having squarefree discriminant, and the set of f such that the associated elliptic curve is semistable.

The following result is a congruence version of Theorem 1.

Theorem 5 *Let $\phi : \mathbb{Z}^2 \to \mathbb{Z}$ be large and locally well approximated.*

$$\sum_{E_{AB}\in\mathcal{E}_X^{\pm}} |\mathrm{Sel}_2(E)|\phi(A,B) = 3\Big(\sum_{E_{AB}\in\mathcal{E}_X^{\pm}} \phi(A,B)\Big) + C(\mathcal{E};\phi)^{\pm}X^{3/4} + O_\epsilon(X^{3/4-\alpha+\epsilon}),$$

for constants $C(\mathcal{E};\phi)^{\pm}$ and the same α as in Theorem 1.

In the theorem above we assume that ϕ is fixed (and suppress the dependence on ϕ in the error term).

We note that the characteristic function of the set of pairs (A,B) such that E_{AB} is a semistable elliptic curve (resp. has squarefree discriminant) is large and locally well approximated. Hence, in addition to families of elliptic curves whose coefficients satisfy finitely many congruence conditions, the above result also includes natural families of elliptic curves defined by the imposition of infinitely many congruence conditions.

We next describe two generalizations of Theorem 2. In our first result, we impose finitely many *splitting conditions* on the set of binary quartic forms being counted. An element f of $V(\mathbb{Z})$, $V(\mathbb{Z}_p)$, of $V(\mathbb{F}_p)$ is said to have *splitting type* $(f_1^{e_1}\dots f_k^{e_k})$ if the reduction of f modulo p splits as the product of irreducible polynomials of degrees f_i raised to the powers e_i. As convention, we will define the splitting type of the zero form in $V(\mathbb{F}_p)$ by (0). For nonzero $\sigma_p = (f_1^{e_1}\dots f_k^{e_k})$, we define its index by $\mathrm{ind}(\sigma_p) := \sum(e_i - 1)f_i$. The splitting types that arise this way are called *quartic splitting types*. Let S be a finite set of primes, and for each $p \in S$, let σ_p be a quartic splitting type. We denote this collection of splitting types $\{\sigma_p\}_{p\in S}$ by Σ. For $i \in \{0,1,2+,2-\}$ and $(I,J) \in \mathbb{Z}^2$, we let $h_\Sigma^{(i)}(I,J)$ denote the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits on irreducible elements $f \in V(\mathbb{Z})$, having invariants I and J, such that $f(x,y)$ has $r-i$ real roots, i pairs of complex conjugate roots, and such that for all $p \in S$, the splitting type of f at p is σ_p. Then we have the following result.

Theorem 6 *Let notation be as above. For $i \in \{0, 1, 2+, 2-\}$, we have*

$$\begin{aligned}\sum_{\substack{(I,J)\in\mathbb{Z}^2\\ H(I,J)<X}} h_\Sigma^{(i)}(I,J) &= \frac{2\zeta(2)}{27\sigma_i}\kappa_{5/6}(\Sigma)C_{5/6}^{\circ}\cdot X^{5/6} + \frac{\zeta(1/2)}{108\sigma_i}\kappa_{3/4}(\Sigma)C_{3/4}^{\circ}\cdot X^{3/4}\\ &\quad + O_\epsilon\left(\ell(\Sigma)X^{2/3+\epsilon} + \ell'(\Sigma)X^{1/2+\epsilon}\right),\end{aligned}$$

where we take $\circ$ to be $\Delta > 0$ when $i \in \{0, 2+, 2-\}$ and $\Delta < 0$ when $i = 1$, and we have $\kappa_*(\Sigma) = \prod_{p\in S}\kappa_*(\sigma_p)$, $\ell(\Sigma) = \prod_p \ell_p(\sigma_p)$, and $\ell'(\Sigma) = \prod_p \ell'_p(\sigma_p)$. The values $\kappa_{5/6}(\sigma_p)$ are listed in Table 1, while the values $\kappa_{3/4}(\sigma_p)$ can be computed from Table 2. We have $\ell_p(\sigma_p) = p^{1-\mathrm{ind}_p(\sigma_p)}$ and $\ell'_p(\sigma_p) = p^{2-\mathrm{ind}_p(\sigma_p)}$, for nonzero splitting types σ_p. We take $\ell_p(0) = -4$ and $\ell'_p(0) = -3$.

We note that Theorem 6 can be easily generalized to collections $\Sigma = (\Sigma_p)_p$, where Σ_p is a set of quartic splitting types (not just a singleton set). Moreover, for some natural choices for Σ_p, the error term can be easily improved. For example, when $\Sigma_p = \Sigma_p^{\mathrm{ur}}$ is the set of all unramified splitting types, the error can be improved to the error coming from the case when $\Sigma_p = \Sigma_p^{\mathrm{ram}}$ is the set of all ramified splitting types. This is because the count for Σ_p^{ur} is the difference between the count when there is no condition on p, and the count for Σ_p^{ram}.

Our next result involves summing a large and locally well approximated function $\phi : V(\mathbb{Z}) \to \mathbb{R}$ over $\mathrm{PGL}_2(\mathbb{Z})$-orbits on irreducible integral binary quartic forms. To state it, we need some additional notation. First let $\nu(\phi)$ denote the *density* of ϕ and, for $a \in \mathbb{Z}$, let $\nu_a(\phi)$ denote the density of ϕ restricted to $V_a(\mathbb{Z})$, the set of binary quartic forms whose x^4-coefficients are a. That is, we define

$$\nu(\phi) := \int_{V(\widehat{\mathbb{Z}})} \phi(v)dv; \qquad \nu_a(\phi) := \int_{V_a(\widehat{\mathbb{Z}})} \phi(v)dv. \tag{8}$$

Above, the measures dv are normalized so that $V(\widehat{\mathbb{Z}})$ and $V_a(\widehat{\mathbb{Z}})$ have volume 1. Next, define the Dirichlet series

$$D^{\pm}(\phi, s) := \sum_{a>0} \frac{\nu_{\pm a}(\phi)}{a^s}. \tag{9}$$

We will prove in §6 that $D^{\pm}(\phi, s)$ has an analytic continuation to the region $\Re(s) > 1/3$, with at most a simple pole at $s = 1$. In particular, the value $D^{\pm}(\phi, 1/2)$ is well defined. For $i \in \{0, 1, 2+, 2-\}$ and $(I, J) \in \mathbb{Z}^2$, we let $h_\phi^{(i)}(I, J)$ denote the sum of $\phi(f)$ over $\mathrm{GL}_2(\mathbb{Z})$-orbits on irreducible elements $f \in V(\mathbb{Z})$, having invariants I and J, such that $f(x, y)$ has $r - i$ real roots and i pairs of complex conjugate roots. We have the following result.

Theorem 7 *Let $\phi : V(\mathbb{Z})\backslash\{\Delta \neq 0\} \to \mathbb{R}$ be a large and locally well approximated function. For $i \in \{0, 1, 2+, 2-\}$, let $V(\mathbb{Z})^{(i)}$ (resp. $V(\mathbb{R})^{(i)}$) denote the set of elements $f \in V(\mathbb{Z})$ (resp. $f \in V(\mathbb{R})$) having nonzero discriminant and exactly $4 - 2i$ real roots in $\mathbb{P}^1_{\mathbb{R}}$. Then for any positive constant α*

acceptable in Theorem 1, we have

$$\sum_{\substack{(I,J)\in\mathbb{Z}^2\\ H(I,J)<X}} h_\phi^{(0)}(I,J) \;=\; \nu(\phi)\frac{2\zeta(2)}{27\sigma_0}C_{5/6}^{\Delta>0}X^{5/6} \;+\; \frac{D^+(\phi,\frac{1}{2})+D^-(\phi,\frac{1}{2})}{216\sigma_0}C_{3/4}^{\Delta>0}X^{3/4} \;+\; O_\epsilon(X^{3/4-\alpha+\epsilon});$$

$$\sum_{\substack{(I,J)\in\mathbb{Z}^2\\ H(I,J)<X}} h_\phi^{(1)}(I,J) \;=\; \nu(\phi)\frac{2\zeta(2)}{27\sigma_1}C_{5/6}^{\Delta<0}X^{5/6} \;+\; \frac{D^+(\phi,\frac{1}{2})+D^-(\phi,\frac{1}{2})}{216\sigma_1}C_{3/4}^{\Delta<0}X^{3/4} \;+\; O_\epsilon(X^{3/4-\alpha+\epsilon});$$

$$\sum_{\substack{(I,J)\in\mathbb{Z}^2\\ H(I,J)<X}} h_\phi^{(2\pm)}(I,J) \;=\; \nu(\phi)\frac{2\zeta(2)}{27\sigma_2}C_{5/6}^{\Delta>0}X^{5/6} \;+\; \frac{D^\pm(\phi,\frac{1}{2})}{108\sigma_2}C_{3/4}^{\Delta>0}X^{3/4} \;+\; O_\epsilon(X^{3/4-\alpha+\epsilon}).$$

In the theorem above we assume that ϕ is fixed (and suppress the dependence on ϕ in the error term).

The average sizes of the 2-torsion in the class groups $\mathrm{Cl}(K)$ and narrow class groups $\mathrm{Cl}^+(K)$ of monogenic cubic fields (K,α) ordered by height are determined in [12]. The above result would imply that this bound of 2-torsion in these class groups admits a secondary term growing as $X^{3/4}$.

Suppose that a large and locally well approximated function $\phi=\prod_p\phi_p$ is such that $\phi_p: V(\mathbb{Z}_p)\to\mathbb{R}$ is invariant under multiplication by units in $\mathbb{Z}_p$. This is the case, for instance, for the functions ϕ arising by imposing splitting conditions on the family of binary quartic forms, for the characteristic function of elements in $V(\mathbb{Z})$ with squarefree discriminant, and the function ϕ corresponding to the 2-torsion in the class groups of monogenic cubic fields. It is easy to see that $D^+(\phi,s)=D^-(\phi,s)$ for such functions ϕ, giving a uniform description of the leading second order constants for the four possible splitting types at infinity. We will also see in the appendix that for such functions ϕ, the Dirichlet series $D^\pm(\phi,s)$ admits an Euler product. However, we do not believe this to be true of general functions ϕ. Finally, we note that the characteristic function of the set of *soluble* binary quartic forms over $\mathbb{Z}_p$ is not invariant under multiplication by units. This is the main reason our secondary term constants of Theorems 1 and 5 are not explicit.

3 Parametrization results

In this section, we collect an assortment of results parametrizing various arithmetic objects using (weighted) group orbits on lattices. Specifically, in §3.1, we parametrize elements in the 2-Selmer group of elliptic curves, following work of Birch and Swinnerton-Dyer [18], Cremona [21], and Bhargava and the first named author [14]. Then in §3.2, we use results of Bhargava [8] and Wood [48] to parametrize quartic rings along with a monogenized cubic resolvent ring. We combine this latter parametrization with a construction of Heilbronn [31] to parametrize 2-torsion elements in the dual of the class group of monogenized cubic fields.

3.1 The 2-Selmer groups of elliptic curves

For $E=E_{AB}\in\mathcal{E}$, we define the following two invariants of E:

$$I(E):=-2^4\cdot 3\cdot A; \qquad J(E):=-2^6\cdot 3^3\cdot B.$$

We denote the elliptic curve having invariants I and J by E^{IJ}; note that the height of E^{IJ} is given by $H(E^{IJ})=H(I,J)/(2^{10}3^3)$. We recall the following two results, proved originally by Birch–Swinnerton-Dyer [18] and further refined by Cremona [21], regarding the connection between the

2-Selmer groups of elliptic curves and PGL_2-orbits on V. The results are stated in the notation in [14, §3.1] (see [14, Theorem 3.2 and Proposition 3.3]). A binary quartic form $f(x,y)$ with coefficients in a field K is said to be *K-soluble* if the equation $z^2 = f(x,y)$ has a nontrivial solution. A binary quartic form $f(x,y)$ in $V(\mathbb{Q})$ is said to be *locally soluble* if f is soluble over $\mathbb{R}$ and over $\mathbb{Q}_p$ for every prime p.

Theorem 3.1 *Let K be a field having characteristic not 2 or 3. Let $E : y^2 = x^3 - \frac{I}{3}x - \frac{J}{27}$ be an elliptic curve over K. Then there exists a bijection between elements in $E(K)/2E(K)$ and $\mathrm{PGL}_2(K)$-orbits of K-soluble binary quartic forms having invariants equal to I and J, given by*

$$(\xi,\eta) + 2E(K) \mapsto \mathrm{PGL}_2(K) \cdot \left(\frac{1}{4}x^4 - \frac{3}{2}\xi x^2y^2 + 2\eta xy^3 + \left(\frac{I}{3} - \frac{3}{4}\xi^2\right)y^4\right).$$

Under this bijection, the identity element in $E(K)/2E(K)$ corresponds to the $\mathrm{PGL}_2(K)$-orbit of binary quartic forms having a linear factor over K.

Furthermore, the stabilizer in $\mathrm{PGL}_2(K)$ of any (not necessarily K-soluble) binary quartic form f in V_K, having nonzero discriminant and invariants I and J, is isomorphic to $E(K)[2]$, where E is the elliptic curve defined by $y^2 = x^3 - \frac{I}{3}x - \frac{J}{27}$.

This leads to the following parametrization of the 2-Selmer groups of elliptic curves over $\mathbb{Q}$.

Theorem 3.2 *Let $E = E^{IJ}$ be an elliptic curve over $\mathbb{Q}$. Then the elements of the 2-Selmer group of E are in one-to-one correspondence with $\mathrm{PGL}_2(\mathbb{Q})$-equivalence classes of locally soluble integral binary quartic forms having invariants equal to I and J.*

Furthermore, the set of integral binary quartic forms that have a rational linear factor and invariants equal to I and J lie in a single $\mathrm{PGL}_2(\mathbb{Q})$-equivalence class, and this class corresponds to the identity element in the 2-Selmer group of E.

Next, we express the number of $\mathrm{PGL}_2(\mathbb{Q})$-equivalence classes of locally soluble integral binary quartic forms having fixed invariants as the weighted count of $\mathrm{PGL}_2(\mathbb{Z})$-orbits on integral binary quartic forms. Moreover, we show that this weight is *local*, i.e., it can be expressed as a product of weights defined over $\mathbb{Z}_p$. The condition of local solubility is clearly local: let $\ell_p : V(\mathbb{Z}_p) \to \mathbb{R}$ be the characteristic function of the set of elements in $V(\mathbb{Z}_p)$ that are soluble over $\mathbb{Q}_p$, and let $\ell : V(\mathbb{Z}) \to \mathbb{R}$ be given by $\ell(f) := \prod \ell_p(f)$. That is, ℓ is the characteristic function of locally soluble elements in $V(\mathbb{Z})$.

Given a binary quartic form $f \in V(\mathbb{Z})$ (resp. $f \in V(\mathbb{Z}_p)$ for some prime p), let $B(f)$ (resp. $B_p(f)$) denote a set of representatives for the action of $\mathrm{PGL}_2(\mathbb{Z})$ (resp. $\mathrm{PGL}_2(\mathbb{Z}_p)$) on the $\mathrm{PGL}_2(\mathbb{Q})$-equivalence class of f (resp. $\mathrm{PGL}_2(\mathbb{Q}_p)$-equivalence class of f) in $V(\mathbb{Z}_p)$. For an integral binary quartic form f in $V(\mathbb{Z})$ or $V(\mathbb{Z}_p)$, respectively define

$$\mathrm{Aut}_{\mathbb{Q}}(f) := \mathrm{Stab}_{\mathrm{PGL}_2(\mathbb{Q})}(f); \qquad \mathrm{Aut}_{\mathbb{Z}}(f) := \mathrm{Stab}_{\mathrm{PGL}_2(\mathbb{Z})}(f);$$
$$\mathrm{Aut}_{\mathbb{Q}_p}(f) := \mathrm{Stab}_{\mathrm{PGL}_2(\mathbb{Q}_p)}(f); \qquad \mathrm{Aut}_{\mathbb{Z}_p}(f) := \mathrm{Stab}_{\mathrm{PGL}_2(\mathbb{Z}_p)}(f).$$

We define the *global weight* $m(f)$ for f in $V(\mathbb{Z})$ and the *local weight* $m_p(f)$ for f in $V(\mathbb{Z}_p)$ to be:

$$\begin{aligned} m(f) &:= \sum_{f' \in B(f)} \frac{\#\mathrm{Aut}_{\mathbb{Q}}(f')}{\#\mathrm{Aut}_{\mathbb{Z}}(f')} = \sum_{f' \in B(f)} \frac{\#\mathrm{Aut}_{\mathbb{Q}}(f)}{\#\mathrm{Aut}_{\mathbb{Z}}(f')}; \\ m_p(f) &:= \sum_{f' \in B_p(f)} \frac{\#\mathrm{Aut}_{\mathbb{Q}_p}(f')}{\#\mathrm{Aut}_{\mathbb{Z}_p}(f')} = \sum_{f' \in B_p(f)} \frac{\#\mathrm{Aut}_{\mathbb{Q}_p}(f)}{\#\mathrm{Aut}_{\mathbb{Z}_p}(f')}. \end{aligned} \tag{10}$$

Moreover, given $f \in V(\mathbb{Z})$ (resp. $f \in V(\mathbb{Z}_p)$), we define $\mathrm{PGL}_2(\mathbb{Q})_f$ (resp. $\mathrm{PGL}_2(\mathbb{Q}_p)_f$) to be the set of elements $\gamma \in \mathrm{PGL}_2(\mathbb{Q})$ (resp. $\gamma \in \mathrm{PGL}_2(\mathbb{Q}_p)$), such that $\gamma \cdot f$ belongs to $V(\mathbb{Z})$ (resp. $V(\mathbb{Z}_p)$). Then the following result follows from [14, Proposition 3.6] and its proof.

Proposition 3.3 *For elements f in $V(\mathbb{Z})$ and $V(\mathbb{Z}_p)$, we have the equalities*

$$m(f) = \#[\mathrm{PGL}_2(\mathbb{Z})\backslash \mathrm{PGL}_2(\mathbb{Q})_f], \quad m_p(f) = \#[\mathrm{PGL}_2(\mathbb{Z}_p)\backslash \mathrm{PGL}_2(\mathbb{Q}_p)_f],$$

respectively. Moreover, we have

$$m(f) = \prod_p m_p(f)$$

for every $f \in V(\mathbb{Z})$.

Suppose an elliptic curve E^{IJ} over $\mathbb{Q}$ is such that $E^{IJ}(\mathbb{Q})[2] = \{0\}$. Then it follows by Theorem 3.1 that if $f \in V(\mathbb{Q})$ has invariants I and J, then $\mathrm{Aut}_{\mathbb{Q}}(f) = \mathrm{Aut}_{\mathbb{Z}}(f) = \{\mathrm{id}\}$. As a consequence, $m(f) = \#B(f)$. First, for integers I and J, let $V(\mathbb{Z})_{IJ}$ denote the set of elements in $V(\mathbb{Z})$ with invariants I and J. An element in $V(\mathbb{Q})$ is said to be *generic* if it is irreducible and its corresponding elliptic curves has trivial 2-torsion. For any set $S \subset V(\mathbb{Q})$ let S^{gen} denote the set of generic elements in S. Then Theorem 3.2 immediately implies the following result.

Theorem 3.4 *Let E be an elliptic curve over $\mathbb{Q}$ with trivial 2-torsion and invariants I and J. Then we have*

$$|\mathrm{Sel}_2(E)| = 1 + \sum_{f \in \frac{V(\mathbb{Z})^{\mathrm{gen}}_{IJ}}{\mathrm{PGL}_2(\mathbb{Z})}} \frac{\ell(f)}{m(f)}.$$

3.2 Quartic rings with monogenic cubic resolvent rings

In this section, we recall results that parametrize certain cubic and quartic rings over $\mathbb{Z}$. We begin with the following parametrization of cubic rings due to Levi [35], Delone–Faddeev [24], and Gan-Gross-Savin [28]. Let $U = \mathrm{Sym}^3(\mathbb{Z}^2)$ denote the space of binary cubic forms. The group GL_2 acts on U via the following twisted action: $\gamma \cdot f(x, y) := f((x, y)\cdot\gamma)/\det(\gamma)$. Then we have the following result.

Proposition 3.5 ([35],[24],[28]) *There is a natural bijection between the set of cubic rings over $\mathbb{Z}$ upto isomorphism and the set of $\mathrm{GL}_2(\mathbb{Z})$-orbits on binary cubic forms.*

In fact, this bijection has a very explicit description. Let C be a cubic ring over $\mathbb{Z}$, and let $L := C/\mathbb{Z}$ be the associated two dimensional lattice. Consider the *index form* $\mathrm{ind} : L \to \mathbb{Z}$ which sents $\alpha \in L$, to the signed index of $\mathbb{Z}[\alpha]$ in C. More precisely, this is defined by $\mathrm{ind}\colon L \ni \alpha \mapsto \alpha \wedge \alpha^2 \in \wedge^2 L$. Since $L \cong \mathbb{Z}^2$ and ind is cubic, this yields an integral binary cubic form $f \in U(\mathbb{Z})$ upto the action of $\mathrm{GL}_2(\mathbb{Z})$. A single element $f \in U(\mathbb{Z})$ corresponds to a cubic ring C along with a basis $\{\alpha_1, \alpha_2\}$ of $C/\mathbb{Z}$. These can be lifted in a unique way to two elements β_1 and β_1 of C whose traces belong to $\{-1, 0, 1\}$, and it is then true that $\{1, \beta_1, \beta_2\}$ is a basis for C as a $\mathbb{Z}$-module.

Next, let $W = 2 \otimes \mathrm{Sym}^2(3)$ denote the space of pairs of ternary quadratic forms. There is a natural action of the group $\mathrm{GL}_2 \times \mathrm{SL}_3$ on W. We represent elements in $W(\mathbb{Z})$ by pairs (A, B) of 3×3 symmetric matrices with coefficients $\frac{1}{2}a_{ij}$ and $\frac{1}{2}b_{ij}$ with $1 \le i \le j \le 3$, where $a_{ii}, b_{ii} \in 2\mathbb{Z}$ and $a_{ij}, b_{ij} \in \mathbb{Z}$ for $i \ne j$. Then the following landmark result of Bhargava [8, Theorem 1] parametrizes pairs (Q, C), where Q is a quartic ring and C is a cubic resolvent ring of Q.

Theorem 3.6 ([8]) *There is a canonical bijection between the set of* $\mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$*-orbits on* $W(\mathbb{Z})$ *and the set of isomorphism classes of pairs* (Q, C)*, where* Q *is a quartic ring and* C *is a cubic resolvent ring of* Q*.*

There is a natural *resolvent* map $\mathrm{Res} : W \to U$ given by $\mathrm{Res}(A, B) := 4\det(Ax - By)$. This map is SL_3-invariant and GL_2-covariant. Moreover, if $(A, B) \in W(\mathbb{Z})$ corresponds to the pair (Q, C) of rings, the binary cubic form corresponding to C is $\mathrm{Res}(A, B) \in U(\mathbb{Z})$.

Let C be a rank-n ring over $\mathbb{Z}$. An element $\alpha \in C$ is said to be a *monogenizer* for C if $C = \mathbb{Z}[\alpha]$. We then say that the pair (C, α) over $\mathbb{Z}$ is an (n-ic) *monogenized rank-n ring*. Two monogenized rings (C, α) and (C', α') are said to be *isomorphic* if there is an isomorphism from C to C' such that it sends α to $\alpha' + n$ for some $n \in \mathbb{Z}$. For a ring R, let $U_1(R)$ denote the set $x^3 + tx^2 + sx + r$, where $r, s, t \in R$. There is a natural action of R on $U_1(R)$ given by $(r \cdot f)(x) := f(x + r)$ for $r \in R$ and $f \in U_1(R)$. Then we have the following result.

Proposition 3.7 *The map* $f(x) \mapsto (\mathbb{Z}[x]/(f(x)), \overline{x})$ *gives a bijection between the set of* $\mathbb{Z}$*-orbits on* $U_1(\mathbb{Z})$ *and the set of pairs* (C, α)*, where* C *is a cubic ring and* $\alpha \in C/\mathbb{Z}$ *is a monogenizer for* C*.*

Let Q be a quartic ring and let C be a monogenized cubic resolvent of Q. Then there exists a monic binary cubic form representing C. Hence, there exists a pair $(A, B) \in W(\mathbb{Z})$, corresponding to (Q, C), such that $\det(A) = 1/4$. The set of integral symmetric 3×3-matrices with determinant $1/4$ forms a single $\mathrm{SL}_3(\mathbb{Z})$-orbit. Hence, the pair (Q, C) can be represented by $(A_0, B) \in W(\mathbb{Z})$, wher A_0 is an anti-diagonal 3×3 matrix with coefficients $1/2$, -1, and $1/2$. Wood [48] considers the following embedding of $V(\mathbb{Z})$ into $W(\mathbb{Z})$:

$$\iota : ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4 \mapsto \left[\begin{pmatrix} & & \frac{1}{2} \\ & -1 & \\ \frac{1}{2} & & \end{pmatrix}, \begin{pmatrix} a & \frac{b}{2} & \\ \frac{b}{2} & c & \frac{d}{2} \\ & \frac{d}{2} & e \end{pmatrix} \right]. \tag{11}$$

The special orthogonal group SO_{A_0} is naturally isomorphic to PGL_2. Moreover, the map ι respects the two actions of $\mathrm{PGL}_2(\mathbb{Z})$ on $V(\mathbb{Z})$ and of $\mathrm{SO}_{A_0}(\mathbb{Z})$ on the set of pairs (A_0, B) in $W(\mathbb{Z})$. Combining the construction of ι with Theorem 3.6 and Proposition 3.7, Wood proves the following result.

Theorem 3.8 *There is a natural bijection between the set of* $\mathrm{PGL}_2(\mathbb{Z})$*-orbits on* $V(\mathbb{Z})$ *and the set of isomorphism classes of pairs* (Q, C, α)*, where* Q *is a quartic ring and* (C, α) *is a monogenized cubic resolvent ring of* Q*.*

Given a binary quartic form $f(x, y)$ in $V(\mathbb{Z})$ or $V(\mathbb{Z}_p)$ for some prime p, we let Q_f and C_f denote the quartic ring and monogenic cubic ring associated to f. Note that C_f only depends on the invariants I and J of f since the coefficients of $\mathrm{Res}(\iota(f))$ are invariants for the action of PGL_2 on V. We denote this cubic ring by R_{IJ}.

4 Secondary terms in the count of integral binary quartic forms

In this section, we determine secondary terms for the counting function of $\mathrm{PGL}_2(\mathbb{Z})$-orbits on generic elements of $V(\mathbb{Z})$, where each orbit is weighted by ϕ, a $\mathrm{PGL}_2(\mathbb{Z})$-invariant periodic function on $V(\mathbb{Z})$. First, in §4.1, we set up notation, and introduce the averaging method from [14]. In §4.2, we give bounds for the number of nongeneric $\mathrm{PGL}_2(\mathbb{Z})$-orbits on integral binary quartic forms with

bounded height and nonzero discriminants. A bound of $O(X^{3/4+\epsilon})$ was proved in [14, Lemma 2.4], but in order to recover a secondary term, we need to improve this. Then in §4.3, we use the "slicing method" developed in [16] to obtain primary and secondary terms for our counting functions and complete the proofs of Theorems 2 and 6 (modulo a volume computation, carried out in §8).

4.1 Preliminaries

Recall that we say an element $f \in V(\mathbb{Q})$ is *generic* if f is irreducible over $\mathbb{Q}$ and the corresponding elliptic curve has trivial 2-torsion. The latter condition is equivalent to the stabilizer of f in $\mathrm{PGL}_2(\mathbb{Q})$ being trivial. By Theorem 3.1, it follows that f is generic if and only if both f and the cubic resolvent polynomial of f are irreducible.

We partition the set of elements in $V(\mathbb{R})$ with nonzero discriminant into the following sets:

$$V(\mathbb{R})\backslash\{\Delta = 0\} = V(\mathbb{R})^{(0)} \cup V(\mathbb{R})^{(1)} \cup V(\mathbb{R})^{(2)} = V(\mathbb{R})^{(0)} \cup V(\mathbb{R})^{(1)} \cup V(\mathbb{R})^{(2+)} \cup V(\mathbb{R})^{(2-)}.$$

Above, for $i \in \{0,1,2\}$, the set $V(\mathbb{R})^{(i)}$ consists of elements $f \in V(\mathbb{R})$ having $4-2i$ distinct real roots in $\mathbb{P}^1(\mathbb{R})$ and i pairs of distinct complex conjugate roots in $\mathbb{P}^1(\mathbb{C})$. The set $V(\mathbb{R})^{(2)}$ consists of definite binary quartic forms, and we further partition it into the set of positive definite binary quartics $V(\mathbb{R})^{(2+)}$ and the set of negative definite binary quartics $V(\mathbb{R})^{(2-)}$. For a set $S \subset V(\mathbb{R})$ and $i \in \{0,1,2+,2-\}$, we define $S^{(i)}$ to be $S \cap V(\mathbb{R})^{(i)}$.

For a $\mathrm{PGL}_2(\mathbb{Z})$-invariant function $\phi : V(\mathbb{Z}) \to \mathbb{R}$, and $i \in \{0,1,2+,2-\}$, we define the counting function $N^{(i)}(\phi; X)$ to be

$$N^{(i)}(\phi; X) := \sum_{\substack{f \in \mathrm{PGL}_2(\mathbb{Z})\backslash V(\mathbb{Z})^{(i),\mathrm{gen}} \\ H(f)<X}} \phi(f).$$

We start with a slightly modified treatment of [14, §2.1,2.3], and provide fundamental domains for the action of $\mathrm{PGL}_2(\mathbb{Z})$ on $V(\mathbb{R})$. The only difference is that we work with the action of PGL_2 on V, while [14, §2] considers the action of GL_2 on V. We fix $i \in \{0,1,2+,2-\}$, and note that $V(\mathbb{R})^{(i)}$ contains only points with positive discriminant when $i \in \{0,2+,2-\}$ and only points with negative discriminant when $i = 1$. Given any pair (I,J) with $\Delta(I,J) > 0$ (resp. $\Delta(I,J) < 0$), the set of elements in $V(\mathbb{R})^{(i)}$, for $i \in \{0,2+,2-\}$ (resp. $i = 1$), having invariants I and J consists of a single $\mathrm{PGL}_2(\mathbb{R})$-orbit. Fundamental sets $L^{(i)}$ for the action of $\mathrm{PGL}_2(\mathbb{R})$ on the set of elements in $V(\mathbb{R})^{(i)}$ having height 1 are given in [14, Table 1]. Let $R_\infty^{(i)}$ denote the set $\mathbb{R}_{>0} \cdot L^{(i)}$. Then the sets $R_\infty^{(i)}$ are fundamental sets for the action of $\mathrm{PGL}_2(\mathbb{R})$ on $V(\mathbb{R})^{(i)}$. Moreover $R_\infty^{(i)}$ contains exactly one element with invariants I and J with $\Delta(I,J) > 0$ when $i \in \{0,2+,2-\}$ and exactly one element with invariants I and J with $\Delta(I,J) < 0$ when $i = 1$.

Let $\mathcal{F}$ be Gauss' fundamental domain for the action of $\mathrm{PGL}_2(\mathbb{Z})$ on $\mathrm{PGL}_2(\mathbb{R})$. Explicitly, we write $\mathcal{F} = \{u(t)k : u \in N(t),\ (t) \in A',\ k \in K\}$, where

$$N(t) := \left\{ u := \begin{pmatrix} 1 & \\ u & 1 \end{pmatrix} : u \in \mathrm{i}(t) \right\}, \quad A' := \left\{ (t) := \begin{pmatrix} t^{-1} & \\ & t \end{pmatrix} : t \geq \frac{\sqrt[4]{3}}{\sqrt{2}} \right\}, \quad K := \mathrm{SO}_2(\mathbb{R}).$$

Above $\mathrm{i}(t)$ is a subset of $[-1/2, 1/2]$ depending on t, and is all of $[-1/2, 1/2]$ when $t \geq 1$. For elements $u \in N(t)$ and $(t) \in A'$, we denote the product $u(t)$ by (u,t).

Let G_0 be a nonempty semialgebraic left K-invariant bounded open subset of $\mathrm{PGL}_2(\mathbb{R})$. Denote the set of elements in $R_\infty^{(i)}$ (resp. $G_0 \cdot R_\infty^{(i)}$) with height less than X by $R_X^{(i)}$ (resp. $S_X^{(i)}$).

Then Bhargava's averaging method, developed in [9, 10] yields the following equality for every $\mathrm{PGL}_2(\mathbb{Z})$-invariant function $\phi : V(\mathbb{Z}) \to \mathbb{R}$:

$$\begin{aligned} N^{(i)}(\phi; X) &= \frac{1}{\sigma_i \mathrm{Vol}(G_0)} \int_{\gamma \in \mathcal{F}} \Big(\sum_{f \in \gamma S_X^{(i)} \cap V(\mathbb{Z})^{\mathrm{gen}}} \phi(f) \Big) d\gamma \\ &= \frac{1}{\sigma_i \mathrm{Vol}(G_0)} \int_{t \geq \frac{4\sqrt{3}}{\sqrt{2}}} \int_{u \in N(t)} \Big(\sum_{f \in (u,t) S_X^{(i)} \cap V(\mathbb{Z})^{\mathrm{gen}}} \phi(f) \Big) t^{-2} du d^{\times} t. \end{aligned} \tag{12}$$

Above, $\sigma_i = 4$ when $i \in \{0, 2+, 2-\}$ and $\sigma_i = 2$ when $i = 1$, $d\gamma = t^{-2} du d^{\times} t dk$ is a Haar measure on $\mathrm{PGL}_2(\mathbb{R})$ with dk normalized so that the volume of K is 1, and the volume of G_0 is computed with respect to $d\gamma$. A version of (12), using the group $\mathrm{GL}_2(\mathbb{R})$ rather than $\mathrm{PGL}_2(\mathbb{R})$, and where ϕ is taken to be 1, is proved in [14, Theorem 2.5]; the proof in our case is identical.

4.2 Bounding the number of nongeneric elements

In this subsection, we bound the number of $\mathrm{PGL}_2(\mathbb{Z})$-orbits on nongeneric integral orbits on integral binary quartic forms having bounded height. By construction, we have $R_X^{(i)} = X^{1/6} R_1^{(i)}$. Since the sets $L^{(i)}$ of [14, Table 1] (and hence the sets $R_1^{(i)}$) are absolutely bounded, it follows that the absolute values of coefficients of the elements in $R_X^{(i)}$ are $\ll X^{1/6}$. Since G_0 is bounded, the same is true of the coefficients of elements in $S_X^{(i)}$. This means that when $t > CX^{1/24}$, for some sufficiently large C, every element in $(u,t)S_X^{(i)} \cap V(\mathbb{Z})$ has x^4-coefficient in $(-1, 1)$, and hence x^4-coefficient equal to 0. Such an element is not generic implying that $(u,t)S_X^{(i)} \cap V(\mathbb{Z})^{\mathrm{gen}}$ is empty. That is, every integral element $f(x,y)$ deep enough in the cusp is nongeneric. Moreover, it is nongeneric because y is a factor of $f(x,y)$.

Let a, b, c, d, and e, denote the coeffients of elements in $V(\mathbb{R})$. That is, for $f(x,y) \in V(R)$, the x^4-, x^3y-, x^2y^2-, xy^3-, and y^4-coefficients of $f(x,y)$ are denoted by $a(f)$, $b(f)$, $c(f)$, $d(f)$, and $e(f)$, respectively. Let $V(\mathbb{Z})^{\mathrm{red}}$ denote the set of reducible elements in $V(\mathbb{Z})$. In the next two lemmas, we bound the number of elements in $V(\mathbb{Z})^{\mathrm{red}}$ that lie within the main body of the fundamental domain.

Lemma 4.1 *We have*

$$\int_{1 \ll t \ll X^{1/24}} \int_{u \in [-1/2, 1/2]} \#\{(u,t)S_X^{(i)} \cap V(\mathbb{Z})^{\mathrm{red}} : a(f) \neq 0\} t^{-2} du d^{\times} t \ll_{\epsilon} X^{2/3+\epsilon}.$$

This follows immediately from the proof of [14, Lemma 2.3].

Lemma 4.2 *For a positive real number $\delta \leq 1/24$, we have*

$$\int_{t \gg 1}^{X^{\delta}} \int_{u \in [-1/2, 1/2]} \#\{f(x,y) \in (u,t)S_X^{(i)} \cap V(\mathbb{Z}) : a(f) = 0\} t^{-2} du d^{\times} t \ll X^{2/3+2\delta}.$$

Proof: For $u \in [-1/2, 1/2]$ and $1 \ll t < X^{\delta}$ we have

$$\#\{f \in (u,t)S_X^{(i)} \cap V(\mathbb{Z}) : a(f) = 0\} \ll t^4 X^{4/6},$$

since $(u,t)S_X^{(i)}$ is contained within the set of elements $f(x,y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$ with $|a| \ll t^{-4}X^{1/6}$, $|b| \ll t^{-2}X^{1/6}$, $|c| \ll X^{1/6}$, $|d| \ll t^2X^{1/6}$, and $|e| \ll t^4X^{1/6}$. Thus

$$\int_{t\gg 1}^{X^\delta}\int_{u\in[-1/2,1/2]} \#\{f \in (u,t)S_X^{(i)} \cap V(\mathbb{Z}) : a(f) = 0\}t^{-2}dud^\times t \ll X^{4/6}\int_{t\gg 1}^{X^\delta} t^2 d^\times t \ll X^{2/3+2\delta},$$

and the lemma follows. □

The next result, which follows immediately from the method of Lemma 4.2, is needed in the sequel.

Lemma 4.3 *We have*

$$\int_{t\gg 1}\int_{u\in[-1/2,1/2]} \#\{f(x,y) \in (u,t)S_X^{(i)} \cap V(\mathbb{Z})^{\mathrm{gen}} : b(f) = 0\}t^{-2}dud^\times t \ll X^{2/3}. \tag{13}$$

Suppose an element $f \in V(\mathbb{Z})$ is irreducible but nongeneric. Then its cubic resolvent polynomial must be reducible. To bound the number of such elements in the fundamental domain, we need the following result which follows from work of Nakagawa (see [36, Theorem 1]).

Proposition 4.4 *Let L be a quartic étale algebra over $\mathbb{Q}$ with ring of integers $\mathcal{O}_L$. Then the number $N(\mathcal{O}_L,k)$ of suborders of $\mathcal{O}_L$ having index k is bounded by*

$$N(\mathcal{O}_L,k) \ll_\epsilon k^\epsilon \prod_{\substack{p^2\nmid \mathrm{Disc}(\mathcal{O}_L)\\ p^3||k}} p \prod_{\substack{p^2\nmid \mathrm{Disc}(\mathcal{O}_L)\\ p^e||k,\ e\geq 4}} p^{\lfloor e/2\rfloor} \prod_{\substack{p^2|\mathrm{Disc}(\mathcal{O}_L)\\ p^e||k,\ e\geq 2}} p^{\lfloor e/2\rfloor} \tag{14}$$

In particular, we have $N(\mathcal{O}_L,k) \ll_\epsilon k^\epsilon N(k)$ where $N(k) := \prod_{p^e\|k} p^{\lfloor e/2\rfloor}$, a bound that is independent of L.

We have the following lemma improving [14, Lemma 2.4].

Lemma 4.5 *The number of $\mathrm{PGL}_2(\mathbb{Z})$-orbits f on $V(\mathbb{Z})$ with $\Delta(f) \neq 0$ and $H(f) < X$ such that the stabilizer of f in $\mathrm{PGL}_2(\mathbb{Q})$ is nontrivial is bounded by $O_\epsilon(X^{2/3+\epsilon})$.*

Proof: If the stabilizer of $f(x,y)$ in $\mathrm{PGL}_2(\mathbb{Q})$ is nontrivial, then the cubic resolvent polynomial $g(x,y)$ is reducible. Let Q and C denote the quartic ring and cubic ring corresponding to f and g, respectively. Let $L = Q\otimes\mathbb{Q}$ and $K = C\otimes\mathbb{Q}$. It follows from a result of Baily (see [2]) that the number of possible quartic $\mathbb{Q}$-algebras associated to a cubic algebra K is $\ll_\epsilon |\mathrm{Cl}(K)[2]||\Delta(K)|^\epsilon$. Since K is isomorphic to $\mathbb{Q}\times F$, for a quadratic field F, we have $|\mathrm{Cl}(K)[2]| \ll_\epsilon |\Delta(K)|^\epsilon$. Therefore, once a reducible integer cubic polynomial g is fixed, the associated quartic algebra L has $\ll_\epsilon |\Delta(g)|^\epsilon$ choices.

A quick application of [12, Lemma 4.3] implies that the number of reducible integral traceless monic cubic polynomials g with $H(g) < X$ is $\ll_\epsilon X^{1/2+\epsilon}$. Fix such a polynomial g as well as one of the $O_\epsilon(X^\epsilon)$ choices for the quartic étale algebra L. The set of $\mathrm{PGL}_2(\mathbb{Z})$-orbits f such that the cubic resolvent polynomial of f is g and the étale quartic algebra corresponding to f is L injects into the set of quartic suborders Q of L with $\mathrm{Disc}(Q) = \Delta(g)$. In particular, the index of Q in $\mathcal{O}_L$, the ring of integers of L, is determined by the pair (g,L), and is equal to $k = (|\Delta(g)|/|\mathrm{Disc}(L)|)^{1/2}$. The previous lemma now implies that the number of possible options for Q given (g,L) is $\ll_\epsilon k^\epsilon N(k) \leq k^{1/2+\epsilon}$.

Fix $\delta > 0$ to be optimized later. We will use partition the relevant set of forms f into two sets: those whose corresponding étale quartic algebra L satisfies $|\mathrm{Disc}(L)| \geq X^\delta$, and those

for which $|\mathrm{Disc}(L)| < X^\delta$. Putting together the previous observations, we see that the number of $\mathrm{PGL}_2(\mathbb{Z})$-orbits f on $V(\mathbb{Z})$ with $\Delta(f) \neq 0$, $H(f) < X$, $\mathrm{Stab}_{\mathrm{PGL}_2(\mathbb{Q})}(f) \neq 1$, and such that the étale quartic algebra L corresponding to f satisfies $|\mathrm{Disc}(L)| \geq X^\delta$ is $\ll_\epsilon X^{1/2+\epsilon} \cdot X^{(1-\delta)/4} = X^{3/4-\delta/4+\epsilon}$.

Meanwhile, the number of quartic étale algebras L with $|\mathrm{Disc}(L)| < X^\delta$ is bounded by $O(X^\delta \log X)$. Each algebra with discriminant D has $\ll_\epsilon (X/D)^{1/2+\epsilon}$ suborders with discriminant less than X (again using Nakagawa's result). This gives a bound of $O_\epsilon(X^{1/2+\delta/2+\epsilon})$ on the number of possible quartic orders with discriminant less than X and whose étale algebras have discriminant bounded by X^δ. Since a quartic order Q occurs for an absolutely bounded number of $\mathrm{PGL}_2(\mathbb{Z})$-orbits of binary quartic forms (see [14, Proposition 2.16]), we see that the number of $\mathrm{PGL}_2(\mathbb{Z})$-orbits f on $V(\mathbb{Z})$ with $\Delta(f) \neq 0$, $H(f) < X$, and such that the étale quartic algebra L corresponding to f satisfies $|\mathrm{Disc}(L)| < X^\delta$ is $\ll_\epsilon X^{1/2+\delta/2+\epsilon}$. Optimizing, we pick $\delta = 1/3$, yielding the result. □

4.3 Slicing and secondary terms

For ease of notation, we let Y denote $X^{1/6}$ throughout this subsection. Let $\phi : V(\mathbb{Z}) \to \mathbb{R}$ be a $\mathrm{PGL}_2(\mathbb{Z})$-invariant function. From (12), we have

$$N^{(i)}(\phi, X) = \frac{1}{\sigma_i \mathrm{Vol}(G_0)} \mathcal{I}(\phi, Y),$$

where $\mathcal{I}(\phi, Y)$ is defined by

$$\mathcal{I}(\phi, Y) := \int_{t \geq \frac{\sqrt[4]{3}}{\sqrt{2}}} \int_{u \in N(t)} \Big(\sum_{f \in Y \cdot (u,t) S^{(i)} \cap V(\mathbb{Z})^{\mathrm{gen}}} \phi(f) \Big) du \frac{d^\times t}{t^2}, \tag{15}$$

where we denote the bounded set $S_1^{(i)}$ by $S^{(i)}$. We break up the integral over t above into the "small t range" and the "large t range" as follows. Let $\psi_1 : \mathbb{R}_{\geq 0} \to [0,1]$ be a smooth function such that $\psi(x) = 1$ for $x \leq 2$ and $\psi_1(x) = 0$ for $x \geq 3$, and denote $1 - \psi_1$ by ψ_2. Let $0 < \kappa < 1/4$ be a real number to be optimized later. We define

$$\begin{aligned} \mathcal{I}^{(1)}(\phi, Y; \kappa) &:= \int_{t \geq \frac{\sqrt[4]{3}}{\sqrt{2}}} \int_{u \in N(t)} \psi_1\Big(\frac{t}{Y^\kappa}\Big) \Big(\sum_{f \in Y \cdot (u,t) S^{(i)} \cap V(\mathbb{Z})} \phi(f) \Big) du \frac{d^\times t}{t^2}, \\ \mathcal{I}^{(2)}(\phi, Y; \kappa) &:= \int_{t \geq \frac{\sqrt[4]{3}}{\sqrt{2}}} \int_{u \in N(t)} \psi_2\Big(\frac{t}{Y^\kappa}\Big) \Big(\sum_{f \in Y \cdot (u,t) S^{(i)} \cap V(\mathbb{Z}) : a(f) \neq 0} \phi(f) \Big) du \frac{d^\times t}{t^2}, \end{aligned} \tag{16}$$

and note that Lemmas 4.1, 4.2, and 4.5 imply the estimate

$$\mathcal{I}(\phi, Y) = \mathcal{I}^{(1)}(\phi, Y; \kappa) + \mathcal{I}^{(2)}(\phi, Y; \kappa) + O_\epsilon(X^{2/3+\kappa/3} + X^{2/3+\epsilon}). \tag{17}$$

We use the next result of Davenport to estimate the number of lattice points in bounded regions.

Proposition 4.6 ([22]) *Let $\mathcal{R}$ be a bounded, semi-algebraic multiset in $\mathbb{R}^n$ having maximum multiplicity m that is defined by at most k polynomial inequalities, each having degree at most ℓ. Let $\mathcal{R}'$ denote the image of $\mathcal{R}$ under any (upper or lower) triangular, unipotent transformation of $\mathbb{R}^n$. Then the number of lattice points (counted with multiplicity) contained in the region $\mathcal{R}'$ is given by*

$$\mathrm{Vol}(\mathcal{R}) + O(\max\{\overline{\mathrm{Vol}}(\mathcal{R})\}, 1\}),$$

where $\overline{\mathrm{Vol}}(\mathcal{R})$ denotes the greatest d-dimensional volume of any projection of $\mathcal{R}$ onto a coordinate subspace obtained by equating $n-d$ coordinates to zero, where d ranges over all values in $\{1, \ldots, n-1\}$. The implied constant in the second summand depends only on n, m, k, and ℓ.

For the rest of this subsection, we restrict our function ϕ to be defined via congruence classes modulo some positive integer m, and will keep track of the dependence of our error terms on m. We handle the cases of small and large t separately.

The small t case: estimates for $\mathcal{I}^{(1)}(\phi, Y; \kappa)$

Let m be a positive integer, let $\overline{\phi} : V(\mathbb{Z}/m\mathbb{Z}) \to \mathbb{R}$ be a bounded (independent of m) function, and denote the lift of $\overline{\phi}$ to $V(\mathbb{Z})$ by ϕ. Recall the definitions of $\nu(\phi)$ and $\nu_a(\phi)$ from (8). We also define $\mathrm{supp}(\phi)$ to denote the cardinality of the support of $\overline{\phi}$ in $V(\mathbb{Z}/m\mathbb{Z})$. Then we have the following lemma.

Lemma 4.7 *Let $\phi : V(\mathbb{Z}) \to \mathbb{R}$ be an absolutely bounded function defined modulo a positive integer m. For $u \in [-1/2, 1/2]$ and $t \gg 1$, we have*

$$\sum_{f \in Y \cdot (u,t)S^{(i)} \cap V(\mathbb{Z})} \phi(f) = \nu(\phi)\mathrm{Vol}(S^{(i)})X^{5/6} + O\Big(\frac{\mathrm{supp}(\phi)t^4X^{2/3}}{m^4} + \frac{\mathrm{supp}(\phi)t^6X^{1/2}}{m^3} + \mathrm{supp}(\phi)t^6\Big).$$

Proof: Partition the support of ϕ in $V(\mathbb{Z})$ into $\mathrm{supp}(\phi)$ translated lattices $L_v := v + mV(\mathbb{Z})$, where v ranges over lifts of elements $\bar{v}$ in the support of $\overline{\phi}$ in $V(\mathbb{Z}/m\mathbb{Z})$. The coefficients of an element $f \in Y \cdot (u,t)S^{(i)}$ satisfy

$$|a(f)| \ll t^{-4}Y; \quad |b(f)| \ll t^{-2}Y; \quad |c(f)| \ll Y; \quad |d(f)| \ll t^2Y; \quad |e(f)| \ll t^4Y.$$

Applying Proposition 4.6 to the set $(Y \cdot (u,t)S^{(i)} - v)/m$ yields

$$\#\big(Y \cdot (u,t)S^{(i)} \cap L_v\big) = \frac{\mathrm{Vol}(S^{(i)})Y^5}{m^5} + O\Big(\frac{t^4Y^4}{m^4} + \frac{t^6Y^3}{m^3} + \frac{t^6Y^2}{m^2} + \frac{t^4Y}{m} + 1\Big).$$

Since we assume that $t \gg 1$, the first two summands in the above error term added to t^6 dominate the others. The lemma then follows by summing the above expression over $\bar{v} \in V(\mathbb{Z}/m\mathbb{Z})$ that are in the support of $\overline{\phi}$. $\square$

Estimating the inner sum of f using Lemma 4.7, we obtain

$$\begin{aligned} \mathcal{I}^{(1)}(\phi, Y; \kappa) &= \nu(\phi)\mathrm{Vol}(S^{(i)})X^{5/6} \int_{t \geq \frac{\sqrt[4]{3}}{\sqrt{2}}} \int_{u \in N(t)} \psi_1\Big(\frac{t}{Y^\kappa}\Big) du \frac{d^\times t}{t^2} \\ &\quad + O\Big(\frac{\mathrm{supp}(\phi)X^{\frac{2}{3}+\frac{\kappa}{3}}}{m^4} + \frac{\mathrm{supp}(\phi)X^{\frac{1}{2}+\frac{2\kappa}{3}}}{m^3} + \mathrm{supp}(\phi)X^{\frac{2\kappa}{3}}\Big), \end{aligned} \tag{18}$$

where we recall that $Y = X^{1/6}$.

The large t case: estimates for $\mathcal{I}^{(2)}(\phi, Y; \kappa)$

To evaluate the contribution from the large range of t, we fiber by the x^4-coefficients of integral binary quartic forms. That is, we write

$$\mathcal{I}^{(2)}(\phi, Y; \kappa) = \sum_{a \neq 0} \int_{t \geq \frac{\sqrt[4]{3}}{\sqrt{2}}} \int_{u \in N(t)} \psi_2\Big(\frac{t}{Y^\kappa}\Big)\Big(\sum_{f \in Y \cdot (u,t)S^{(i)} \cap V_a(\mathbb{Z})} \phi(f)\Big) du \frac{d^\times t}{t^2}. \tag{19}$$

As before, we assume that ϕ is the lift of some function $\overline{\phi} : V(\mathbb{Z}/m\mathbb{Z}) \to \mathbb{R}$ which is absolutely bounded (independent of m). For any integer a and any ring R, let $V_a(R)$ denote the set of elements in $V(R)$ whose x^4-coefficients are equal to a. For $a \in \mathbb{Z}$, recall that $\nu_a(\phi)$ denotes the density of ϕ restricted to $V_a(\mathbb{Z})$, and for $a \in \mathbb{Z}$, let $\mathrm{supp}(\phi, a)$ denote the cardinality of the support of $\overline{\phi}$ restricted to $V_a(\mathbb{Z}/m\mathbb{Z})$. For a set $T \subset V(\mathbb{R})$, we let $T|_a$ denote $T \cap V_a(\mathbb{R})$. Then from a similar argument in the proof of Lemma 4.7, applying Proposition 4.6 yields the following result.

Lemma 4.8 *With notation as above, we have*

$$\sum_{f \in Y \cdot (u,t)S^{(i)} \cap V_a(\mathbb{Z})} \phi(f) = \nu_a(\phi) t^4 X^{2/3} \mathrm{Vol}\Big(S^{(i)}|_{t^4 a/X^{1/6}}\Big) + O\Big(\frac{\mathrm{supp}(\phi, a) t^6 X^{1/2}}{m^3} + \mathrm{supp}(\phi, a) t^6\Big).$$

We return to (19), and note that for $a \neq 0$, the set $(Y \cdot (u,t)S^{(i)})|_a$ is empty unless $t \ll (Y/|a|)^{1/4}$. Since $\psi_2(t/Y^\kappa) = 0$ unless $t \gg Y^\kappa$, only $a \neq 0$ with $|a| \ll Y^{1-4\kappa}$ contributes to (19). We truncate the sum over a and integral over t in (19) using these observations, and apply Lemma 4.8 to estimate the sum over f. Integrating the resulting error term over t, u, and summing over a gives an error that is $\ll$

$$\Big(\frac{X^{1/2}}{m^3} + 1\Big) \sum_{\substack{a \neq 0 \\ |a| \ll Y^{1-4\kappa}}} \mathrm{supp}(\phi, a) \int_{t \ll (Y/|a|)^{1/4}} t^4 d^\times t \ll \Big(\frac{X^{2/3}}{m^3} + X^{1/6}\Big) \sum_{\substack{a \neq 0 \\ |a| \ll Y^{1-4\kappa}}} \frac{\mathrm{supp}(\phi, a)}{|a|}.$$

Meanwhile, the main term contribution to $\mathcal{I}^{(2)}(\phi, Y; \kappa)$ is

$$\begin{aligned} & X^{2/3} \sum_{a \neq 0} \nu_a(\phi) \int_{t \geq 1} \psi_2\Big(\frac{t}{Y^\kappa}\Big) \mathrm{Vol}\Big(S^{(i)}|_{t^4 a/Y}\Big) t^2 d^\times t \\ = \; & \frac{X^{3/4}}{4} \sum_{a \neq 0} \frac{\nu_a(\phi)}{|a|^{\frac{1}{2}}} \int_{\substack{\alpha \in \mathbb{R}^\times \\ a\alpha > 0}} \psi_2\Big(\frac{|\alpha|^{\frac{1}{4}} Y^{\frac{1}{4}-\kappa}}{|a|^{\frac{1}{4}}}\Big) \mathrm{Vol}(S^{(i)}|_\alpha) |\alpha|^{\frac{1}{2}} d^\times \alpha \end{aligned} \tag{20}$$

by the change of variables $\alpha = t^4 a/Y$.

We next have the following lemma which follows immediately since $\nu_a(\phi)$ only depends on the residue class of a modulo m.

Lemma 4.9 *Define the functions*

$$D^\pm(\phi, s) := \sum_{a > 0} \frac{\nu_{\pm a}(\phi)}{a^s}.$$

Then the functions $D^\pm(\phi, s)$ admit meromorphic continuations to $\mathbb{C}$, and are holomorphic with the exception of possible simple poles at $s = 1$ with residue $\nu(\phi)$.

Then with a computation identical to [16, (Equation 41)], we obtain for any $M > 1$

$$\begin{aligned} \sum_{\pm a > 0} \frac{\nu_a(\phi)}{|a|^{\frac{1}{2}}} \psi_2\Big(\frac{|\alpha|^{\frac{1}{4}} Y^{\frac{1}{4}-\kappa}}{|a|^{\frac{1}{4}}}\Big) \; = \; & D^\pm(\phi, 1/2) + 4\nu(\phi) \widetilde{\psi_2}(-2) (|\alpha| Y^{1-4\kappa})^{\frac{1}{2}} \\ & + O_M(\min((|\alpha| Y^{1-4\kappa})^{-M}, 1)). \end{aligned} \tag{21}$$

Above, $\widetilde{\psi_2}$ denotes the Mellin transform of ψ_2. We will choose $\kappa < 1/4$ and so the error term is negligible. The contribution of the second summand of (21) to the right hand side of (20) is equal to

$$\nu(\phi)X^{5/6}\widetilde{\psi_2}(-2)Y^{-2\kappa}\int_{\alpha\in\mathbb{R}^\times}\mathrm{Vol}(S^{(i)}|_\alpha)d\alpha = \nu(\phi)\mathrm{Vol}(S^{(i)})X^{5/6}\int_{t>0}\psi_2\Big(\frac{t}{Y^\kappa}\Big)t^{-2}d^\times t,$$

where the equality follows from the definition of the Mellin transform. Finally, the contribution of the first summand in (21) to (20) is

$$\frac{1}{4}\Big(D^+(\phi,1/2)\mathcal{V}^+(S^{(i)}) + D^-(\phi,1/2)\mathcal{V}^-(S^{(i)})\Big)X^{3/4}, \text{ where}$$

$$\mathcal{V}^\pm(S^{(i)}) := \int_{\pm\alpha>0}\mathrm{Vol}(S^{(i)}|_\alpha)\frac{1}{|\alpha|^{1/2}}d\alpha.$$

Therefore, we have

$$\begin{aligned}\mathcal{I}^{(2)}(\phi,Y;\kappa) &= \nu(\phi)\mathrm{Vol}(S^{(i)})X^{5/6}\int_{t>0}\psi_2\Big(\frac{t}{Y^\kappa}\Big)t^{-2}d^\times t\\ &\quad+\frac{1}{4}\Big(D^+(\phi,1/2)\mathcal{V}^+(S^{(i)}) + D^-(\phi,1/2)\mathcal{V}^-(S^{(i)})\Big)X^{3/4}\\ &\quad+O\left(\Big(\frac{X^{2/3}}{m^3}+X^{1/6}\Big)\sum_{0\neq|a|\ll Y^{1-4\kappa}}\frac{\mathrm{supp}(\phi,a)}{|a|}\right).\end{aligned} \tag{22}$$

We define the quantities

$$M^{(i)}_{5/6}(\phi) := \frac{\nu(\phi)\mathrm{Vol}(\mathcal{F})\mathrm{Vol}(S^{(i)})}{\sigma_i\mathrm{Vol}(G_0)};\quad M^{(i)}_{3/4}(\phi) := \frac{D^+(\phi,1/2)\mathcal{V}^+(S^{(i)}) + D^-(\phi,1/2)\mathcal{V}^-(S^{(i)})}{4\sigma_i\mathrm{Vol}(G_0)}.$$

Adding the right hand sides of (18) and (22) yields

$$N^{(i)}(\phi,X) = M^{(i)}_{5/6}(\phi)X^{5/6} + M^{(i)}_{3/4}(\phi)X^{3/4} + O_\epsilon(X^{2/3+\epsilon} + X^{2/3+\kappa/3} + E(\phi;X;\kappa)), \tag{23}$$

for every κ, where the error term $E(\phi;X)$ is given by

$$E(\phi;X;\kappa) = \Big(\frac{X^{\frac{2}{3}+\frac{2\kappa}{3}}}{m^4} + \frac{X^{\frac{1}{2}+\frac{4\kappa}{3}}}{m^3} + X^{\frac{4\kappa}{3}}\Big)\mathrm{supp}(\phi) + \Big(\frac{X^{\frac{2}{3}}}{m^3}+X^{\frac{1}{6}}\Big)\sum_{\substack{0\neq a\\ |a|\ll Y^{1-4\kappa}}}\frac{\mathrm{supp}(\phi,a)}{|a|}.$$

For any fixed m and ϕ, (23) holds for all positive X. This automatically implies that $M^{(i)}_1(\phi)$ and $M^{(i)}_2(\phi)$ are independent of the choice of G_0.

We are now ready to prove the analogues, Theorems 2 and 6, of Shintani's result in the case of binary quartic forms.

Proof of Theorem 2: The theorem requires us to provide a secondary term in the estimate of $N^{(i)}(\phi,X)$ when $\phi(f) = 1$ for all f. In this case, we have $m = 1$, $\nu(\phi) = \nu(a,\phi) = 1$ for all a and

hence $D^{\pm}(\phi, s) = \zeta(s)$. The result now follows by taking $\kappa = \epsilon$ in (23) in conjunction with the computations of $M^{(i)}_{5/6}(\phi)$ and $M^{(i)}_{3/4}(\phi)$ in §8. □

Proof of Theorem 6: For this result, we take $m = \prod_{p\in S} p$ and write $\phi : V(\mathbb{Z}/m\mathbb{Z}) \to \mathbb{R}$ as $\prod_{p\in S} \phi_p$, where $\phi_p : V(\mathbb{Z}/p\mathbb{Z}) \to \mathbb{R}$ is the characteristic function of the set of elements having splitting type σ_p. The value of the main term follows from (23). The value of the secondary term follows in conjunction with the special values of $D(\phi, 1/2)$ computed in Table 2. The error term is easily seen to also follow from (23). □

Remark 4.10 We note that the error terms for Theorem 6 in the Σ-aspects can be improved using the equidistribution results we prove in §6, specifically, Corollary 6.8. In fact, since the functions ϕ_p under consideration are more than strongly invariant–they are invariant under multiplication by all units (not merely squares) in $\mathbb{F}_p$–we can get even better error bounds by improving Corollary 6.8. Finally, for many applications, it is necessary to take ϕ_p to be certain natural linear combinations of characteristic functions of elements with a particular splitting types. For such functions, it is often the case that further cancellations in the Fourier transform, leading to further improvements to the error terms, may be obtained. For some such examples, see [47], [33], [17], [7].

5 Constructing periodic approximations of local functions

In this section, we prove that m_p and ℓ_p are and almost well approximated by periodic functions (and well approximated except in the case of ℓ_2). We do this by explicitly constructing the periodic approximations $m_p^{(k)}$ and $\ell_p^{(k)}$ in §5.1 and §5.3, respectively. In §5.2, for $p \geq 5$, we give a natural interpretation of $m_p^{(k)}$ using the Bruhat–Tits tree of $\mathrm{PGL}_2(\mathbb{Q}_p)$. Finally, at the end of §5.3, we prove that ℓ_p/m_p is large and almost well appoximated, and well approximated for $p \geq 3$.

5.1 Approximating the number of $\mathrm{PGL}_2(\mathbb{Z}_p)$-orbits in a $\mathrm{PGL}_2(\mathbb{Q}_p)$-equivalence class

Proposition 5.1 *The function m_p is well approximated by periodic functions.*

Proof: For $k \geq 0$, let $\mathcal{C}^{(k)} = \mathrm{PGL}_2(\mathbb{Z}_p) \begin{pmatrix} p^k & 0 \\ 0 & 1 \end{pmatrix} \mathrm{PGL}_2(\mathbb{Z}_p) \subset \mathrm{PGL}_2(\mathbb{Q}_p)$, where we use the same notation $\begin{pmatrix} p^k & 0 \\ 0 & 1 \end{pmatrix}$ for an element in $\mathrm{GL}_2(\mathbb{Q}_p)$ and its image in $\mathrm{PGL}_2(\mathbb{Q}_p)$. Then it is well known that $\mathrm{PGL}_2(\mathbb{Q}_p) = \bigsqcup_{k\geq 0} \mathcal{C}^{(k)}$ and that the size of $\mathrm{PGL}_2(\mathbb{Z}_p)\backslash\mathcal{C}^{(k)}$ is finite for each k. For $f \in V(\mathbb{Z}_p)$, we put $\mathcal{C}_f^{(k)} = \{g \in \mathcal{C}^{(k)} \mid g \cdot f \in V(\mathbb{Z}_p)\}$ and define $m_p^{(k)}(f)$ to be the size of $\mathrm{PGL}_2(\mathbb{Z}_p)\backslash\mathcal{C}_f^{(k)}$. Then $m_p = \sum_{k\geq 0} m_p^{(k)}$ and $m_p^{(0)}$ is identically 1.

We show that the function $m_p^{(k)}$ is periodic with period p^{2k}. Suppose $f_1, f_2 \in V(\mathbb{Z}_p)$ are congruent modulo p^{2k}. We let $h = f_1 - f_2 \in p^{2k}V(\mathbb{Z}_p)$. For $g \in \mathcal{C}^{(k)}$, $g \cdot h \in V(\mathbb{Z}_p)$, so $g \cdot f_1 \in V(\mathbb{Z}_p)$ if and only if $g \cdot f_2 \in V(\mathbb{Z}_p)$. Therefore $\mathcal{C}_{f_1}^{(k)} = \mathcal{C}_{f_2}^{(k)}$, implying $m_p^{(k)}(f_1) = m_p^{(k)}(f_2)$.

Suppose that $m_p^{(k)}(f) \neq 0$. Then $\mathcal{C}_f^{(k)} \neq \emptyset$. Take $g = \gamma_1 \begin{pmatrix} p^k & 0 \\ 0 & 1 \end{pmatrix} \gamma_2 \in \mathcal{C}_f^{(k)}$, for elements $\gamma_1, \gamma_2 \in \mathrm{PGL}_2(\mathbb{Z}_p)$ and let $f' = \gamma_2 f \in V(\mathbb{Z}_p)$. Then $\begin{pmatrix} p^k & 0 \\ 0 & 1 \end{pmatrix} f' \in V(\mathbb{Z}_p)$. Let $f' = (a_0, b_0, c_0, d_0, e_0)$.

Then $\begin{pmatrix} p^k & 0 \\ 0 & 1 \end{pmatrix} f' = (a_0/p^{2k}, b_0/p^k, c_0, d_0p^k, e_0p^{2k})$ and so $p^{2k} \mid a_0$ and $p^k \mid b_0$. By (24), this implies $p^{2k} \mid \Delta(f')$. Therefore $p^{2k} \mid \Delta(f)$ as well since $\Delta(f) = \Delta(f')$. $\square$

5.2 Counting nodes in the Bruhat–Tits tree of $\mathrm{PGL}_2(\mathbb{Q}_p)$

Let $p \geq 5$ be prime. In this subsection, which is unnecessary for the proofs of the main results, we give an alternative and more natural description of $m_p^{(k)}$ in terms of the Bruhat–Tits tree $\mathcal{T} = (\mathcal{V}, \mathcal{D})$ of $\mathrm{PGL}_2(\mathbb{Q}_p)$. Recall that $\mathcal{T}$ is an undirected simple graph whose nodes $\mathfrak{n} \in \mathcal{V}$ correspond to homothety classes of lattices in $\mathbb{Q}_p^2$, where we recall that two lattices L_1 and L_2 are *homothetic* if there exists $\theta \in \mathbb{Q}_p^\times$ with $\theta L_1 = L_2$. Given a pair (L, L') of lattices in $\mathbb{Q}_p^2$, there exist integers a and b and a basis $\{v_1, v_2\}$ of L such that $\{p^a v_1, p^b v_2\}$ is a basis of L'. Then it is easy to check that a and b are invariants of the pair (L, L'), and that $|a - b|$ remains invariant even when we replace L and L' by homothetic lattices. This yields a function

$$\mathrm{inv} : \mathcal{V} \times \mathcal{V} \to \mathbb{Z}_{\geq 0}.$$

Moreover, $\mathrm{inv}(\mathfrak{n}_1, \mathfrak{n}_2) = 0$ if and only if $\mathfrak{n}_1 = \mathfrak{n}_2$. We say that two nodes $\mathfrak{n}_1, \mathfrak{n}_2 \in \mathcal{V}$ are *neighbors* if $\mathrm{inv}(\mathfrak{n}_1, \mathfrak{n}_2) = 1$, and there is an edge in $\mathcal{D}$ between two nodes if and only if they are neighbors. It is well known that this set of vertices and edges makes $\mathcal{T}$ into a regular tree, where each vertex has degree $p + 1$. Furthermore, the distance between two nodes $\mathfrak{n}_1$ and $\mathfrak{n}_2$ is $\mathrm{inv}(\mathfrak{n}_1, \mathfrak{n}_2)$.

The group $\mathrm{GL}_2(\mathbb{Q}_p)$ acts on the set of lattices in $\mathbb{Q}_p^2$ via $\gamma \cdot L := \{\gamma v : v \in L\}$. This action descends to an action of $\mathrm{PGL}_2(\mathbb{Q}_p)$ on the set of homothety classes of lattices, and hence an action of $\mathrm{PGL}_2(\mathbb{Q}_p)$ on $\mathcal{V}$. This action is transitive since the aforementioned action of $\mathrm{GL}_2(\mathbb{Q}_p)$ is transitive. Denote the node corresponding to the homothety class of the lattice $\mathbb{Z}_p^2$ by $\mathfrak{o}$. Then it is easy to see that the stabilizer in $\mathrm{PGL}_2(\mathbb{Q}_p)$ of $\mathfrak{o}$ is $\mathrm{PGL}_2(\mathbb{Z}_p)$. This gives a bijection

$$\Phi : \mathrm{PGL}_2(\mathbb{Q}_p)/\mathrm{PGL}_2(\mathbb{Z}_p) \to \mathcal{V},$$

sending the $\mathrm{PGL}_2(\mathbb{Z}_p)$-orbit of $\gamma \in \mathrm{PGL}_2(\mathbb{Q}_p)$ to the homothety class of $\gamma \cdot \mathfrak{o}$. For an integer $k \geq 1$, let $\mathfrak{o}_k$ denote the node corresponding to the homothety class of $p^k\mathbb{Z}_p \oplus \mathbb{Z}_p$. Then $\mathfrak{o}_k$ has distance k from $\mathfrak{o}$. It is well known that $\mathrm{PGL}_2(\mathbb{Z}_p)$ acts transitively on the set of nodes having distance k from $\mathfrak{o}$, and hence every node having distance k from $\mathfrak{o}$ is $\mathrm{PGL}_2(\mathbb{Z}_p)$-equivalent to $\mathfrak{o}_k$.

To describe the connection between $\mathcal{T}$ and binary quartic forms, we define the following notion: Denote the determinant of L by $d(L)$. Then we say that $f \in V(\mathbb{Q}_p)$ is *integral-valued with respect to* L if $d(L)^2 \mid f(v)$ for every $v \in L$. Since this notion is homothety invariant, it descends to a notion of f being *integral valued* with respect to nodes in $\mathcal{V}$. We next prove that this notion of integrality respects the action of $\mathrm{PGL}_2(\mathbb{Q}_p)$.

Proposition 5.2 *Let $f \in V(\mathbb{Q}_p)$, $\mathfrak{n} \in \mathcal{V}$ and $\gamma \in \mathrm{PGL}_2(\mathbb{Q}_p)$, be any elements. Then f is integral-valued with respect to $\gamma \cdot \mathfrak{n}$ if and only if $\gamma \cdot f$ is integral with respect to $\mathfrak{n}$.*

Proof: Let $L \subset \mathbb{Q}_p^2$ and $\widetilde{\gamma} \in \mathrm{GL}_2(\mathbb{Q}_p)$ be representatives of $\mathfrak{n}$ and γ, respectively. By definition, $\gamma \cdot f$ is integral-valued with respect to $\mathfrak{n}$ if and only if $d(L)^2 \mid (\gamma \cdot f)(x, y)$ for every $(x, y) \in L$. However, we have $(\gamma \cdot f)(x, y) = f((x, y) \cdot \widetilde{\gamma})/(\det(\widetilde{\gamma}))^2$. Therefore, we have that $\gamma \cdot f$ is integral-valued with respect to $\mathfrak{n}$ if and only if $(\det(\widetilde{\gamma})d(L))^2 \mid f(x, y)$ for every $(x, y) \in \widetilde{\gamma}L$, which is true if and only if f is integral-valued with respect to $\gamma \cdot \mathfrak{n}$, as necessary. $\square$

We next have the following consequence describing the image of $\mathrm{PGL}_2(\mathbb{Q}_p)$ under Φ.

Corollary 5.3 *Let $f \in V(\mathbb{Q}_p)$ be any element. Then $\gamma \in \mathrm{PGL}_2(\mathbb{Q}_p)$ belongs to $\mathrm{PGL}_2(\mathbb{Q}_p)_f$ (equivalently, $\gamma \cdot f \in V(\mathbb{Z}_p)$) if and only if f is integral with respect to $\Phi(\gamma)$.*

Proof: We know that f is integral with respect to $\Phi(\gamma) = \gamma \cdot \mathfrak{o}$ if and only if $\gamma \cdot f$ is integral with respect to $\mathfrak{o}$ which is true if and only if $\gamma \cdot f \in V(\mathbb{Z}_p)$ as necessary. $\square$

It is only here that we use $p \geq 5$. It is needed since for $p = 2$ and $p = 3$, there exist elements $f \in V(\mathbb{Q}_p) \backslash V(\mathbb{Z}_p)$ for which $f(v) \in \mathbb{Z}_p$ for all $v \in \mathbb{Z}_p$.

Finally, denote the set of nodes in $\mathcal{V}$ with respect to which $f \in V(\mathbb{Q}_p)$ is integral by $\mathcal{V}(f)$. An immediate consequence of Corollary 5.3 is that $m_p(f)$ is equal to $\#\mathcal{V}(f)$. For an integer $k \geq 0$, let $\mathcal{V}_k$ denote the set of nodes in $\mathcal{V}$ that have distance k from $\mathfrak{o}$. It is easy to check that $m_p^{(k)} = \#(\mathcal{V}(f) \cap \mathcal{V}_k)$.

5.3 Approximating local solubility

Let p be a prime and recall that we defined $\ell_p : V(\mathbb{Z}_p) \to \{0, 1\}$ to be the characteristic function of the set of elements in $V(\mathbb{Z}_p)$ that are soluble. In this section, we prove that ℓ_p is well approximated by periodic functions if p is odd, and is almost well approximated by periodic functions if $p = 2$.

We first show the following:

Proposition 5.4 *Let $k \geq 0$.*

1. *Suppose p is odd. If $f \in V(\mathbb{Z}_p)$ is insoluble but there exist a soluble $f_0 \in V(\mathbb{Z}_p)$ such that $f \equiv f_0 \pmod{p^{2k}}$, then $p^{2k+2} \mid \Delta(f)$.*
2. *Suppose $p = 2$. If $f \in V(\mathbb{Z}_p)$ is insoluble but there exist a soluble $f_0 \in V(\mathbb{Z}_p)$ such that $f \equiv f_0 \pmod{p^{2k+2}}$, then $p^{2k+2} \mid \Delta(f)$.*

For the proof, we begin with some preliminary results.

Lemma 5.5 *Let $g(x) \in \mathbb{Z}_p[x]$ and $a \in \mathbb{Z}_p$. Let $\lambda = \mathrm{ord}_p(g(a))$ and $\mu = \mathrm{ord}_p(g'(a))$. If $\lambda > 2\mu$, then $\{g(a + \xi) \mid \xi \in p^{\lambda-\mu}\mathbb{Z}_p\} = p^\lambda \mathbb{Z}_p$.*

This is a version of Hensel's lemma. We have $g(a + \xi) = g(a) + \xi g'(a) + \xi^2 h(a, \xi)$ for a polynomial $h(x, y) \in \mathbb{Z}_p[x, y]$, and the proof is standard with this identity. Next, we have the following lemma, which will also be used in the next section, regarding congruences for the discriminant polynomial $\Delta(f)$ for $f = (a, b, c, d, e) \in V$:

Lemma 5.6 *We have*

$$\Delta(f) \equiv b^2(c^2d^2 + 18bcde - 4bd^3 - 4c^3e - 27b^2e^2) \pmod{(ae, ad^2)}, \tag{24}$$

$$\Delta(f) \equiv 4ac^3(4ce - d^2) \pmod{(a^2, ab, b^2)}, \tag{25}$$

$$\Delta(f) \equiv 0 \pmod{(a^2, abc, abd, ac^2, b^4, b^3d, b^2c^2)}. \tag{26}$$

This immediately follows from the concrete form the discriminant:

$$\begin{aligned}\Delta(f) =& 256a^3e^3 \\ &- 192a^2bde^2 - 128a^2c^2e^2 + 144a^2cd^2e - 27a^2d^4 \\ &+ 144ab^2ce^2 - 6ab^2d^2e - 80abc^2de + 18abcd^3 + 16ac^4e - 4ac^3d^2 \\ &- 27b^4e^2 + 18b^3cde - 4b^3d^3 - 4b^2c^3e + b^2c^2d^2.\end{aligned}$$

As an immediate consequence, we obtain the following:

Lemma 5.7 *Let $k \geq 1$. Suppose the coefficients of $f(x,y) = ax^4+bx^3y+cx^2y^2+dxy^3+ey^4 \in V(\mathbb{Z}_p)$ satisfy one of the following divisibility properties:*

- *$p^{2k+1} \mid a$, $p^{k+1} \mid b$, $p \mid d$, and $p \mid e$;*
- *$p \mid b$, $p \mid c$, $p^k \mid d$, and $p^{2k} \mid e$.*

Then $p^{2k+2} \mid \Delta(f)$.

Proof: This follows from (24) and (26). Note that $\Delta(a,b,c,d,e) = \Delta(e,d,c,b,a)$. □

We now prove Proposition 5.4.

Proof: Recall from the proof of [14, Proposition 3.18] that if f is insoluble, the splitting type of $f \pmod p$ is either (1^21^2), (2^2), (1^4) or (0). In particular $p^2 \mid \Delta(f)$ so we have the assertion for $k = 0$. For the rest of the proof we assume $k \geq 1$. By suitably multiplying an element in $\mathbb{Q}_p^\times$ to the variables and further transforming the variables by an elemnt in $\mathrm{PGL}_2(\mathbb{Z}_p)$, we may assume that $e_0 = f_0(0,1)$ is a squared element. Let $f(x,y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$. Then $e \equiv e_0 \pmod{p^{2k}}$. If $p^{2k} \nmid e_0$, then the congruence condition implies that $e/e_0 \in 1 + p\mathbb{Z}_p$ if p is odd and $e/e_0 \in 1 + 8\mathbb{Z}_2$ if $p = 2$. In particular $e/e_0 = u^2$ for some $u \in \mathbb{Z}_p^\times$ and so $e = u^2e_0$ is a squared element. This contradicts to f being not locally soluble. Thus $p^{2k} \mid e_0$ and so $p^{2k} \mid e$. If $p^k \nmid d$, then Lemma 5.5 implies that f is locally soluble, again a contradiction. Thus $p^k \mid d$.

Since f is insoluble, the splitting type of $f \pmod p$ is either (1^21^2), (2^2), (1^4) or (0). If the splitting type is (2^2), then $e = f(0,1)$ must be in $\mathbb{Z}_p^\times$, so this can not occur. If the splitting type is either (1^4) or (0), then $p \mid b$ and $p \mid c$, so by Lemma 5.7, $p^{2k+2} \mid \Delta(f)$ as needed. Suppose the splitting type of f is (1^21^2). Applying a $\mathrm{PGL}_2(\mathbb{Z}_p)$-transformation if necessary, we may assume that the coefficientsof f satisfy $p \mid a$, $p \mid b$, $p \nmid c$, $p^k \mid d$, and $p^{2k} \mid e$. Since f is insoluble and local solubility is a $\mathrm{PGL}_2(\mathbb{Q}_p)$-invariant property, it follows that $f_1 = \left(\begin{smallmatrix} p^k & \\ & 1 \end{smallmatrix}\right) f$ is insoluble. Now the coefficients $(a_1,b_1,c_1,d_1,e_1) = (p^{2k}a, p^kb, c, d/p^k, e/p^{2k})$ are all integral and satisfy $p^{2k+1} \mid a_1$, $p^{k+1} \mid b_1$, and $p \nmid c_1$. Since f_1 is insoluble, the splitting type of f_1 must again be (1^21^2) (the only other options, (1^22) or (1^211), would imply that f_1 is soluble). Therefore, applying another $\mathrm{PGL}_2(\mathbb{Z}_p)$-transformation, we may assume that in addition we also have $p \mid d_1$ and $p \mid e_1$. Applying Lemma 5.7, it follows that $p^{2k+2} \mid \Delta(f_1) = \Delta(f)$ as necessary. □

We are now ready to prove the following:

Proposition 5.8 *If p is an odd prime, then the function ℓ_p is well approximated by periodic functions. If $p = 2$, then the function ℓ_p is almost well approximated by periodic functions.*

Proof: Let $k \geq 0$. For $f \in V(\mathbb{Z}_p) \setminus \{\Delta = 0\}$, we let $L_p^{(k)}(f) = 0$ if any $g \in V(\mathbb{Z}_p)$ satisfying $g \equiv f \pmod{p^{2k}}$ is not locally soluble, and $L_p^{(k)}(f) = 1$ otherwise. Then we define the functions $\ell_p^{(k)}$ on $V(\mathbb{Z}_p)$ by setting $\ell_p^{(0)} := L_p^{(0)}$ and $\ell_p^{(k)} := L_p^{(k)} - L_p^{(k-1)}$ for $k \geq 1$. By definition $L_p^{(k)}$ is periodic with period p^{2k} and is $\mathrm{PGL}_2(\mathbb{Z}_p)$-invariant, so is $\ell_p^{(k)}$. We show that ℓ_p is well approximated by $\ell_p^{(k)}$ for odd p, and almost well approximated by $\ell_p^{(k)}$ for $p = 2$.

By definition, $\ell_p^{(0)} = L_p^{(0)}$ is identically 1, and $L_p^{(k)}(f) \geq L_p^{(k+1)}(f)$ for all k and f. If f is soluble then $L_p^{(k)}(f) = 1$ for all k, meanwhile if f is insoluble and $p^{2k+2} \nmid \Delta(f)$, then Proposition 5.4 implies that $L_p^{(k)}(f) = 0$ if p is odd, and $L_p^{(k+1)}(f) = 0$ if $p = 2$. Thus $\lim_{k\to\infty} L_p^{(k)}(f) = \ell_p(f)$ for all $f \in V(\mathbb{Z}_p) \setminus \{\Delta = 0\}$. Let $k \geq 1$, and suppose $\ell_p^{(k)}(f) \neq 0$. This happens only when $L_p^{(k-1)}(f) = 1$

and $L_p^{(k)}(f) = 0$. Then Proposition 5.4 implies that $p^{2k} \mid \Delta(f)$ if p is odd, and $p^{2k-2} \mid \Delta(f)$ if $p = 2$. $\square$

Finally, we deduce that ℓ_p/m_p is well approximated for all odd primes, and almost well approximated for $p = 2$. To do this, we prove that the product of two (almost) well approximated functions is (almost) well approximated, and that under certain conditions, the inverse of a large and (almost) well approximated function is also (almost) well approximated.

Lemma 5.9 *Suppose that $\phi : V(\mathbb{Z}_p)\backslash\{\Delta = 0\} \to \mathbb{R}$ and $\phi' : V(\mathbb{Z}_p)\backslash\{\Delta = 0\} \to \mathbb{R}$ are large and (almost) well approximated. Then their product $\phi\phi'$ is also large and (almost) well approximated.*

Suppose $\phi : V(\mathbb{Z}_p)\backslash\{\Delta = 0\} \to \mathbb{R}_{>0}$ is bounded away from 0, large and (almost) well approximated via the series of functions $\phi^{(k)}$. Let $\Phi^{(k)}$ denote the partial sums of $\phi^{(k)}$, i.e., $\Phi^{(k)} := \sum_{n=0}^{k} \phi^{(k)}$, and assume that the $\Phi^{(k)}$ are also bounded away from 0 by an absolute constant. That is, we have $\Phi^{(k)}(f) \geq c > 0$ for all f, where c is an absolute constant. Then $1/\phi$ is large and (almost) well approximated.

Proof: We define the partial sums $\Phi^{(k)} := \sum_{n=0}^{k} \phi^{(k)}$ and $\Phi'^{(k)} := \sum_{n=0}^{k} \phi'^{(k)}$; define $\Psi^{(k)} := \Phi^{(k)}\Phi'^{(k)}$; and finally define $\psi^{(0)}$ to be identically 1 and $\psi^{(k)} := \Psi^{(k)} - \Psi^{(k-1)}$ for $k \geq 1$. It is then easy to check that $\phi\phi'$ is large and (almost) well approximated via the functions $\psi^{(k)}$.

Consider the functions $\Psi^{(k)} = (\Phi^{(k)})^{-1}$, and define $\psi^{(0)}$ to be identically 1, and $\psi^{(k)} := \Psi^{(k)} - \Psi^{(k-1)}$ for $k \geq 1$. We claim that $1/\phi$ is large and (almost) well approximated via the functions $\psi^{(k)}$. Indeed, Properties (a) and (b) are immediately seen to be satisfied. To verify Property (c) in the case when ϕ is well approximated, note that if $f \in V(\mathbb{Z})$ is an element with $p^{2k} \nmid \Delta(f)$, then $\phi^{(k)}(f) = 0$, which implies that $\Phi^{(k)}(f) = \Phi^{(k-1)}(f)$, which in turn implies that $\psi^{(k)}(f) = 1/\Phi^{(k)}(f) - 1/\Phi^{(k)}(f) = 0$ as necessary. The proof in the case when ϕ is almost well approximated is identical. $\square$

Suppose $\chi : \mathbb{Z}_p^2\backslash\{\Delta \neq 0\} \to \mathbb{R}$ is large and (almost) well approximated by periodic functions. Then it is clear that the function $\mathrm{Inv}_\chi := \chi \circ \mathrm{Inv} : V(\mathbb{Z}_p)\backslash\{\Delta = 0\} \to \mathbb{R}$ is also (almost) well approximated. Here, for a ring R, the function $\mathrm{Inv} : V(R) \to R^2$ is defined by setting $\mathrm{Inv}(f) = (I(f), J(f))$. We then have the following consequence of Propositions 5.1, 5.8, and Lemma 5.9

Corollary 5.10 *Let $\chi : \mathbb{Z}_p^2\backslash\{\Delta \neq 0\} \to \mathbb{R}$ be (almost) well approximated by periodic functions. Then $\mathrm{Inv}_\chi \cdot \ell_p/m_p$ is large and (almost) well approximated by periodic functions when $p \geq 3$, and almost well approximated when $p = 2$.*

6 Density estimates and Fourier analysis on $V(\mathbb{Z}/n\mathbb{Z})$

As we will see in §8, the error terms obtained in Section 4 are insufficiently strong for the purposes of summing $\mathrm{PGL}_2(\mathbb{Z})$-orbits on irreducible binary quartic forms weighted by a locally well approximated function. In fact, the obtained error terms in both the main body and the cuspidal region are insufficiently strong. The main goals of this section are: 1: to prove that the previously introduced Dirichlet series $D^{\pm}(\phi, s)$ have analytic continuation to the left of $\Re(s) = 1$, so as to show that their values at 1/2 are well defined, and 2: to obtain cancellation in the Fourier transforms of $\mathrm{PGL}_2(\mathbb{Z}/p^2\mathbb{Z})$-invariant functions on $V(\mathbb{Z}/p^2\mathbb{Z})$. These results will then be combined in §8 with equidistribution techniques to improve the main body count.

This section is organized as follows. In Section 6.1, we obtain upper bounds on the density of the set of elements in $V(\mathbb{Z}_p)$ whose discriminants are divisible by p^{2k} for $k \geq 1$. This will allow us to analytically continue the Dirichlet series $D^{\pm}(\phi, s)$ to the left of $\Re(s) = 1$. Then in Section 6.2, we obtain nontrivial cancellation in the Fourier transform of $\mathrm{PGL}_2(\mathbb{Z}/p^2\mathbb{Z})$-invariant functions on $V(\mathbb{Z}/p^2\mathbb{Z})$.

For a function $\phi\colon V(\mathbb{Z}_p) \to \mathbb{R}$, similar to (8), we define

$$\nu(\phi) := \int_{V(\mathbb{Z}_p)} \phi(v)dv; \qquad \nu_a(\phi) := \int_{V_a(\mathbb{Z}_p)} \phi(v)dv. \tag{27}$$

Above, the measures dv are normalized so that $V(\mathbb{Z}_p)$ and $V_a(\mathbb{Z}_p)$ have volume 1.

6.1 Bounds on the density of elements in $V(\mathbb{Z}_p)$ with small discriminant

For $k \geq 1$, let $\chi_{p^{2k}} : V(\mathbb{Z}/p^{2k}\mathbb{Z}) \to \mathbb{R}$ denote the characteristic function of any set of elements $f \in V(\mathbb{Z}/p^{2k}\mathbb{Z})$ with $\Delta(f) = 0$. In this subsection, we begin by proving the following result.

Proposition 6.1 *We have*

$$\nu(\chi_{p^2}) \ll \frac{1}{p^2}, \qquad \nu(\chi_{p^4}) \ll \frac{1}{p^4}, \qquad \nu(\chi_{p^{2k}}) \ll \frac{k}{p^{3k/2}},$$

for $k \geq 3$.

Proof: We begin with the case $k \geq 3$. Fix $0 \neq a \in \mathbb{Z}_p$, and denote the set of binary quartic forms $f \in V(\mathbb{Z}_p)$ with $a(f) = a$ by $V_a(\mathbb{Z}_p)$. Let $S_a(2k) \subset V_a(\mathbb{Z}_p)$ denote the set of elements $f \in V_a(\mathbb{Z}_p)$ with $p^{2k} \mid \Delta(f)$. To obtain a bound on the density of $S_a(2k)$, we consider the map

$$\begin{array}{rcl} \mathrm{mon}_a : V_a(\mathbb{Z}_p) & \to & V_1(\mathbb{Z}_p) \\ ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4 & \mapsto & x^4 + bx^3y + acx^2y^2 + a^2dxy^3 + a^3ey^4. \end{array}$$

The Jacobian change of variables of this map is clearly a^6, and it is easily seen that $\Delta(\mathrm{mon}_a(f)) = a^6\Delta(f)$. If $p^\ell \parallel a$, then it follows that for $g(x,y) \in \mathrm{mon}_a(S_a(2k))$, we have $p^{2k+6\ell} \mid \Delta(g)$. In particular, we have $\mathrm{mon}_a(S_a(2k)) \subset S_1(2k+6\ell)$.

It is proved in [1, Lemma 5.1] that the density of $S_1(2k+6\ell)$ is $\ll p^{-3k/2-9\ell/2}$. We claim that the density of $\mathrm{mon}_a(S_a(2k)) \subset S_1(2k+6\ell)$ is in fact smaller, and bounded by $\ll p^{-3k/2-5\ell}$. For this, we use the additional fact that every $f(x,y) \in \mathrm{mon}_a(S_a(2k))$ has x^2y^2-coefficient divisible by p^ℓ. The set $S_1(2k+6\ell)$ is preserved by the transformation $f(x,y) \mapsto f(x+ry,y)$ for any $r \in \mathbb{Z}_p$. This yields a surjective map $\mathbb{Z}_p \times S_1^{(0)}(2k+6\ell) \to S_1(2k+6\ell)$, where $S_1^{(0)}(2k+6\ell)$ is the set of elements in $S_1(2k+6\ell)$ whose x^3y-coefficient belongs to $\{0,1,2,3\}$. The density of $S_1^{(0)}(2k+6\ell)$ satisfies the same bound as the density of $S_1(2k+6\ell)$, and is thus $\ll p^{-3k/2-9\ell/2}$. We have the following lemma

Lemma 6.2 *Fix any $f(x,y) \in S_1^{(0)}(2k+6\ell)$. Then the density of $r \in \mathbb{Z}_p$ such that $f(x+ry,y)$ belongs to $\mathrm{mon}_a(S_a(2k))$ is $\ll p^{-\ell/2}$.*

Proof: We prove the lemma in the case when the x^4y-coefficient is 0; the other cases are identical. In this case, we need to upper bound the set of $r \in \mathbb{Z}_p$ such that $p^\ell \mid (6r^2 + c)$ for any fixed $c \in \mathbb{Z}_p$, and the set of such r clearly has density $\ll p^{-\ell/2}$. $\square$

We therefore have

$$\nu(S_a(2k)) = p^{6\ell}\nu(\mathrm{mon}_a(S_a(2k))) \ll p^{6\ell}p^{-\ell/2}\nu(S_1^{(0)}(2k+6\ell)) \ll p^{6\ell}p^{-\ell/2}p^{-3k/2-9\ell/2} = p^{\ell-3k/2}.$$

The required result (for $k \geq 3$) follows immediately by integrating over a.

For $k = 2$ and $k = 3$, it is sufficient to bound the densities of the sets $T_\delta := \{f \in V(\mathbb{Z}_p) : p^\delta \parallel \Delta(f)\}$ for $\delta = 2$, 3, 4, and 5. For this, we use the Jacobian change of variables formula of [14, Proposition 3.11] (also stated in Proposition 8.3) to obtain

$$\nu(T_\delta) \ll \int_{\substack{(I,J)\in\mathbb{Z}_p^2 \\ p^\delta \parallel \Delta(I,J)}} \#\frac{\mathrm{Inv}^{-1}(I,J)}{\mathrm{PGL}_2(\mathbb{Z}_p)} dI dJ. \tag{28}$$

Let $g_{I,J}$ denote the monic cubic polynomial with invariants I and J, and let $R_{I,J}$ denote the corresponding cubic ring. Let $\mathcal{Q}_{I,J}$ denote the set of quartic algebras over $\mathbb{Z}_p$ with resolvent $R_{I,J}$. We have the following result regarding the size of $\mathcal{Q}_{I,J}$.

Lemma 6.3 *We have*

$$\#\frac{\mathrm{Inv}^{-1}(I,J)}{\mathrm{PGL}_2(\mathbb{Z}_p)} \leq \#\mathrm{Stab}_{\mathrm{GL}_2(\mathbb{Z}_p)}(g_{I,J})\#\mathcal{Q}_{I,J}.$$

Proof: The set $\mathcal{Q}_{I,J}$ is in bijection with $\mathrm{GL}_2(\mathbb{Z}_p)\times\mathrm{SL}_3(\mathbb{Z}_p)$-orbits on the set $S_{I,J}$ of pairs $(A,B) \in W(\mathbb{Z}_p)$ such that the ring associated to $4\det(Ax-By)$ is $R_{I,J}$. The map ι defined in (11) gives an injection from $\mathrm{Inv}^{-1}(I,J)$ to $S_{I,J}$. Since PGL_2 is isomorphic to SO_{A_0} (see the discussion following the definition of ι) this induces an injection $\mathrm{PGL}_2(\mathbb{Z}_p)\backslash\mathrm{Inv}^{-1}(I,J) \to \mathrm{SL}_3(\mathbb{Z}_p)\backslash S_{I,J}$. Moreover, the image of this injection is contained in $S'_{I,J}$, the set of pairs $(A,B) \in S_{I,J}$ with monic cubic resolvent. The lemma follows since it is easy to see that the size of a set of $\mathrm{GL}_2(\mathbb{Z}_p)$-equivalent elements in $\mathrm{SL}_3(\mathbb{Z}_p)\backslash S'_{I,J}$ is bounded by $\#\mathrm{Stab}_{\mathrm{GL}_2(\mathbb{Z}_p)}(g_{I,J})$. □

Set $K_{I,J} := R_{I,J}\otimes\mathbb{Q}_p$, and note that there are $O(1)$ étale quartic algebras K_4 over $\mathbb{Q}_p$ with resolvent $K_{I,J}$. Then every element in $\mathcal{Q}_{I,J}$ is a suborder in some such K_4 having index $\leq p^2$, and index $\leq p$ if $\Delta(K_4) \geq p^2$. (This latter condition follows since we are assuming that $\delta \leq 5$.) The previously used result of Nakagawa [36, Theorem 1] implies that there are $O(1)$ such suborders in K_4. The result now follows from (28) in conjunction with [43, Proposition 3.8], which upper bounds the density of pairs (A,B) such that $p^\delta \mid \Delta(x^3+Ax+B)$ by $O(p^{-2})$ when $\delta = 2$ or $\delta = 3$ and by $O(p^{-4})$ when $\delta = 4$ or $\delta = 5$. □

We now collect several consequences of Proposition 6.1. The most important of these is the analytic continuation of $D(\phi,s)$ for large and locally well approximated functions $\phi : V(\mathbb{Z}) \to \mathbb{R}$. To establish this continuation, we need the following lemma.

Lemma 6.4 *Fix an integer $k \geq 1$ and an element $a \in \mathbb{Z}/p^{2k}\mathbb{Z}$. We write $a = up^\ell$, where $u \in \mathbb{Z}/p^{2k}\mathbb{Z}$ is a unit and $\ell \leq 2k$. Then we have the following bounds.*

$$\nu_u(\chi_{p^2}), \nu_{up}(\chi_{p^2}) \ll \frac{1}{p^2}, \quad \nu_{up^2}(\chi_{p^2}) \ll \frac{1}{p}, \quad \nu_{up^\ell}(\chi_{p^4}) \ll \frac{1}{p^{4-\ell}}, \quad \nu_{up^\ell}(\chi_{p^{2k}}) \ll \min\Big(\frac{k}{p^{3k/2-\ell}}, 1\Big),$$

for $k \geq 3$.

Proof: The bound on $\nu_u(\chi_{p^2})$ follows from an analysis of the possible splitting types, while the last two bounds are a direct consequence of Proposition 6.1. We consider the remaining two bounds.

We first consider $\nu_{up^2}(\chi_{p^2})$. Let $f = (a,b,c,d,e) \in V(\mathbb{Z}_p)$ and assume $p^2 \mid a$ Then by (24),

$$\Delta(f) \equiv b^2(c^2d^2 + 18bcde - 4bd^3 - 4c^3e - 27b^2e^2) \pmod{p^2}.$$

Note that $\Delta_3 := c^2d^2 + 18bcde - 4bd^3 - 4c^3e - 27b^2e^2$ is the discriminant of $bx^3 + cx^2y + dxy^2 + ey^3$. The density of $(b,c,d,e) \in \mathbb{Z}_p^4$ with $p^2 \mid b^2$ is $O(p^{-1})$, while with $p^2 \mid \Delta_3$ is $O(p^{-2})$, as desired.

Finally we consider $\nu_{up}(\chi_{p^2})$. Let $f = (a,b,c,d,e) \in V(\mathbb{Z}_p)$ with $a = up$ for some $u \in \mathbb{Z}_p^\times$, and assume $p^2 \mid \Delta(f)$. Then $(f \pmod p) \in V(\mathbb{F}_p)$ has a multiple root in $\mathbb{P}^1$. Note that $(1:0) \in \mathbb{P}^1(\mathbb{F}_p)$ is a root. Suppose $(1:0)$ is a multiple root. Then $p \mid b$. Thus by (25),

$$\Delta(f) \equiv 4ac^3(4ce - d^2) \pmod{p^2}.$$

Thus $p^2 \mid \Delta(f)$ if and only if $p \mid c$ or $p \mid (4ce - d^2)$. The $(b,c,d,e) \in \mathbb{Z}_p^4$ satisfying the conditions have density $O(p^{-2})$ as desired. Suppose $(1:0) \in \mathbb{P}^1(\mathbb{F}_p)$ is a simple root. Then $(f \pmod p)$ has one multiple root in $\mathbb{P}^1(\mathbb{F}_p)$. Further suppose that the multiple root is $(0:1)$. Then $p \mid d$ and $p \mid e$. Thus by (25) with $\Delta(a,b,c,d,e) = \Delta(e,d,c,b,a)$, we have

$$\Delta(f) \equiv -4b^2c^3e \pmod{p^2}.$$

So $p^2 \mid \Delta(f)$ if and only if $p^2 \mid e$ or $p \mid bc$. The density of $(b,c,d,e) \in \mathbb{Z}_p^4$ satisfying the conditions is $O(p^{-3})$. There are p possibilities for the multiple root of $(f \pmod p)$, and the densities all coincide. Thus the total density of $(b,c,d,e) \in \mathbb{Z}_p^4$ for $p^2 \mid \Delta(f)$ is $O(p^{-2})$, completing the proof. $\square$

Let $\phi : V(\mathbb{Z}) \to \mathbb{R}$ be a large and locally well approximated function via $\phi(\cdot) = \sum_n \phi(n;\cdot)$. We clearly have the equality $D^\pm(\phi,s) = \sum_n D^\pm(\phi(n;\cdot),s)$ in the region $\Re(s) > 1$ of absolute convergence. Recall that $\phi(n;\cdot)$ is defined by congruence conditions modulo n^2, and is supported on the set of elements $f \in V(\mathbb{Z}_p)$ satisfying $n^2 \mid c\Delta(f)$, for some positive integer c. We have the following result on $D(\phi(n;\cdot),s)$:

Proposition 6.5 *Keep the above notation. The functions $D^\pm(\phi(n;\cdot),s)$ have analytic continuation to the whole plane, with at most a simple pole at $s = 1$ with residue $r_n = \nu(\phi(n;\cdot))$. Moreover, we have the bound*

$$\left| D^\pm(\phi(n;\cdot),s) - \frac{r_n}{s-1} \right| \ll_\epsilon \frac{n^\epsilon}{n_1^{3/2+\sigma/2} n_2^{4\sigma} m_3^{1/3+\sigma/6}}(1+|t|), \tag{29}$$

for $s = \sigma + it$ with $\sigma > 1/3$, where we write $n = n_1 n_2^2 m_3$ with n_1 and n_2 squarefree, m_3 cubefull ($p \mid m_3$ implies $p^3 \mid m_3$), and n_1, n_2, m_3 pairwise relatively prime.

Proof: We write $\psi := \phi(n,\cdot)$. We break up the Dirichlet series into summands corresponding to each divisor d of n^2:

$$D^\pm(\psi,s) = \sum_{a \geq 1} \frac{\nu_{\pm a}(\psi)}{a^s} = \sum_{d \mid n^2} \sum_{\substack{a \geq 1 \\ (a,n^2)=d}} \frac{\nu_{\pm a}(\psi)}{a^s} = \sum_{d \mid n^2} \frac{1}{d^s} \sum_{\substack{a \geq 1 \\ (a,n^2)=1}} \frac{\nu_{\pm ad}(\psi)}{a^s}.$$

Since ψ is $\mathrm{PGL}_2(\mathbb{Z}/n^2\mathbb{Z})$ invariant, it follows that $\nu_{\pm a}(\psi) = \nu_{\pm u^2 a}(\psi)$ for each $u \in \mathbb{Z}$ with $(u,n^2) = 1$. It follows that for any d, the value of $\nu_{\pm ad}(\psi)$ only depends on the residue class of a modulo n^2/d. Moreover, the inner product of $\nu(\psi|_{ad})$ with a character χ of $(\mathbb{Z}/(n^2/d)\mathbb{Z})^\times$ is 0 unless χ is

quadratic. Therefore, each summand in the rightmost term of the above equation is a weighted sum of $O(n^\epsilon)$ quadratic Dirichlet L-functions $L(s,\chi)$, where χ is a quadratic character on $(\mathbb{Z}/(n^2/d)\mathbb{Z})^\times$. The analytic continuation of $D^\pm(\psi,s)$ follows immediately.

To prove the bound, we start with observing that when χ is a (quadratic) character on $(\mathbb{Z}/(n^2/d)\mathbb{Z})^\times$, the conductor of $L(s,\chi)$ is $\ll \mathrm{rad}(n^2/d)$, where rad denotes the radical function. Applying the convexity bound, we obtain for $s=\sigma+it$ with $\sigma\in(0,1-\delta)$ for any positive δ, the following estimate:

$$\Big|\sum_{\substack{a\geq 1\\(a,n^2)=1}}\frac{\nu_{\pm ad}(\psi)}{a^s}\Big| \ll_\epsilon \max_a|\nu_{\pm ad}(\psi)|\mathrm{rad}(n^2/d)^{1/2-\sigma/2+\epsilon}(1+|t|).$$

Next, we bound $\max_a|\nu_{\pm ad}(\psi)|$ using Lemma 6.4. As above, we write $n=n_1n_2^2m_3$. We also write $d=d_1d_2d_3$, where $d_1\mid n_1^2$, $d_2\mid n_2^4$, and $d_3\mid m_3^2$. Therefore, we have

$$\begin{aligned}
&\frac{1}{d^\sigma}\max_a|\nu_{\pm ad}(\psi)|\mathrm{rad}(n^2/d)^{1/2-\sigma/2+\epsilon}\\
&\ll_\epsilon n^\epsilon\Big(\prod_{p^2\mid d_1}p^{-1-2\sigma}\prod_{p\mid n_1^2/d_1}p^{-3/2-\sigma/2}\Big)\cdot\Big(\frac{d_2^{1-\sigma}}{n_2^4}\prod_{p\mid n_2^4/d_2}p^{1/2-\sigma/2}\Big)\cdot d_3^{-\sigma}\min\Big(\frac{d_3}{m_3^{3/2}},1\Big)m_3^{(1-\sigma)/6}\\
&\ll_\epsilon \frac{n^\epsilon}{n_1^{3/2+\sigma/2}n_2^{4\sigma}m_3^{1/3+\sigma/6}},
\end{aligned}$$

where the last estimate follows from a straightforward check. The result now follows noting that the number of divisors of n^2 is bounded by $O_\epsilon(n^\epsilon)$. □

The above proposition has the following immediate corollary, which follows by noting that the sum over n of the right hand side in (29) converges absolutely for $\sigma>1/3$ and that the sum of the residues $r_n(=\nu(\phi(n;\cdot))$ also converges absolutely.

Corollary 6.6 *Let $\phi: V(\mathbb{Z})\to\mathbb{R}$ be a large and locally well approximated function. Then $D^\pm(\phi,s)$ has an analytic continuation to $\Re(s)>1/3$ with only a possible simple pole at $s=1$. In particular, the value of $D(\phi,1/2)$ is well defined.*

6.2 Equidistribution of strongly invariant sets in $V(\mathbb{Z}/p^2\mathbb{Z})$

For a positive integer n, and a function $\phi: V(\mathbb{Z}/n\mathbb{Z})\to\mathbb{C}$, let $\widehat{\phi}: V^*(\mathbb{Z}/n\mathbb{Z})\to\mathbb{C}$ be its Fourier transform normalized as follows:

$$\widehat{\phi}(h)=\frac{1}{n^5}\sum_{f\in V(\mathbb{Z}/n\mathbb{Z})}\phi(f)e\Big(\frac{[f,h]}{n}\Big)=\frac{1}{n^5}\sum_{f\in V(\mathbb{Z}/n\mathbb{Z})}\phi(f)\exp\Big(\frac{2\pi i[f,h]}{n}\Big), \tag{30}$$

where $[\cdot,\cdot]$ is the natural bilinear form $V(\mathbb{Z}/n\mathbb{Z})\times V^*(\mathbb{Z}/n\mathbb{Z})\to\mathbb{Z}/n\mathbb{Z}$. For the rest of the section, we take $n=p^k$ to be a prime power. We need our result only for $k=2$, but since the argument is identical we work for general k. Let $S\subset V(\mathbb{Z}/n\mathbb{Z})$ be a strongly invariant set, i.e., its characteristic function χ_S is strongly invariant. Clearly we have $\widehat{\chi_S}(0)=|S|/n^5$ and we have the "trivial" bound $|\widehat{\chi_S}(h)|\leq|S|/n^5$ for any $h\in V^*(\mathbb{Z}/n\mathbb{Z})$. (In fact, these trivial bounds clearly hold for every subset S.) The first result of this subsection improves upon this trivial bound for strongly invariant sets in $V(\mathbb{Z}/p^k\mathbb{Z})$ not intersecting $pV(\mathbb{Z}/p^k\mathbb{Z})$. To simplify notation for the next proposition, we write

V, V^*, GL_1, PGL_2 for $V(\mathbb{Z}/p^k\mathbb{Z})$, $V^*(\mathbb{Z}/p^k\mathbb{Z})$, $\mathrm{GL}_1(\mathbb{Z}/p^k\mathbb{Z})$, and $\mathrm{PGL}_2(\mathbb{Z}/p^k\mathbb{Z})$, respectively. We state the result in terms of bounding the *orbital exponential sum* $\mathcal{G}_{p^k}(f,h)$ associated to $f \in V$ and $h \in V^*$ defined as

$$\mathcal{G}_{p^k}(f,h) := \frac{1}{|\mathrm{GL}_1||\mathrm{PGL}_2|} \sum_{t\in \mathrm{GL}_1} \sum_{g\in \mathrm{PGL}_2} \exp\Big(2\pi i \cdot \frac{t^2[gf,h]}{p^k}\Big). \tag{31}$$

Then we prove the following proposition.

Proposition 6.7 *Let notation be as above, and let $f \in V$ be an element which is not a multiple of p. Then we have*

$$\mathcal{G}_{p^k}(f,h) \ll \begin{cases} 1 & h = 0, \\ p^{-1/2} & h \in p^{k-1}V^*, h \neq 0, \\ p^{-1} & h \notin p^{k-1}V^*. \end{cases} \tag{32}$$

Proof: The classical quadratic Gauss sum is defined by

$$\mathcal{Q}_{p^k}(a) := \frac{1}{|\mathrm{GL}_1|} \sum_{t\in \mathrm{GL}_1} \exp\Big(2\pi i \cdot \frac{t^2 a}{p^k}\Big), \qquad a \in \mathbb{Z}/p^k\mathbb{Z}. \tag{33}$$

The explicit formula of the quadratic Gauss sum is well known. In particular, we have

$$\mathcal{Q}_{p^k}(a) \ll \begin{cases} 1 & a = 0, \\ p^{-1/2} & p^{k-1} \mid a, a \neq 0, \\ p^{-1} & p^{k-1} \nmid a. \end{cases} \tag{34}$$

Then by definition, the orbital exponential sum is expressed as

$$\mathcal{G}_{p^k}(f,h) = \frac{1}{|\mathrm{PGL}_2|} \sum_{g\in \mathrm{PGL}_2} \mathcal{Q}_{p^k}([gf,h]). \tag{35}$$

For simplicity we first consider the case $h \notin pV^*$. We show that except for $O(p^{-1})$-proposition of g in PGL_2, $[gf,h]$ is not divisible by p. Then (34) implies that $\mathcal{G}_{p^k}(f,h) \ll p^{-1}$. To study $[gf,h]$, it will be convenient to have an explicit description of V^*. Since the estimate (32) holds automatically for $p = 2,3$, we assume $p \neq 2,3$. For $f = (f_0,f_1,f_2,f_3,f_4), h = (h_0,h_1,h_2,h_3,h_4) \in V$, let $[f,h] = f_0h_0 + 4^{-1}f_1h_1 + 6^{-1}f_2h_2 + 4^{-1}f_3h_3 + f_4h4$. This bilinear form is PGL_2-invariant in the sense that $[gf,h] = [f,g^Th]$ holds for all $f,h \in V$ and $g \in \mathrm{PGL}_2$, where g^T is the matrix transpose of g. Via the map $V \ni h \mapsto [\cdot,h] \in V^*$ we identify V^* with V, and regard h as an element in V.

Let $\mathrm{PGL}_2^\bullet \subset \mathrm{PGL}_2$ consists of elements whose $(1,2)$-entry (i.e., the upper right entry) is in $(\mathbb{Z}/p^k\mathbb{Z})^\times$. Then the proportion of elements $g \in \mathrm{PGL}_2$ not in $\mathrm{PGL}_2^\bullet$ is $O(p^{-1})$, and hence has a negligible contribution to (32). Any $g \in \mathrm{PGL}_2^\bullet$ is expressed uniquely in the form

$$g = l_m a_s w l_n := \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix} \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}$$

for $m,n \in \mathbb{Z}/p^k\mathbb{Z}$ and $s \in (\mathbb{Z}/p^k\mathbb{Z})^\times$. We consider $[gf,h] = [a_s w l_n f, l_m^T h]$. The x^4-coefficient of wl_nf is $(wl_nf)(1,0) = f(n,1)$, and the x^4-coefficient of l_m^Th is $(l_m^Th)(1,0) = h(1,m)$. Therefore, we have

$$[gf,h] = \frac{1}{s^2}(f(n,1)h(1,m)s^4 + p_{m,n}(s)). \tag{36}$$

for a degree ≤ 3 polynomial $p_{m,n}(s)$ in s with coefficients in $(\mathbb{Z}/p^k\mathbb{Z})[m,n]$. Since $f \pmod p$ and $h \pmod p$ are both not the zero form by our assumption, they have at most 4 roots in $\mathbb{P}^1(\mathbb{F}_p)$. Therefore, except for $O(p^{-1})$-propotion of g, $f(n,1)h(1,m)$ is not divisible by p. For such majority of g, the reduction modulo p of the polynomial of s in (36) is of degree 4, and has at most 4 roots in $\mathbb{F}_p$. Therefore, except for $O(p^{-1})$-propotion of s (and hence of g), $[gf,h]$ is not divisible by p. This finishes the proof of (32) for $h \notin pV$.

This argument works for general $h \neq 0$: Let us write $h = p^l h'$, where $l < k$ and not all the coefficients of h' are divisible by p. Then $[gf,h] = p^l[gf,h']$. (Here h' is well defined as an element in $V(\mathbb{Z}/p^{k-l}\mathbb{Z})$ and we also regard $[gf,h'] \in \mathbb{Z}/p^{k-l}\mathbb{Z}$.) Exactly by the same argument, we confirm that except for $O(p^{-1})$-proportion of g, $[gf,h']$ is not divisible by p. Thus again (34) implies (32), concluding the proof of the proposition. □

Proposition 6.7 combined with the first bound of Proposition 6.1 immediately implies the following result.

Corollary 6.8 *Let p be a prime, and let $\phi : V(\mathbb{Z}/p^2\mathbb{Z}) \to \mathbb{R}$ be a strongly invariant bounded function whose support is contained within the set of elements $f \in V(\mathbb{Z}/p^2\mathbb{Z})$ with $\Delta(f) = 0$. Then we have*

$$\widehat{\phi}(0) \ll \frac{1}{p^2}, \qquad \widehat{\phi}(ph) \ll \frac{1}{p^{2+1/2}}, \qquad \widehat{\phi}(h) \ll \frac{1}{p^3},$$

for $h \in V^*(\mathbb{Z}/p^2\mathbb{Z}), h \notin pV^*(\mathbb{Z}/p^2\mathbb{Z})$.

7 Uniformity estimates

Recall that for a positive integer m, the set of generic elements in $V(\mathbb{Z})$ whose discriminants are divisible by m^2 is denoted by $\mathcal{W}_m$. In this section, we prove Theorem 3 by giving a bound for the number of $\mathrm{PGL}_2(\mathbb{Z})$-orbits on $\mathcal{W}_m$ (on average over m) having bounded height.

7.1 Fibering over quartic fields

Let $f \in V(\mathbb{Z})$ be an integral binary quartic form, and let (Q_f, C_f, α_f) be the triple corresponding to the $\mathrm{PGL}_2(\mathbb{Z})$-orbit of f under the bijction of Theorem 3.8. If f is generic, then Q_f and C_f are integral domains, and thus are orders in a quartic field and a cubic field, respectively. We define L_f to be $Q_f \otimes \mathbb{Q}$, and let $\mathcal{O}_f$ denote the maximal order in L_f. Let R_f be the (unique) cubic resolvent ring of $\mathcal{O}_f$. Define the *index* $\mathrm{ind}(f)$ of f to be the index of Q_f in $\mathcal{O}_f$ (equivalently, the index of C_f in R_f), and define the *strongly divisible factor* $\mathrm{sd}(f)$ of f to be the product of primes p such that $p^2 \mid \Delta(\mathcal{O}_f)$. For positive integers m_1 and m_2 we define

$$\mathcal{W}_{m_1,m_2} := \{f \in \mathcal{W}_{m_1 m_2} : m_1 \mid \mathrm{ind}(f),\, m_2 \mid \mathrm{sd}(f)\}.$$

For positive real numbers M_1 and M_2, let $N(X; M_1, M_2)$ denote the following quantity:

$$N(X; M_1, M_2) := \sum_{\substack{m_1 \in [M_1, 2M_1] \\ m_2 \geq M_2}} |\mu(m_2)| \# \frac{\{f \in \mathcal{W}_{m_1,m_2} : H(f) < X\}}{\mathrm{PGL}_2(\mathbb{Z})}. \tag{37}$$

Note that $\Delta(f)/(\mathrm{ind}(f)^2\mathrm{sd}(f)^2) = \Delta(\mathcal{O}_f)/\mathrm{sd}(f)^2$ is squarefree upto some bounded product C_6 of 2's and 3's. As a consequence, we have the inclusion $\mathcal{W}_m \subset \bigcup_{m \mid C_6 m_1 m_2} \mathcal{W}_{m_1,m_2}$, where the m_2 are

squarefree. Therefore, for a real number $M > 1$, we have

$$\sum_{m \geq M} \# \frac{\{f \in \mathcal{W}_m : H(f) < X\}}{\mathrm{PGL}_2(\mathbb{Z})} \ll \sum_{\substack{2^k = M_1 \leq X^{1/2} \\ M_2 = M/(2C_6 M_1)}} N(X; M_1, M_2), \tag{38}$$

where M_1 ranges over powers of 2 betewen 1 and $X^{1/2}$, while $M_2 = M/(2C_6 M_1)$. In particular, the sum over M_1, M_2 has length $O(\log(X))$.

We repackage the count $N(X; M_1, M_2)$ by fibering over maximal quartic orders. To this end, let $V(\mathbb{Z})^{\mathrm{gen}}_{H<X}$ denote the set of generic elements in $V(\mathbb{Z})$ having height less than X. We consider the map

$$\begin{aligned} \Psi_X : \mathrm{PGL}_2(\mathbb{Z}) \backslash V(\mathbb{Z})^{\mathrm{gen}}_{H<X} &\to \{\text{maximal quartic order with its (unique) cubic resolvent order}\} \\ f &\mapsto (\mathcal{O}_f, R_f). \end{aligned} \tag{39}$$

The ring C_f is a monogenized cubic subring of R_f, with monogenizer $\alpha_f \in R_f$. For a cubic ring R, we let R^{red} denote the set of elements $\alpha \in R$ such that $\mathrm{Tr}(\alpha) \in \{-1, 0, 1\}$. Since the monogenizer of C_f is defined upto addition by an element of $\mathbb{Z}$, we may in fact assume that $\alpha_f \in R_f^{\mathrm{red}}$. We introduce a *height* on R_f by setting $h(\alpha) := \max(|\alpha'|_v)$, where α' is the unique $\mathbb{Z}$-translate of α with $\alpha' \in R^{\mathrm{red}}$, and v ranges over the archimedian valuations of $K = C \otimes \mathbb{Q}$; when v is a complex place, we take $|x|_v$ to be the absolute value of x. (An equivalent height is to take the length of α' in $R_f \otimes_{\mathbb{Z}} \mathbb{R}$.) A standard computation reveals that we have $h(\alpha_f) \ll H(f)^{1/6}$.

We study the fibers of the map (39). Clearly, we have an injection

$$\Psi_X^{-1}(\mathcal{O}, R) \hookrightarrow \{(Q, \alpha) : Q \subset \mathcal{O},\, \alpha \in R^{\mathrm{red}},\, h(\alpha) \ll X^{1/6},\, [\mathcal{O} : Q] = [R : \mathbb{Z}[\alpha]]\}. \tag{40}$$

In fact, the above map remains an injection even if we also specify that $\mathbb{Z}[\alpha]$ is a cubic resolvent ring of Q. However, we will not be using this fact in our estimates. Recall that by Bhargava's work [8], the set of pairs $(\mathcal{O}, R)$, where $\mathcal{O}$ is quartic ring, and R is a cubic resolvent ring of $\mathcal{O}$, is in bijection with $\mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$-orbits on $W(\mathbb{Z})$. We denote the set of elements in $W(\mathbb{Z})$ corresponding to pairs $(\mathcal{O}, R)$, where $\mathcal{O}$ is a maximal integral domain and R is an integral domain, by $W(\mathbb{Z})^{\mathrm{mg}}$. For $(A, B) \in W(\mathbb{Z})^{\mathrm{mg}}$, let $w_X(A, B)$ to be the size of the right hand side of (40), where (where (A, B) to the pair of rings $(\mathcal{O}, R)$. For positive real numbers Y and M_2, let $S_W(Y, M_2)$ denote the set of $\mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$-orbits on the set of elements $(A, B) \in W(\mathbb{Z})^{\mathrm{mg}}$ such that $|\Delta(A, B)| \ll Y$, and $m_2^2 \mid \Delta(A, B)$ for some squarefree positive integer $m_2 \geq M_2$. Then we have the following result.

Proposition 7.1 *With notation as above, we have*

$$N(X; M_1, M_2) \ll \sum_{(A,B) \in S_W(X/M_1^2, M_2)} w_X(A, B). \tag{41}$$

Proof: For integers $m_1 > M_1$ and $m_2 > M_2$ Let $f \in \mathcal{W}_{m_1, m_2}$ be a binary quartic form with $H(f) \leq X$. Note that we have $|\Delta(Q_f)| = |\Delta(f)| \ll H(f)$. Hence, we have $|\Delta(R_f)| = |\Delta(\mathcal{O}_f)| \ll X/M_1^2$, and moreover, $m_2^2 \mid \Delta(\mathcal{O}_f)$. Hence, $(\mathcal{O}_f, R_f) \in S_W(X/M_1^2, M_2)$. Therefore, we have

$$N(X; M_1, M_2) \ll \sum_{(A,B) \in S_W(X/M_1^2, M_2)} \# \Psi_X^{-1}(A, B) \ll \sum_{(A,B) \in S_W(X/M_1^2, M_2)} w_X(A, B),$$

as needed. $\square$

7.2 Preliminary results on cubic rings

Evaluating $w_X(A,B)$ requires counting elements $\alpha \in R^{\mathrm{red}}$, where R is a cubic ring. To this end, we collect some basic results on cubic rings and their shapes.

We begin by introducing some notation. Let g be an irreducible integral binary cubic form, whose $\mathrm{GL}_2(\mathbb{Z})$-orbit corresponds to the cubic order R under the Dalone–Faddeev parametrization. Let K denote the fraction field of R. We identify $R \otimes \mathbb{R}$ with $\mathbb{R}^3$ by choosing the isomorphism $\mathbb{C} \cong \mathbb{R}^2$ given by $a+ib \mapsto (a+b, a-b)$ when K is complex. The embedding $R \to R \otimes \mathbb{R}$ maps R into a lattice. Denote the lengths of successive minima of this lattice by 1 (corresponding the to element $1 \in R$), $\ell_1(R)$, and $\ell_2(R)$. (Here, we have normalized the length on $R \otimes \mathbb{R}$ such that $1 \in R$ has length 1 in $R \otimes \mathbb{R}$.) We define the *skewness* of R to be $\mathrm{sk}(R) := \ell_2(R)/\ell_1(R)$. We begin with the following lemma.

Lemma 7.2 *Let R be a cubic order. Then we have*

$$\ell_1(R) \asymp \frac{|\mathrm{Disc}(R)|^{\frac{1}{4}}}{\mathrm{sk}(R)^{\frac{1}{2}}}; \quad \ell_2(R) \asymp |\mathrm{Disc}(R)|^{\frac{1}{4}}\mathrm{sk}(R)^{\frac{1}{2}}; \quad \ell_2(R) \ll \ell_1(R)^2; \quad 1 \leq \mathrm{sk}(R) \ll |\mathrm{Disc}(R)|^{\frac{1}{6}}.$$

Proof: First note that by Minkowski's theorem, we have $\ell_1(R)\ell_2(R) \asymp \sqrt{|\mathrm{Disc}(R)|}$. The first two claims follow immediately from this and the definition of $\mathrm{sk}(R)$. Suppose $\beta \in R \backslash \mathbb{Z}$ is the element having length $\ell_1(C)$. Then β^2 has length $\asymp \ell_1^2(R)$. Since $\ell_2(R)$ is the second successsive minima, we have $\ell_1(R)^2 \gg \ell_2(R)$, which yields the third claim. The final claim is a consequence of the first three claims. $\square$

Next, recall that an integral irreducible binary cubic form $g(x,y) \in U(\mathbb{Z})$ corresponds to a cubic order R, together with a basis $\{\overline{\beta_1}, \overline{\beta_2}\}$ of $R/\mathbb{Z}$. We say that $g(x,y)$ is *almost Minkowski reduced* if the basis $\{\beta_1, \beta_2\}$ is almost Minkowski reduced, where we say that some set of linearly independent vectors in $\mathbb{R}^n$ is *almost Minkowski reduced* if the angles between every pair of them is bounded below by a positive absolute constant. Note that if $\{1, \beta_1, \beta_2\}$ is an almost Minkowski reduced basis of R, then $\{\overline{\beta_1}, \overline{\beta_2}\}$ is an almost Minkowski reduced basis of $R/\mathbb{Z}$. Conversely, if $\{\overline{\beta_1}, \overline{\beta_2}\}$ is an almost Minkowski reduced basis of $R/\mathbb{Z}$, then $\overline{\beta_1}$ and $\overline{\beta_2}$ admit lifts to β_1 and β_2, respectively, such that $\{1, \beta_1, \beta_2\}$ is an almost Minkowski reduced basis of R. Indeed, to ensure this, β_1 and β_2 should simply be taken to be elements in R^{red}.

The notion of being almost Minkowsky reduced can be generalized to elements of $U(\mathbb{R})$: Namely, given $h(x,y) \in U(\mathbb{R})$ with nonzero discriminant, an analogue of the Delone–Faddeev parametrization over $\mathbb{R}$ (see for example [12]) yields an étale cubic algebra R_h over $\mathbb{R}$ along with a basis $\{\overline{\beta_1}, \overline{\beta_2}\}$ for $R_h/\mathbb{R}$. The space $R_h/\mathbb{R}$ can be naturally identified with the set of traceless elements in R_h, and this identification allows us to restrict the standard norm and inner product on $\mathbb{R}^3$ to $R_h/\mathbb{R}$. We say that $h(x,y)$ is *almost Minkowski reduced* if $\{\overline{\beta_1}, \overline{\beta_2}\}$ is almost Minkowski reduced. We define $\overline{\mathrm{sk}}(h)$ to be $|\overline{\beta_1}|/|\overline{\beta_2}|$, and note that if $h(x,y) \in U(\mathbb{Z})$, corresponding to the cubic order R, is almost Minkowski reduced, then $\overline{\mathrm{sk}}(h) \asymp \mathrm{sk}(R)$. We now have the following result.

Lemma 7.3 *Let $B \subset U(\mathbb{R})$ be a bounded set whose boundary does not intersect the discriminant zero locus. Then every element in $\mathcal{F} \cdot B$ is almost Minkowski reduced. Moreover, for $\gamma = utk \in \mathcal{F}$ and $h \in B$, we have $\mathrm{sk}(\gamma \cdot h) \asymp t^2$.*

Proof: We start by noting three facts. First, the assumptions on B ensure that the angles between the two basis vectors corresponding to each element of B are uniformly bounded away from 0. Hence

every element in B is almost Minkowski reduced. Second, let $\gamma \in \mathrm{GL}_2(\mathbb{R})$ and let $h(x,y) \in U(\mathbb{R})$ be an element corresponding to the basis $\{\overline{\beta_1}, \overline{\beta_2}\}$ of R_h. Then $\{\gamma\overline{\beta_1}, \gamma\overline{\beta_2}\}$ is the basis corresponding to $\gamma \cdot h$. Third, if $\{\overline{\beta_1}, \overline{\beta_2}\}$ is almost Minkowski reduced, then so is $\{\gamma\overline{\beta_1}, \gamma\overline{\beta_2}\}$ for every $\gamma \in \mathcal{F}$. The first assertion follows from these three facts. The second assertion is an immediate consequence of the second fact. $\square$

We conclude this subsection by explaining how these results can be used to help evaluate the right hand side of (41). Let $g(x,y)$ be an irreducible integral binary cubic form corresponding to the cubic order R be a cubic order and basis $\{1, \beta_1, \beta_2\}$, with $\beta_1, \beta_2 \in R^{\mathrm{red}}$. There is a natural bijection between R^{red} and $\mathbb{Z}^2$, where we map $(r,s) \in \mathbb{Z}^2$ to the unique $\mathbb{Z}$-translate α of $r\beta_1 + s\beta_2$ such that $\mathrm{Tr}(\alpha) \in \{-1, 0, 1\}$. Suppose the basis $\{1, \beta_1, \beta_2\}$ is almost Minkowsky reduced. Then the condition $h(\alpha) \ll X^{1/6}$ translates to $|r| \ll X^{1/6}/\ell_1(R)$ and $|s| \ll X^{1/6}/\ell_2(R)$. Suppose $(A,B) \in W(\mathbb{Z})^{\mathrm{mg}}$ with cubic resolvent $g(x,y)$ corresponds to the pair $(\mathcal{O}, R)$. Assume that $g(x,y)$ corresponds to an almost Minkowski reduced basis $\{1, \beta_1, \beta_2\}$ of R. The quantity $w_X(A,B)$ is equal to the number of pairs (Q, α), where Q is a suborder of $\mathcal{O}$, $\alpha \in R$ with $h(\alpha) \ll X^{1/6}$ and $[\mathcal{O} : Q] = [R : \mathbb{Z}[\alpha]]$. The height condition implies that there exist integers r_1 and r_2 with $|r_1| \ll X^{1/6}/\ell_1(R)$ and $|r_2| \ll X^{1/6}/\ell_2(R)$ such that $\alpha = r_1\beta_1 + r_2\beta_2 + n$ for some integer n. Moreover, since $g(x,y)$ is the index form on $R/\mathbb{Z}$, it follows that $[R : \mathbb{Z}[\alpha]] = |g(r_1, r_2)|$. Therefore, with notation as above, we have the following bound:

$$w_X(A,B) \ll \sum_{\substack{(r_1,r_2)\in\mathbb{Z}^2 \\ |r_1|\ll X^{1/6}/\ell_1(R) \\ |r_2|\ll X^{1/6}/\ell_2(R)}} \#\{Q \subset \mathcal{O} : [\mathcal{O} : Q] = |g(r_1, r_2)|\}. \tag{42}$$

We turn now to the right hand side of (41). It is necessary to sum over $(A,B) \in S_W(X/M_1^2, M_2)$. Such an element (A,B) corresponds to a pair $(\mathcal{O}, R)$, where $\mathcal{O}$ is a maximal quartic order, and R is the cubic resolvent of $\mathcal{O}$. We will break the ranges of the possible values of $|\Delta(R)|$ and $\mathrm{sk}(R)$ into dyadic ranges. By Lemma 7.2, a pair of such dyadic ranges $[Y, 2Y]$ and $[Z, 2Z]$ determines the sizes of $\ell_1(R)$ and $\ell_2(R)$ up to absolutely bounded multiplicative constants. Furthermore, Y will be bounded above by X/M_1^2.

Let $S_W(Y, Z; M_2)$ denote the set of $\mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$-orbits on the set of elements $(A,B) \in W(\mathbb{Z})^{\mathrm{mg}}$, corresponding to $(\mathcal{O}, R)$, such that $Y \le |\Delta(A,B)| < 2Y$, $Z \le \mathrm{sk}(R) < 2Z$, and the discriminant of (A,B) is strongly divisible by m_2^2 for some squarefree positive integer $m_2 \ge M_2$. Then an application of Bhargava's averaging method in [9] yields an absolutely bounded ball $\mathcal{B} \subset W(\mathbb{R})$, with support bounded away from the discriminant-0 locus of $W(\mathbb{R})$, such that

$$\sum_{(A,B)\in S_W(Y,Z;M_2)} w_X(A,B) \ll \int_{\substack{\lambda \asymp Y^{1/12} \\ t\asymp Z^{1/2} \\ s_1,s_2\gg 1}} \Big(\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}_{M_2}} w_X(A,B)\Big) \frac{d^\times s d^\times t d^\times \lambda}{s_1^6 s_2^6 t^2}, \tag{43}$$

where $s = (s_1, s_2)$, $d^\times s := d^\times s_1 d^\times s_2$, (t, s_1, s_2) is the diagonal element $(g_2, g_3) \in \mathrm{GL}_2(\mathbb{R}) \times \mathrm{SL}_3(\mathbb{R})$ with $g_2 = \mathrm{diag}(t^{-1}, t)$ and $g_3 = \mathrm{diag}(s_1^{-2}s_2^{-1}, s_1 s_2^{-1}, s_1 s_2^2)$, and $W(\mathbb{Z})^{\mathrm{mg}}_{M_2}$ denotes the subset of $W(\mathbb{Z})^{\mathrm{mg}}$ whose discriminants are strongly divisible by m_2^2 for some squarefree $m_2 \ge M_2$. In the next subsection, we use (42) and (43) to obtain the uniformity estimate.

7.3 Bounds in the small Y range

Let X, Y, and Z be positive real numbers with $Y \ll X/M_1^2$ and $Z \ll Y^{1/6}$. (This bound on Z comes from Lemma 7.2.) In this subsection, we will break up our uniformity estimate into three

parts (the sum over r_1, r_2 with $r_1 r_2 \neq 0$, the sum over r_2 with $r_1 = 0$, and the sum over r_1 with $r_2 = 0$). For the first two parts to be nonzero, we will show that we must have $Y \ll X^{2/3}$. In this subsection, we obtain bounds on these two parts.

We begin with the following lemma. Recall that we define $N(k)$ for positive integers k in Proposition 4.4. For convenience we define it for negative integers k in the same manner. In the sequel we use the following bound.

Lemma 7.4 *We have*

$$\sum_{\substack{0\neq|d|\ll D\\ 0\neq|r|\ll R}} N(dr^3) \ll_\epsilon D^{1+\epsilon}R^{2+\epsilon}.$$

The above lemma follows from standard methods, and we omit the proof.

For $\lambda \asymp Y^{1/12}$, $t \asymp Z^{1/2}$, and $s_1, s_2 \gg 1$, the cubic resolvent ring R associated to every element in $(t, s_1, s_2)\lambda\mathcal{B} \cap W(\mathbb{Z})^{\mathrm{mg}}$ satisfies (by design) $|\Delta(R)| \asymp Y$ and $\mathrm{sk}(R) \asymp Z$. Therefore, from Lemma 7.2, we also have $\ell_1(R) \asymp Y^{1/4}/Z^{1/2}$ and $\ell_2(R) \asymp Y^{1/4}Z^{1/2}$. We define the quantities

$$R_1 := X^{1/6}Z^{1/2}/Y^{1/4}; \qquad R_2 := X^{1/6}/(Y^{1/4}Z^{1/2}). \tag{44}$$

Then the ranges of $|r_1|$ and $|r_2|$ in the right hand side of (42), for every $(A, B) \in (t, s_1, s_2)\lambda\mathcal{B} \cap W(\mathbb{Z})^{\mathrm{mg}}$, are $|r_1| \ll R_1$ and $|r_2| \ll R_2$.

We define the functions $w_X^{(1)}$, $w_X^{(2)}$ and $w_X^{(3)}$ from $W(\mathbb{Z})^{\mathrm{mg}} \to \mathbb{R}$ by

$$w_X^{(1)}(A,B) = \sum_{\substack{r_1r_2\neq 0\\ r_i\ll R_i}} N(g(r_1,r_2)),\ \ w_X^{(2)}(A,B) = \sum_{r_2\ll R_2} N(g(0,r_2)),\ \ w_X^{(3)}(A,B) = \sum_{r_1\ll R_1} N(g(r_1,0))$$

where $g(x, y)$ is the cubic resolvent form corresponding to (A, B) and the sum is over integers r_1 and r_2, not both zero. By Proposition 4.4, we have

$$w_X(A,B) \ll_\epsilon X^\epsilon \left(w_X^{(1)}(A,B) + w_X^{(2)}(A,B) + w_X^{(3)}(A,B)\right). \tag{45}$$

Bounding the contribution of $w_X^{(1)}(A, B)$

Let X, Y, and Z be fixed, and take $\lambda \asymp Y^{1/12}$, $t \asymp Z^{1/2}$, and $s_1, s_2 \gg 1$. Set $s = (s_1, s_2)$. For

$$\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}} w_X^{(1)}(A,B) \tag{46}$$

to be nonzero, we must have $R_2 \gg 1$, which in turn implies $X^{1/6} \gg Y^{1/4}Z^{1/2}$. Thus, only pairs (A, B) with $\Delta(A, B) \ll X^{2/3}$ are being counted. With such a strong bound on the discriminant, we will not need any savings from the fact that only pairs (A, B) whose discriminants are strongly divisible by some m_2^2 (with $m_2 > M_2$) are being counted.

Fix an element $(A, B) \in (t, s)\lambda\mathcal{B}$, and let $g(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$ denote its cubic resolvent form. The coefficients of $g(x, y)$ and the values of $g(r_1, r_2)$, for $|r_1| \ll R_1$ and $|r_2| \ll R_2$, satisfy the following bound:

$$|a| \ll \frac{Y^{1/4}}{Z^{3/2}}; \quad |b| \ll \frac{Y^{1/4}}{Z^{1/2}}; \quad |c| \ll Y^{1/4}Z^{1/2}; \quad |d| \ll Y^{1/4}Z^{3/2}; \quad |g(r_1,r_2)| \ll \frac{X^{1/2}}{Y^{1/2}}. \tag{47}$$

We denote the coefficients of elements $(A,B) \in W(\mathbb{R})$ by a_{ij}, b_{ij}. The action of torus elements $(t,s)\lambda$ on $W(\mathbb{R})$ multiply each coefficient c_{ij} by a factor which we denote by $w(c_{ij})$. We have

$$w(a_{11}) = \frac{\lambda}{ts_1^4 s_2^2}; \quad w(a_{12}) = \frac{\lambda}{ts_1 s_2^2}; \quad w(a_{13}) = \frac{\lambda s_2}{ts_1}; \quad w(a_{22}) = \frac{\lambda s_1^2}{ts_2^2}; \quad w(b_{11}) = \frac{\lambda t}{s_1^4 s_2^2}. \tag{48}$$

If $(A,B) \in W(\mathbb{Z})$ satisfies either $a_{11} = b_{11} = 0$, or $a_{11} = a_{12} = a_{13} = 0$, or $a_{11} = a_{12} = a_{22} = 0$, then (A,B) is not generic (in the first case, A and B have a common rational zero, implying that the quartic ring corresponding to them is not an integral domain; in the second case, we have $\det(A) = 0$ implying that the cubic resolvent ring is not an integral domain). Thus, for $(t,s_1,s_2)\lambda\mathcal{B} \cap W(\mathbb{Z})^{\text{gen}}$ to be nonempty, we must have $w(b_{11}) \gg 1$, $w(a_{13}) \gg 1$ and $w(a_{22}) \gg 1$ (which in turn implies that $w(c_{ij}) \gg 1$ for all c_{ij} other than a_{11} and a_{12}). These weight conditions will be used to constrain the ranges of s_1 and s_2 in terms of t and λ.

We obtain bounds on (46) using two methods. In the first method, we use the trivial bound $N(k) \ll_\epsilon |k|^{1/2+\epsilon}$. This yields the estimate

$$w_X^{(1)}(A,B) \ll_\epsilon R_1 R_2 (X/Y)^{1/4+\epsilon} \ll_\epsilon \frac{X^{7/12+\epsilon}}{Y^{3/4}},$$

where we are using the bound on $|g(r_1,r_2)|$ from (47). Therefore, using Proposition 4.6 to bound the number of generic elements in $(t,s)\lambda\mathcal{B} \cap W(\mathbb{Z})$, we obtain

$$\begin{aligned} \sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\text{mg}}} w_X^{(1)}(A,B) &\ll_\epsilon X^{7/12+\epsilon} Y^{-3/4} (\lambda^{12} + \lambda^{11} t s_1^4 s_2^2 + \lambda^{10} t^2 s_1^5 s_2^4) \\ &\ll X^{7/12+\epsilon} \left(Y^{1/4} + t s_1^4 s_2^2 Y^{1/6} + t^2 s_1^5 s_2^4 Y^{1/12} \right). \end{aligned}$$

Noting that we have $Z \ll Y^{1/6}$, and integrating this over t, s, and λ yields

$$\begin{aligned} \int_{\substack{\lambda \asymp Y^{1/12} \\ t \asymp Z^{1/2} \\ s_1, s_2 \gg 1}} \Big(\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\text{mg}}} w_X^{(1)}(A,B) \Big) \frac{d^\times s d^\times t d^\times \lambda}{s_1^6 s_2^6 t^2} &\ll_\epsilon X^{7/12+\epsilon} \Big(\frac{Y^{1/4}}{Z^2} + \frac{Y^{1/6}}{Z} + Y^{1/12} \Big) \\ &\ll X^{7/12+\epsilon} Y^{1/4}. \end{aligned} \tag{49}$$

Since $Y \ll X^{2/3}$, we are already at the border of what is needed.

Our second method is based on the following idea. As A, B, r_1, and r_2 vary, the value of $g(r_1,r_2) = 4\det(Ar_1 + Br_2)$ should equidistribute. That would mean that our estimate $N(k) \ll_\epsilon |k|^{1/2+\epsilon}$ is inefficient most of the time. (For example, when k is squarefree, we have $N(k) = 1$!) To realize this idea, we fiber first over r_1 and r_2, then over possible values of $k = 4\det(Ar_1 + Br_2)$ (with each k weighted by $N(k)$). For each triple (r_1,r_2,k), we count the number of possible matrices $Ar_1 + Br_2$ with determinant $k/4$ (using, crucially, [41, Theorem 4.8]). Finally, for fixed r_1, r_2, and $Ar_1 + Br_2$, we estimate the number of (A,B): Namely, we bound by considering

$$\begin{aligned} \sum_{(A,B)} w_X^{(1)}(A,B) &= \sum_{(A,B)} \sum_{(r_1,r_2)} N(4\det(Ar_1 + Br_2)) \\ &= \sum_{(r_1,r_2)} \sum_k N(k) \sum_{\substack{C \in S(\mathbb{Z}) \\ 4|\det(C)| = k}} \#\{(A,B) : Ar_1 + Br_2 = C\}. \end{aligned}$$

There are $\ll R_1 R_2$ choices for the pair (r_1,r_2), and from the last bound in (47), k ranges $0 \neq |k| \ll X^{1/2}/Y^{1/2}$. We estimate the number of $C = Ar_1 + Br_2$ with $\det(C) = k/4$ as follows: Let

S denote the space of 3×3 symmetric matrices. Let $\mathcal{B}' \subset S(\mathbb{R})$ be a bounded set of 3×3-symmetric matrices. Let $k \neq 0$ be any fixed integer. For $T > 0$, we write

$$N_k(s_1, s_2; T) \;:=\; \{C \in (s_1, s_2)T\mathcal{B}' \cap S(\mathbb{Z}) : 4\det(C) = k\};$$

Then we have the following result proved in [41, Theorem 4.8].

Proposition 7.5 *For a real number $T > 1$ and integer $k \neq 0$, we have*

$$N_k(s_1, s_2; T) \ll_\epsilon (s_1^3 + s_2^3)T^{3+\epsilon} + s_1^4 s_2^5 T^{2+\epsilon},$$

independent of k.

Note that for $(A, B) \in (t, 1, 1)\lambda\mathcal{B}$, and $r_1 \ll R_1$ and $r_2 \ll R_2$, the coefficients of $r_1A + r_2B$ are $\ll X^{1/6}/Y^{1/6}$. Therefore, the number of integral symmetric matrices $C = r_1A + r_2B$, with $(A, B) \in (t, s)\lambda\mathcal{B}$ satisfying $4\det(C) = k$, is

$$\ll_\epsilon (s_1^3 + s_2^3)\frac{X^{1/2+\epsilon}}{Y^{1/2}} + s_1^4 s_2^5 \frac{X^{1/3+\epsilon}}{Y^{1/3}}.$$

The number of choices for the coefficients b_{11}, b_{12}, a_{13}, a_{22}, a_{23}, and a_{33} of $(A, B) \in (t, s)\lambda\mathcal{B} \cap W(\mathbb{Z})^{\mathrm{mg}}$ is $\ll Y^{1/2}/Z$, since the ranges of these coefficients are all $\gg 1$. The condition $r_1r_2 \neq 0$ implies that the choice of these coefficients, along with $Ar_1 + Br_2$ determines (A, B). Putting this together and applying Lemma 7.4, we obtain

$$\begin{aligned}\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}} w_X^{(1)}(A, B) \;&\ll_\epsilon\; R_1R_2\frac{X^{1/2}}{Y^{1/2}}\Big((s_1^3 + s_2^3)\frac{X^{1/2+\epsilon}}{Y^{1/2}} + s_1^4 s_2^5\frac{X^{1/3+\epsilon}}{Y^{1/3}}\Big)\frac{Y^{1/2}}{Z}\\ &=\; (s_1^3 + s_2^3)\frac{X^{4/3+\epsilon}}{YZ} + s_1^4 s_2^5\frac{X^{7/6+\epsilon}}{Y^{5/6}Z}.\end{aligned}$$

Integrating this over t, s, and λ yields

$$\int_{\substack{\lambda\asymp Y^{1/12}\\ t\asymp Z^{1/2}\\ s_1,s_2\gg 1}}\Big(\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}} w_X^{(1)}(A, B)\Big)\frac{d^\times s d^\times t d^\times\lambda}{s_1^6 s_2^6 t^2} \ll_\epsilon \frac{X^{4/3+\epsilon}}{YZ^2} + \frac{X^{7/6+\epsilon}}{Y^{5/6}Z^2} \ll \frac{X^{4/3+\epsilon}}{Y}. \tag{50}$$

We combine (49) and (50) to obtain

$$\begin{aligned}\int_{\substack{\lambda\asymp Y^{1/12}\\ t\asymp Z^{1/2}\\ s_1,s_2\gg 1}}\Big(\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}} w_X^{(1)}(A, B)\Big)\frac{d^\times s d^\times t d^\times\lambda}{s_1^6 s_2^6 t^2} \;&\ll_\epsilon\; X^\epsilon \min\{X^{7/12}Y^{1/4}, \frac{X^{4/3}}{Y}\}\\ &\leq\; X^{11/15+\epsilon},\end{aligned} \tag{51}$$

which is sufficiently small.

Bounding the contribution of $w_X^{(2)}(A, B)$

As before, let X, Y, and Z be fixed, and take $\lambda \asymp Y^{1/12}$, $t \asymp Z^{1/2}$, and $s_1, s_2 \gg 1$. Set $s = (s_1, s_2)$, and consider $(A, B) \in (t, s)\lambda\mathcal{B} \cap W(\mathbb{Z})^{\mathrm{mg}}_{M_2}$. We have

$$w_X^{(2)}(A, B) = \sum_{|r|\ll R_2} N(4\det(B)r^3). \tag{52}$$

Fiber over r (going up to R_2 in size), and $d = \det(B)$ (going up to $Y^{1/4}Z^{3/2}$ in size). For each fixed d, the number of choices for B is

$$\ll_\epsilon (s_1^3 + s_2^3)Y^{1/4+\epsilon}Z^{3/2} + s_1^4 s_2^5 Y^{1/6+\epsilon}Z$$

from Proposition 7.5. Meanwhile, the number of choices for A is bounded depending on the ranges of s_1 and s_2. If $w(a_{11}) \gg 1$, then there are $\ll Y^{1/2}/Z^3$ choices for A; if $w(a_{11}) < 1$ but $w(a_{12}) \gg 1$, then there are $\ll Y^{1/2}/Z^3 \cdot w(b_{11})/w(a_{11})$ choices for A; if $w(a_{11}) < 1$ and $w(a_{12}) < 1$, then there are $\ll Y^{1/2}/Z^3 \cdot w(b_{11})w(b_{12})/(w(a_{11})w(a_{12}))$ choices for A. Here, we are allowed to multiply by $w(b_{11})$ and $w(b_{12})$ since they are both $\gg 1$. This implies that the number of choices for A is $\ll Y^{1/2}/Z$. Therefore, applying Lemma 7.4, we have

$$\begin{aligned}\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}} \sum_{|r|\ll R_2} N(4\det(B)r^3) &\ll_\epsilon X^\epsilon R_2^2 Y^{3/4}Z^{1/2}\big((s_1^3+s_2^3)Y^{1/4}Z^{3/2} + s_1^4 s_2^5 Y^{1/6}Z\big)\\ &= X^{1/3+\epsilon}\left((s_1^3+s_2^3)Y^{1/2}Z + s_1^4 s_2^5 Y^{5/12}Z^{1/2}\right).\end{aligned}$$

Integrating this over s, t, and λ, and using the bound $Y \ll X^{2/3}$ implies that we have

$$\int_{\substack{\lambda\asymp Y^{1/12}\\ t\asymp Z^{1/2}\\ s_1,s_2\gg 1}}\Big(\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}} w_X^{(2)}(A,B)\Big)\frac{d^\times s d^\times t d^\times \lambda}{s_1^6 s_2^6 t^2} \ll_\epsilon X^{2/3+\epsilon}, \tag{53}$$

which is sufficiently small.

7.4 Bounding the contribution of $w_X^{(3)}$ using a uniform Ekedahl sieve

Finally, we obtain upper bounds on

$$\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}_{M_2}} w_X^{(3)}(A,B) = \sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}_{M_2}} \sum_{|r|\ll R_1} N(4\det(A)r^3). \tag{54}$$

This is the region from which we expect the biggest contribution. In the previous estimates (51) and (53) we had $R_2 \gg 1$ which implies $Y \ll X^{2/3}$. Thus, in particular we already have $M_1 \gg X^{1/6}$ and did not need any savings from M_2. In this final term (54), the situation is very different. In particular, we have to consider values of Y going all the way up to X, in which case all our savings will come from M_2.

Our strategy to bound (54) is similar to our previous methods. We fiber over r (going up to R_1) and $a = 4\det(A)$ (going up to $Y^{1/4}/Z^{3/2}$). As before, once $a \neq 0$ is fixed, the number of A that can arise with $4\det(A) = a$ is bounded in Proposition 7.5 by

$$(s_1^3+s_2^3)\frac{Y^{1/4+\epsilon}}{Z^{3/2}} + s_1^4 s_2^5 \frac{Y^{1/6+\epsilon}}{Z}.$$

Once A has been fixed (with $4\det(A) = a$), we bound the number of possible B's with $(A,B) \in (t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}_{M_2}$ using the Ekedhal sieve, as developed in [11]. The first thing we note is that if $p \mid A$ for some prime p, then (A,B) is not maximal for any B. Therefore, we can assume that the $\mathbb{F}_p$-rank of A is 1, 2, or 3 for every prime p. Suppose first that $p \mid a = \det(A)$, implying that the $\mathbb{F}_p$-rank of A is 1 or 2. In this case, the condition $p^2 \mid \Delta(A,B)$ for mod p reasons is at least a codimension 1 condition on B. Moreover, it is easy to see that the polynomial cutting out this

condition involves one of the coefficients b_{22}, b_{23}, and b_{33}. Next suppose that $p \nmid a$. In this case, the condition $p^2 \mid \Delta(A,B)$ for mod p reasons is a codimension 2-condition on B, and we may apply the Ekedahl sieve from [11, Proof of Theorem 3.3] without change. In other words, we proceed as follows. We have fixed A with $4\det(A) = a$. We are interested in counting the number of B's in a certain domain such that m^2 divides the discriminant of (A,B) for mod m reasons, where m is squarefree and $m > M_2$. We fiber over the possible values m' of the gcd of a and m. Each such m' imposes a mod-m' condition on one of b_{22}, b_{23}, and b_{33} (the smallest range of which is that of b_{22}). Moreover, there must exist an $m'' \geq M_2/m'$ which imposes a mod-m'' codimension-2 condition on B (involving b_{33}). Carrying this out, we obtain

$$\begin{aligned}
&\#\{B : (A,B) \in (t,s)\lambda\mathcal{B} \cap W(\mathbb{Z})^{\mathrm{mg}}_{M_2}\} \\
\ll_\epsilon\ & X^\epsilon \sum_{\substack{m' \mid a \\ |\mu(m')|=1}} w(b_{11})w(b_{12})w(b_{13})\max\Big(\frac{w(b_{22})}{m'},1\Big)w(b_{23})\max\Big(\frac{w(b_{33})}{\max(M_2/m',1)},1\Big) \\
\leq\ & X^\epsilon \sum_{\substack{m' \mid a \\ |\mu(m')|=1}} \Big(\frac{\lambda^6 t^6}{M_2} + \frac{\lambda^5 t^5 s_1^{-2} s_2^{-4}}{m'} + \frac{\lambda^5 t^5 s_1^{-2} s_2^2}{\max(M_2/m',1)} + \lambda^4 t^4 s_1^{-4} s_2^{-2}\Big) \\
\ll_\epsilon\ & X^\epsilon\Big(\frac{Y^{1/2}Z^3}{M_2} + s_1^{-2}s_2^2 Y^{5/12}Z^{5/2}\Big).
\end{aligned}$$

Combining the above discussion with the bound

$$\sum_{a \ll Y^{1/4}/Z^{3/2}} \sum_{r \ll R_1} N(ar^3) \ll_\epsilon X^\epsilon R_1^2 Y^{1/4}/Z^{3/2}$$

implied by Lemma 7.4, we obtain

$$\begin{aligned}
&\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}_{M_2}} \sum_{0\neq|r|\ll R_1} N(4\det(A)r^3) \\
\ll_\epsilon\ & X^\epsilon \frac{R_1^2 Y^{1/4}}{Z^{3/2}}\Big(s_1^3 s_2^3 \frac{Y^{1/4}}{Z^{3/2}} + s_1^4 s_2^5 \frac{Y^{1/6}}{Z}\Big)\Big(\frac{Y^{1/2}Z^3}{M_2} + s_1^{-2}s_2^2 Y^{5/12}Z^{5/2}\Big) \\
=\ & X^\epsilon\Big(\frac{s_1^3 s_2^3 X^{1/3}Y^{1/2}Z}{M_2} + s_1^4 s_2^5 \frac{X^{1/3}Y^{5/12}Z^{3/2}}{M_2} + s_1 s_2^5 X^{1/3}Y^{5/12}Z^{1/2} + s_1^2 s_2^7 X^{1/3}Y^{1/3}Z\Big).
\end{aligned}$$

Recall that we always have $Z \ll Y^{1/6}$ and $R_1 \gg 1$. The latter condition implies $Z^{1/2} \gg Y^{1/4}/X^{1/6}$. Multiplying the last summand in the above displayed equation by $1 \ll \sqrt{w(a_{22})} \asymp s_1 s_2^{-1} Y^{1/12} Z^{-1/2}$, and integrating over λ, t, and s yields

$$\begin{aligned}
&\int_{\substack{\lambda \asymp Y^{1/12} \\ t \asymp Z^{1/2} \\ s_1, s_2 \gg 1}} \Big(\sum_{(A,B)\in(t,s)\lambda\mathcal{B}\cap W(\mathbb{Z})^{\mathrm{mg}}} \sum_{|r|\ll R_1} N(4\det(A)r^3)\Big) \frac{d^\times s d^\times t d^\times \lambda}{s_1^6 s_2^6 t^2} \\
\ll_\epsilon\ & \frac{X^{1/3+\epsilon}Y^{1/2}}{M_2} + \frac{X^{1/3+\epsilon}Y^{5/12}}{Z^{1/2}} \qquad (55) \\
\ll_\epsilon\ & \frac{X^{1/3+\epsilon}Y^{1/2}}{M_2} + X^{1/2+\epsilon}Y^{1/6}.
\end{aligned}$$

We are now ready to prove the main result of this section.

Proof of Theorem 3: Combining (51), (53), and (55) yields

$$\int_{\substack{\lambda \asymp Y^{1/12} \\ t \asymp Z^{1/2} \\ s_1, s_2 \gg 1}} \Big(\sum_{(A,B) \in (t,s)\lambda \mathcal{B} \cap W(\mathbb{Z})^{\mathrm{mg}}_{M_2}} w_X(A,B) \Big) \frac{d^\times s d^\times t d^\times \lambda}{s_1^6 s_2^6 t^2} \ll_\epsilon \frac{X^{1/3+\epsilon} Y^{1/2}}{M_2} + X^{1/2+\epsilon} Y^{1/6} + X^{11/15+\epsilon}.$$

From Proposition 7.1, (43), and summing the above equation over Y and Z in the dyadic ranges $Y \leq X/M_1^2$ and $Z \ll Y^{1/6}$, we obtain

$$N(X; M_1, M_2) \ll_\epsilon \frac{X^{5/6+\epsilon}}{M_1 M_2} + \frac{X^{2/3+\epsilon}}{M_1^{1/3}} + X^{11/15+\epsilon}.$$

Finally, applying (38) completes the proof of Theorem 3. □

8 Proofs of the main results

In this section, we prove our main results.

8.1 Summing large and locally well approximated functions over $\mathrm{PGL}_2(\mathbb{Z})\backslash V(\mathbb{Z})$

In this section, we prove the following result.

Theorem 8.1 *Let ϕ be a large and locally well approximated function. Then, for $i \in \{0, 1, 2+, 2-\}$, we have*

$$N^{(i)}(\phi, X) = M^{(i)}_{5/6}(\phi) X^{5/6} + M^{(i)}_{3/4}(\phi) X^{3/4} + O_\epsilon(X^{3/4-1/3804+\epsilon}).$$

We assume that ϕ is fixed (and suppress the dependence on this constant in all the error terms in this section).

Proof: The proof of this result combines the methods and results of §4, §6, and §7, and is carried out in the following steps.

Step 1: Cutting off the tail. We begin by noting that for $f \in V(\mathbb{Z})\backslash\{\Delta \neq 0\}$, we have

$$\phi(f) = \prod_p \phi_p(f) = \sum_{n \geq 1} \phi(n; f), \quad \text{where} \quad \phi(n; f) := \prod_{p^k \| n} \phi_p^{(k)}(f).$$

The function $\phi(n; \cdot)$ is defined modulo n^2 and supported on the set of elements $f \in V(\mathbb{Z})$ with $n^2 \mid C\Delta(f)$ for a fixed positive integer C. Let $\delta > 0$ be a constant to be optimized later. Applying the uniformity estimate in Theorem 3, we obtain

$$\sum_{n \geq X^{1/12+\delta}} N^{(i)}(\phi(n; \cdot); X) \ll_\epsilon X^{3/4-\delta+\epsilon} + X^{3/4-1/60+\epsilon}, \tag{56}$$

since $\sum_{n \geq 1} |\phi(n, f)| \leq \sum_{n^2 | \Delta(f)} 1 \ll_\epsilon |\Delta(f)|^\epsilon$. As in §4, set $Y := X^{1/6}$. Combining (12) and (56), we have

$$\begin{aligned} N^{(i)}(\phi; X) &= \frac{1}{\sigma_i \mathrm{Vol}(G_0)} \sum_{n < X^{1/12+\delta}} \mathcal{I}(\phi(n; \cdot), Y) + O_\epsilon(X^{3/4-\delta+\epsilon} + X^{3/4-1/60+\epsilon}) \\ &= \frac{1}{\sigma_i \mathrm{Vol}(G_0)} \sum_{n < X^{1/12+\delta}} \big(\mathcal{I}^{(1)}(\phi(n; \cdot), Y; \kappa) + \mathcal{I}^{(2)}(\phi(n; \cdot), Y; \kappa) \big) \\ &\quad + O_\epsilon(X^{3/4-\delta+\epsilon} + X^{3/4-1/60+\epsilon}), \end{aligned} \tag{57}$$

for some $0 < \kappa < 1/4$ to be chosen later, where the quantities $\mathcal{I}(\cdot, Y)$ and $I^{(i)}(\cdot, Y; \kappa)$ are defined in (15) and (16), respectively. We refer to the sum over $\mathcal{I}^{(1)}(\phi(n;\cdot), Y; \kappa)$ as the main ball contribution to $N^{(i)}(\phi; X)$ and to the sum over $\mathcal{I}^{(2)}(\phi(n;\cdot), Y; \kappa)$ as the cuspidal contribution.

Step 2: Estimates for the main ball. Part 1: n has large squarefree part. Let $n < X^{1/12+\delta}$ be a positive integer. Denote the squarefree part of n by n_1. We use the bounds on the Fourier coefficients of $\phi(n;\cdot)$ obtained in Proposition 6.8 to estimate the quantity

$$N_n(S^{(i)}; Y, (u,t), \phi) := \sum_{f \in Y(u,t)S^{(i)} \cap V(\mathbb{Z})} \phi(n; f)$$

for $u \in [-1/2, 1/2]$ and $t \gg 1$. Note that we have $(u,t)S^{(i)} = (t) \cdot ((u/t^2, 1) \cdot S^{(i)})$, and that u/t^2 is absolutely bounded.

Let $\eta > 0$ be a sufficiently small real number, to be optimized later. Define the set $S^\sharp \subset S := (u/t^2, 1)S^{(i)}$ to be the (compact) set of elements $v \in \mathbb{R}^n$ such that $|v - w| \geq \eta$ for all elements w in the boundary of S. Define the set $S^\flat \supset S$ to be the (open) set of elements $v \in \mathbb{R}^n$ such that $|v - w| < \eta$ for some w in $\overline{S}$, the closure of S. Then there exist C^∞ functions $\Psi^\sharp := \Psi^\sharp_{S,\eta} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ and $\Psi^\flat := \Psi^\flat_{S,\eta} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ satisfying the following properties.

(a) The function $\Psi^\sharp$ is 1 in a neighbourhood of $S^\sharp$ and its support is contained in S.

(b) The function $\Psi^\flat$ is 1 in a neighbourhood of $\overline{S}$ and its support is contained in $S^\flat$.

(c) The partial derivatives of $\Psi^\sharp$ and $\Psi^\flat$ satisfy the following bounds:

$$|\partial^\alpha(\Psi^\sharp)|, |\partial^\alpha(\Psi^\flat)| \ll_{S,\alpha} \eta^{-|\alpha|}.$$

(d) The Fourier transforms of $\Psi^\sharp$ and $\Psi^\flat$ satisfy the following bounds:

$$|\widehat{\Psi^\sharp}(w)|, |\widehat{\Psi^\flat}(w)| \ll_M \min\big(1, (\eta|w|)^{-M}\big),$$

for $M > 0$.

Properties (a), (b), and (c) are consequences of [32, Theorem 1.4.1, Equation (1.4.2)]. Property (d) is a standard consequence of Property (c). Furthermore, since u/t^2 is absolutely bounded, the above error terms are independent of u. Define the auxilliary counting functions

$$\begin{aligned} N^\sharp_n(S; Y, (t), \phi) &:= \sum_{f \in V(\mathbb{Z})} \Psi^\sharp\Big(\frac{(t)^{-1} f}{Y}\Big)\phi(n; f), \\ N^{\mathrm{err}}_n(S; Y, (t), \phi) &:= \sum_{f \in V(\mathbb{Z})} (\Psi^\flat - \Psi^\sharp)\Big(\frac{(t)^{-1} f}{Y}\Big)|\phi(n; f)|, \end{aligned}$$

and note that we have

$$N_n(S^{(i)}; Y, (u,t), \phi(n;)) = N^\sharp_n(S; Y, (t), \phi) + O(N^{\mathrm{err}}_n(S; Y, (t), \phi)). \tag{58}$$

We use twisted Poisson summation to write

$$N^\sharp_n(S; Y, (t), \phi) = Y^5 \sum_{w \in V^*(\mathbb{Z})} \widehat{\Psi^\sharp}\Big(\frac{(t) \cdot Yw}{n^2}\Big)\widehat{\phi(n;\cdot)}(w).$$

From Property (d) of the function $\Psi^\sharp$, it follows that, up to negligible error, we can restrict the above sum to $w = (a, b, c, d, e) \in V^*(\mathbb{Z})$ satisfying

$$|a| \ll_\epsilon \frac{t^4 n^2}{\eta Y^{1-\epsilon}}; \quad |b| \ll_\epsilon \frac{t^2 n^2}{\eta Y^{1-\epsilon}}; \quad |c| \ll_\epsilon \frac{n^2}{\eta Y^{1-\epsilon}}; \quad |d| \ll_\epsilon \frac{n^2}{\eta t^2 Y^{1-\epsilon}}; \quad |e| \ll_\epsilon \frac{n^2}{\eta t^4 Y^{1-\epsilon}}.$$

Indeed, for the rest of the w's, we obtain the necessary saving from the superpolynomial decay of $\widehat{\Psi^\sharp}$. Hence, upto a negligible error, we have

$$\begin{aligned} |N_n^\sharp(S; Y, (t), \phi) - \nu(\phi(n;))\widehat{\Psi^\sharp}(0)Y^5| &\ll_\epsilon \eta^{-5} t^6 X^{5/6+10\delta+\epsilon} \cdot \frac{\nu(\phi(n;))}{n_1}, \\ |N_n^{\mathrm{err}}(S; Y, (t), \phi) - \nu(|\phi(n;)|)(\widehat{\Psi^\flat}(0) - \widehat{\Psi^\sharp}(0))Y^5| &\ll_\epsilon \eta^{-5} t^6 X^{5/6+10\delta+\epsilon} \cdot \frac{\nu(|\phi(n;)|)}{n_1}, \end{aligned} \tag{59}$$

where we use Corollary 6.8 to estimate $\widehat{\phi(n;)}$ and $\widehat{|\phi|(n;)}$. Let χ_{n^2} denote the characteristic function of the set of elements f such that $n^2 \mid \Delta(f)$, and note that some bounded constant multiple of $\chi_{n^2/(c,n^2)}$ dominates $|\phi(n;)|$, for some positive integer c determined by ϕ. Combining (58) and (59), we obtain

$$\begin{aligned} N_n(S^{(i)}; Y, (t), \phi) &= \nu(\phi(n;))\mathrm{Vol}(S^{(i)})X^{5/6} \\ &\quad + O\big(\nu(\chi_{n^2})\eta X^{5/6}\big) + O_\epsilon\Big(t^6\eta^{-5}\frac{\nu(\chi_{n^2})}{n_1}X^{5/6+10\delta+\epsilon}\Big). \end{aligned}$$

Finally, integrating the left hand side of the above equation over u and t, we obtain the following estimate for $\mathcal{I}^{(1)}(\phi(n;), Y; \kappa)$:

$$\begin{aligned} \mathcal{I}^{(1)}(\phi(n;\cdot), Y; \kappa) &= \nu(\phi(n;\cdot))\mathrm{Vol}(S^{(i)})X^{5/6}\int_{t\geq \frac{\sqrt[4]{3}}{\sqrt{2}}}\int_{u\in N(t)}\psi_1\Big(\frac{t}{Y^\kappa}\Big)du\frac{d^\times t}{t^2} \\ &\quad + O\big(\nu(\chi_{n^2})\eta X^{5/6}\big) + O_\epsilon\Big(\eta^{-5}\frac{\nu(\chi_{n^2})}{n_1}X^{5/6+10\delta+2\kappa/3+\epsilon}\Big). \end{aligned} \tag{60}$$

Step 3: Estimates for the main ball. Part 2: n has small squarefree part. When n has small squarefree part, we simply use (18) to write

$$\begin{aligned} \mathcal{I}^{(1)}(\phi(n;\cdot), Y; \kappa) &= \nu(\phi(n;\cdot))\mathrm{Vol}(S^{(i)})X^{5/6}\int_{t\geq \frac{\sqrt[4]{3}}{\sqrt{2}}}\int_{u\in N(t)}\psi_1\Big(\frac{t}{Y^\kappa}\Big)du\frac{d^\times t}{t^2} \\ &\quad + \mathrm{supp}(\phi(n;))O\Big(\frac{X^{2/3+\kappa/3}}{n^8} + \frac{X^{1/2+2\kappa/3}}{n^6} + X^{2\kappa/3}\Big). \\ &= \nu(\phi(n;\cdot))\mathrm{Vol}(S^{(i)})X^{5/6}\int_{t\geq \frac{\sqrt[4]{3}}{\sqrt{2}}}\int_{u\in N(t)}\psi_1\Big(\frac{t}{Y^\kappa}\Big)du\frac{d^\times t}{t^2} \\ &\quad + O\Big(n^2\nu(\chi_{n^2})X^{2/3+\kappa/3} + n^4\nu(\chi_{n^2})X^{1/2+2\kappa/3} + n^{10}\nu(\chi_{n^2})X^{2\kappa/3}\Big). \end{aligned} \tag{61}$$

Step 4: Estimates for the main ball. Part 3: Summing over n. Write $n = n_1 n_2^2 m_3$, where n_1, n_2, and m_3 are pairwise coprime, n_1 and n_2 are squarefree, and m_3 is cubefull. Fix $\theta > 0$ to be

optimized later. We sum $\mathcal{I}^{(1)}(\phi(n;\cdot),Y;\kappa)$ over $1\le n<X^{1/12+\delta}$, using (60) when $n_1>X^{1/12-\theta}$ and using (61) when $n_1\le X^{1/12-\theta}$. This yields

$$\sum_{1\le n<X^{1/12+\delta}}\mathcal{I}^{(1)}(\phi(n;\cdot),Y;\kappa)=\nu(\phi)\mathrm{Vol}(S^{(i)})X^{5/6}\int_{t\ge\frac{4\sqrt{3}}{\sqrt{2}}}\int_{u\in N(t)}\psi_1\Big(\frac{t}{Y^\kappa}\Big)du\frac{d^\times t}{t^2}+O_\epsilon(\text{error}),\quad(62)$$

where the error term is given by

$$\begin{aligned}\text{error}=&\sum_{\substack{n<X^{1/12+\delta}\\ n_1>X^{1/12-\theta}}}\Big(\nu(\chi_{n^2})\eta X^{5/6}+\eta^{-5}\frac{\nu(\chi_{n^2})}{n_1}X^{5/6+10\delta+2\kappa/3+\epsilon}\Big)\\ +&\sum_{\substack{n<X^{1/12+\delta}\\ n_1\le X^{1/12-\theta}}}\Big(n^2\nu(\chi_{n^2})X^{2/3+2\kappa/3}+n^4\nu(\chi_{n^2})X^{1/2+2\kappa/3}+n^{10}\nu(\chi_{n^2})X^{2\kappa/3}\Big)\\ +&X^\epsilon\sum_{n\ge X^{1/12+\delta}}\nu(\chi_{n^2})X^{5/6}.\end{aligned}$$

From Proposition 6.1, we have the bound $\nu(\chi_{n^2})\ll n_1^{-2}n_2^{-4}m_3^{-3/2}\log n$. Using this, we bound the first summand of the first line in the error as being $\ll_\epsilon$

$$\eta X^{5/6+\epsilon}\sum_{n_1>X^{1/12-\theta}}n_1^{-2}\sum_{n_2^2\ll\frac{X^{1/12+\delta}}{n_1}}n_2^{-4}\sum_{m_3\ll\frac{X^{1/12+\delta}}{n_1n_2^2}}m_3^{-3/2}\ll\eta X^{3/4+\theta+\epsilon}.\quad(63)$$

Similarly, the second summand in the first line is $\ll_\epsilon$

$$\eta^{-5}X^{5/6+10\delta+2\kappa/3+\epsilon}\sum_{n_1>X^{1/12-\theta}}n_1^{-3}\sum_{n_2^2\ll\frac{X^{1/12+\delta}}{n_1}}n_2^{-4}\sum_{m_3\ll\frac{X^{1/12+\delta}}{n_1n_2^2}}m_3^{-3/2}\ll\eta^{-5}X^{2/3+2\theta+10\delta+2\kappa/3+\epsilon}.$$

To bound the three summands in the second line, we note that $\sum_{m<M}m^\alpha=O(M^{\alpha+1/3})$ for $\alpha>0$, where in the sum m runs through all cubefull numbers less than M. Thus, the three summands are respectively $\ll_\epsilon$

$$\begin{aligned}X^{2/3+\kappa/3+\epsilon}\sum_{n_1\le X^{1/12-\theta}}\sum_{n_2^2\ll\frac{X^{1/12+\delta}}{n_1}}\sum_{m_3\ll\frac{X^{1/12+\delta}}{n_1n_2^2}}m_3^{1/2}&\ll_\epsilon X^{3/4-\theta/6+5\delta/6+\kappa/3+\epsilon},\\ X^{1/2+2\kappa/3+\epsilon}\sum_{n_1\le X^{1/12-\theta}}n_1^2\sum_{n_2^2\ll\frac{X^{1/12+\delta}}{n_1}}n_2^4\sum_{m_3\ll\frac{X^{1/12+\delta}}{n_1n_2^2}}m_3^{5/2}&\ll_\epsilon X^{3/4-\theta/6+17\delta/6+2\kappa/3+\epsilon},\\ X^{2\kappa/3}\sum_{n_1\le X^{1/12-\theta}}n_1^8\sum_{n_2^2\ll\frac{X^{1/12+\delta}}{n_1}}n_2^{16}\sum_{m_3\ll\frac{X^{1/12+\delta}}{n_1n_2^2}}m_3^{17/2}&\ll_\epsilon X^{3/4-\theta/6+53\delta/6+2\kappa/3+\epsilon}.\end{aligned}\quad(64)$$

Finally, the last line in the error term is $\ll X^{3/4-\delta+\epsilon}$.

Optimizing, we pick $\eta=X^{-1/72+\theta/6+5\delta/3+\kappa/9}$ and $\theta=1/96+43\delta/8+5\kappa/12$ to obtain

$$\begin{aligned}\mathcal{I}^{(1)}(\phi,Y;\kappa)&=\nu(\phi)\mathrm{Vol}(S^{(i)})X^{5/6}\int_{t\ge\frac{4\sqrt{3}}{\sqrt{2}}}\int_{u\in N(t)}\psi_1\Big(\frac{t}{Y^\kappa}\Big)du\frac{d^\times t}{t^2}\\ &\quad+O_\epsilon\Big(X^{3/4-1/576+83\delta/24+5\kappa/72+\epsilon}+X^{3/4-\delta+\epsilon}\Big).\end{aligned}\quad(65)$$

Step 5: Estimates for the cusp. Part 1: n is large. We begin with the following result:

Lemma 8.2 *Let $1 \leq F_0, F_1, F_2, F_3, F_4$ be real numbers, and let $B_{\vec{F}}$ denote the set of elements $f(x,y) = f_0x^4 + f_1x^3y + f_2x^2y^2 + f_3xy^3 + f_4y^4 \in V(\mathbb{R})$ such that $-F_i \leq f_i \leq F_i$ for each $i \in \{0,1,2,3,4\}$. For a positive integer n, we have*

$$\#\{f \in V(\mathbb{Z})^{\mathrm{gen}} \cap B_{\vec{F}} : a(f)b(f) \neq 0,\, n^2 \mid \Delta(f)\} \ll_\epsilon n^\epsilon \nu(\chi_{n^2}) M(\vec{F}, n), \tag{66}$$

where $M(\vec{F}, n) := F_0F_1(\max(F_2, n^2)\max(F_3, n^2)\max(F_4, n^2))$

Proof: For a ring R, and fixed elements $a \in R$ and $b \in R$, let $V(R)_{a,b}$ denote the set of elements $f \in V(R)$ with $a(f) = a$ and $b(f) = b$. For $m \geq 1$ with prime factorization $m = p_1^{k_1} \cdots p_\ell^{k_\ell}$, let $\nu_{a,b}(\chi_{m^2})$ denote the density in $V(\widehat{\mathbb{Z}})_{a,b}$ of the set of elements $f \in V(\widehat{\mathbb{Z}})_{a,b}$ with $m^2 \mid \Delta(f)$. It is clear that we have

$$\nu_{a,b}(\chi_{m^2}) = \prod_{i=1}^{\ell} \nu_{a,b}(\chi_{p_i^{2k_i}}).$$

Let p be a prime. If λ and r are elements of $\mathbb{Z}_p^\times$, and $f \in V(\mathbb{Z}_p)$, then $\Delta(f(x,y)) = \Delta(\lambda f(x, ry))$. As a consequence, it follows that for any $k \geq 1$, we have $\nu_{a,b}(\chi_{p^k}) = \nu_{a',b'}(\chi_{p^k})$ as long as $v_p(a) = v_p(a')$ and $v_p(b) = v_p(b')$. Therefore, for nonzero elements a and b of $\mathbb{Z}_p$, we have the bound

$$\nu_{a,b}(\chi_{p^k}) \ll \nu(\chi_{p^k}) v_p(a) v_p(b).$$

Returning to the sets $B_{\vec{F}}$, we simply fiber over a and b with $ab \neq 0$, and note that the LHS of (66) is equal to

$$\begin{aligned}
& \sum_{1 \leq |a| \leq F_0} \sum_{1 \leq |b| \leq F_1} \#\{f \in V(\mathbb{Z})^{\mathrm{gen}}_{a,b} \cap B_{\vec{F}} : n^2 \mid \Delta(f)\} \\
\ll & \sum_{1 \leq |a| \leq F_0} \sum_{1 \leq |b| \leq F_1} \nu_{a,b}(\chi_{n^2}) \max(F_2, n^2)\max(F_3, n^2)\max(F_4, n^2) \\
\ll & \sum_{\substack{\mathrm{rad}(d_a) \mid n \\ \mathrm{rad}(d_b) \mid n}} \sum_{\substack{|a| \leq F_0,\, |b| \leq F_1 \\ d_a \mid a,\, d_b \mid b}} d_a d_b \nu(\chi_{n^2}) \max(F_2, n^2)\max(F_3, n^2)\max(F_4, n^2) \\
\ll & \sum_{\substack{d_a \leq F_0, d_b \leq F_1 \\ \mathrm{rad}(d_a) \mid n,\, \mathrm{rad}(d_b) \mid n}} \nu(\chi_{n^2}) M(\vec{F}, n) \\
\ll_\epsilon & \; n^\epsilon \nu(\chi_{n^2}) M(\vec{F}, n),
\end{aligned}$$

as necessary. □

Pick $\theta > 0$ to be optimized later (this θ is of course independent of the θ which was optimized in the previous step). Let n be an integer such that $X^{1/12-\theta} \leq n < X^{1/12+\delta}$. We use Proposition 8.2 (and Lemma 4.3 for the elements f with $b(f) = 0$) to write

$$\mathcal{I}^{(2)}(\phi(n;\cdot), Y; \kappa) \ll_\epsilon X^{2/3} + \int_{Y^\kappa \ll t \ll Y^{1/4}} n^\epsilon \nu(\chi_{n^2}) \frac{Y^2}{t^6} \max(Y, n^2)\max(t^2Y, n^2)\max(t^4Y, n^2) \frac{d^\times t}{t^2}. \tag{67}$$

Suppose first that $n < X^{1/12}$. Then we have $Y \geq n^2$, and so the integral in the equation above is $\ll X^{5/6-\kappa/3+\epsilon}\nu(\chi_{n^2})$. Next suppose that $n \geq X^{1/12}$. We will finally choose $\kappa \geq 6\delta$, and assume

this now. In that case, we have $t^2Y \gg n^2$. Thus the integral is $\ll X^{3/4-\kappa/3+\epsilon}n^2\nu(\chi_{n^2})$. Summing this over $X^{1/12-\theta} \le n < X^{1/12+\delta}$ therefore yields

$$\begin{aligned}\sum_{n=X^{1/12-\theta}}^{X^{1/12+\delta}} \mathcal{I}^{(2)}(\phi(n;\cdot),Y;\kappa) &\ll_\epsilon X^{2/3+\epsilon} + X^{5/6-\kappa/3+\epsilon}\sum_{n=X^{1/12-\theta}}^{X^{1/12}}\nu(\chi_{n^2}) \\ &\quad + X^{3/4-\kappa/3+\epsilon}\sum_{n=X^{1/12}}^{X^{1/12+\delta}} n^2\nu(\chi_{n^2}) \\ &\ll_\epsilon X^{2/3+\epsilon} + X^{3/4+\theta-\kappa/3+\epsilon} + X^{3/4+\delta-\kappa/3+\epsilon}.\end{aligned} \tag{68}$$

Above, we estimate the sums over n of $\nu(\chi_{n^2})$ and $n^2\nu(\chi_{n^2})$ just as in (63) and (64), respectively. Moreover, the contribution of $X^{2/3}$ in (67) comes from Lemma 4.3, which bounds the number of relevant forms $f(x,y)$ with $b(f)=0$. When summed over n, each such form is counted at most $O_\epsilon(X^\epsilon)$ times, with a weight bounded by 1. Thus the sum over of the error term from Lemma 4.3 can be bounded by $O_\epsilon(X^{2/3+\epsilon})$ as stated.

Step 6: Estimates for the cusp. Part 2: n is small. Let $n \le X^{1/12-\theta}$ be a fixed positive integer. We use (22) to write $\mathcal{I}^{(2)}(\phi(n;\cdot),Y;\kappa)$ as a sum of two main terms along with an error that is $\ll_\epsilon$

$$X^{2/3+\epsilon} + \frac{X^{2/3}}{n^6}\sum_{0\neq|a|\ll Y^{1-4\kappa}}\frac{\mathrm{supp}(\phi(n;\cdot),a)}{|a|} \ll_\epsilon X^{2/3+\epsilon} + \frac{X^{2/3}}{n^6}\sum_{0\neq|a|\ll Y^{1-4\kappa}}\frac{\mathrm{supp}(\chi_{n^2},a)}{|a|}. \tag{69}$$

Let $q=p^k$ be a prime power. We know that $\mathrm{supp}(\chi_{q^2},a)=\mathrm{supp}(\chi_{q^2},b)$ if the valuations of a and b at p are the same. It follows that if $p^\ell \parallel a$ for $\ell \le 2k$, then we have

$$\mathrm{supp}(\chi_{q^2},a) \ll \frac{\mathrm{supp}(\chi_{q^2})}{q^{2k-\ell}} = \frac{\mathrm{supp}(\chi_{q^2})}{q^{2k}}\cdot(a,q^2),$$

which, by the Chinese remainder theorem, implies that we have

$$\mathrm{supp}(\chi_{n^2},a) \ll_\epsilon \frac{\mathrm{supp}(\chi_{n^2})}{n^{2-\epsilon}}(a,n^2).$$

Hence, we have

$$\begin{aligned}\sum_{0\neq|a|\ll Z}\frac{\mathrm{supp}(\chi_{n^2},a)}{|a|} &\ll \sum_{d|n^2}\sum_{\substack{0\neq|a|\ll Z\\(a,n^2)=d}}\frac{\mathrm{supp}(\chi_{n^2},a)}{|a|} \\ &\ll_\epsilon \frac{\mathrm{supp}(\chi_{n^2})}{n^{2-\epsilon}}\sum_{d|n^2}\sum_{\substack{0\neq|a|\ll Z\\ d|a}}\frac{d}{|a|} \\ &\ll_\epsilon \frac{\mathrm{supp}(\chi_{n^2})}{n^{2-\epsilon}}Z^\epsilon.\end{aligned}$$

Therefore, the error term in (69), summed up over $n < X^{1/12-\theta}$, is $\ll_\epsilon$

$$X^{3/4-\theta+\epsilon} + X^{2/3+\epsilon}\sum_{n<X^{1/12-\theta}} m_3^{1/2} \ll X^{3/4-\theta+\epsilon}.$$

Optimizing, we pick $\theta = \kappa/6$. Combining this with (68) yields

$$\begin{aligned}\mathcal{I}^{(2)}(\phi, Y; \kappa) &= \nu(\phi)\mathrm{Vol}(S^{(i)})X^{5/6}\int_{t>0}\psi_2\Big(\frac{t}{Y^\kappa}\Big)t^{-2}d^\times t \\ &\quad +\frac{1}{4}(D^+(\phi,1/2)\mathcal{V}^+(S^{(i)}) + D^-(\phi,1/2)\mathcal{V}^-(S^{(i)})) \\ &\quad +O_\epsilon(X^{3/4-\kappa/6+\epsilon} + X^{3/4+\delta-\kappa/3+\epsilon}).\end{aligned} \tag{70}$$

Step 7: Putting it all together. Combining (57), (65), and (70), we obtain

$$\begin{aligned}N^{(i)}(\phi, X) &= \frac{1}{\sigma_i\mathrm{Vol}(G_0)}\Big(\nu(\phi)\mathrm{Vol}(\mathcal{F})\mathrm{Vol}(S^{(i)})X^{5/6} + \frac{1}{4}\sum_{\circ=\pm}D^\circ(\phi,1/2)\mathcal{V}^\circ(S^{(i)})X^{3/4}\Big) \\ &\quad +O_\epsilon(X^{3/4-\delta+\epsilon} + X^{3/4-1/576+83\delta/16+5\kappa/72+\epsilon} + X^{3/4-\kappa/6+\epsilon} + X^{3/4+\delta-\kappa/3+\epsilon}).\end{aligned}$$

Optimizing, we pick $\kappa = 6\delta$ and $\delta = 1/3804$ to bound the error term by $O_\epsilon(X^{3/4-1/3804+\epsilon})$, thereby obtaining the result. □

8.2 Computing the primary and secondary terms

We begin with the following Jacobian change of variables formula proved in [14, Proposition 3.11]. Let ω be a generator of the rank-1 module of top degree differentials of PGL_2 over $\mathbb{Z}$. Let dv be Euclidean measure on V. Then we have

Proposition 8.3 ([14]) *Let F be $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Q}_p$ for some prime p. Let R be an open subset of $F \times F$, and let $s : R \to V(F)$ be a continuous function such that the invariants of $s_{I,J} := s(I, J)$ are I and J. Then for any measurable function $\phi : V(F) \to \mathbb{R}$, we have*

$$\int_{v\in\mathrm{PGL}_2(F)\cdot s(R)}\phi(v)dv = \Big|\frac{1}{27}\Big|\int_R\int_{\mathrm{PGL}_2(F)}\phi(g\cdot s_{I,J})\omega(g)dIdJ,$$

where we regard $\mathrm{PGL}_2(F)\cdot s(R)$ as a multiset and $|\cdot|$ denotes the absolute value of elements in F.

Next, we evaluate the quantities $\mathcal{V}^\pm(S^{(i)})$ defined in §4 by

$$\mathcal{V}^\pm(S^{(i)}) := \int_{\pm s>0}\frac{\mathrm{Vol}(S^{(i)}|_s)}{|s|^{1/2}}ds,$$

where $S^{(i)} = S_1^{(i)} = G_0R_1^{(i)}$. To this end, we have the following lemma.

Lemma 8.4 *Let F be $\mathbb{R}$ or $\mathbb{Q}_p$ for some prime p. Let $f \in V(K)\backslash\{\Delta = 0\}$ be an F-soluble binary quartic form with invariants $I(f) = I$ and $J(f) = J$, and denote E^{IJ} by E. Then we have*

$$\int_{(x,z)\in C_f(F)}\frac{dx}{|z|} = \int_{(x,y)\in E(F)}\frac{dx}{|y|},$$

where $C_f : z^2 = f(x, 1)$ is the genus-1 curve corresponding to f.

Proof: By assumption, f is soluble over F, which implies $C_f(F)$ is nonempty and hence contains some point Q. Since C_f is a trivial principal homogeneous space for E over F, it follows that $E(F)$ has a simple transitive action on $C_f(F)$. The map $\phi_Q : E(F) \to C_f(F)$, sending P to $P(Q)$ is a bijection. For the purpose of proving the lemma, it is enough to show that the Jacobian of ϕ_Q (with respect to the measures dx/z and dx/y on $C_f(F)$ and $E(F)$, respectively), is equal to 1.

Let $\overline{\phi}_Q : E(\overline{F}) \to C_f(\overline{F})$ be the map sending P to $P(Q)$, where $\overline{F}$ is the algebraic closure of F. The Jacobians of ϕ_Q and $\overline{\phi}_Q$ are clearly equal. We now prove that the Jacobian of $\overline{\phi}_Q$ is 1, as follows: first, by replacing f by a $\mathrm{PGL}_2(\overline{F})$-translate, if necessary, we may assume that $f(x,y) = x^3y - I/3xy^3 - J/27y^4$. In this case, the curve $C_f(\overline{F})$ can be naturally identified with $E(\overline{F})$ by sending (x,z) to (x,y). Under this identification, we simply have $\overline{\phi}_Q(P) = P + Q$, which has Jacobian 1 as necessary. The lemma follows. □

Next, we prove the following result.

Proposition 8.5 *For $i \in \{0, 1, 2+, 2-\}$, we have*

$$\frac{\mathcal{V}^+(S^{(i)}) + \mathcal{V}^-(S^{(i)})}{\mathrm{Vol}(G_0)} = \frac{C^{\circ}_{3/4}}{27\pi}, \tag{71}$$

where we take $\circ$ to be $\Delta > 0$ when $i \in \{0, 2+, 2-\}$ and $\circ$ to be $\Delta < 0$ when $i = 1$.

Proof: We start with writing

$$\begin{aligned}
\frac{\mathcal{V}^+(S^{(i)}) + \mathcal{V}^-(S^{(i)})}{\mathrm{Vol}(G_0)} &= \frac{1}{\mathrm{Vol}(G_0)} \int_{G_0 \cdot R_1^{(i)}} \frac{df}{\sqrt{|a(f)|}} \\
&= \frac{1}{27} \int_{f_{IJ} \in R_1^{(i)}} \int_{\gamma \in G_0} \frac{1}{\sqrt{|a(\gamma \cdot f_{IJ})|}} d\gamma dI dJ \\
&= \frac{1}{27} \int_{f_{IJ} \in R_1^{(i)}} \int_{\theta \in K} \frac{1}{\sqrt{|a(\theta \cdot f_{IJ})|}} d\theta dI dJ,
\end{aligned} \tag{72}$$

where the first equality follows from the definition of $S^{(i)} = G_0 \cdot R_1^{(i)}$, the second equality from a direct application of Proposition 8.3, and the third from the fact that the left hand side of (71) is independent of the right K-invariant set G_0; this last fact is deduced from the fact that the leading constants of the two terms in the right hand side of (23) must be independent of G_0. Above, as in §4, the measure $d\theta$ is Haar-measure on $K = \mathrm{SO}_2(\mathbb{R})$ normalized to have volume 1. For $f \in V(\mathbb{R}) \backslash \{\Delta = 0\}$ with invariants I and J, we compute the innermost integral above to be

$$\begin{aligned}
\int_{\theta \in K} \frac{1}{\sqrt{|a(\theta \cdot f)|}} d\theta &= \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \frac{1}{\sqrt{|f(\cos(\theta), \sin(\theta))|}} d\theta \\
&= \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \frac{1}{\sin^2(\theta)\sqrt{|f(\mathrm{cotan}(\theta), 1)|}} d\theta \\
&= \frac{1}{\pi} \int_{x=-\infty}^{\infty} \frac{1}{\sqrt{|f(x,1)|}} dx,
\end{aligned}$$

using the change of variables $x = \mathrm{cotan}(\theta)$. Now the set $\{(x, \sqrt{|f(x,1)|}) : x \in \mathbb{R}\}$ parametrizes precisely half of the real points on the two hypperelliptic curves C_f and C_{-f} (the other half being

parametrized by the set $\{(x, -\sqrt{|f(x,1)|}) : x \in \mathbb{R}\}$). Thus, from Lemma 8.4 and the definition of $\widetilde{\Omega}(I,J)$ in (2), we have

$$\begin{aligned}\int_{\theta\in K}\frac{1}{\sqrt{|a(\theta\cdot f)|}}d\theta &= \frac{1}{\pi}\int_{\substack{(x,z)\in C_f\\ z>0}}\frac{dx}{z}+\frac{1}{\pi}\int_{\substack{(x,z)\in C_{(-f)}\\ z>0}}\frac{dx}{z}\\ &= \frac{1}{\pi}\int_{\substack{(x,y)\in E^{I,J}\\ y>0}}\frac{dx}{y}+\frac{1}{\pi}\int_{\substack{(x,y)\in E^{I,-J}\\ y>0}}\frac{dx}{y}\\ &= \frac{1}{\pi}\widetilde{\Omega}(I,J).\end{aligned}$$

The result now follows immediately from (72) and the definition of $C^{\circ}_{3/4}$ in (3). □

It is now possible to compute the values of both $\mathcal{V}^+(S^{(i)})$ and $\mathcal{V}^-(S^{(i)})$:

Corollary 8.6 *We have $\mathcal{V}^{\pm}(S^{(i)}) = c^{\pm,i}\mathrm{Vol}(G_0)C^{\circ}_{3/4}/(27\pi)$, where $c^{\pm,0} = c^{\pm,1} = 1/2$, $c^{\pm,2\pm} = 1$, $c^{\pm,2\mp} = 0$, and where we take $\circ$ to be $\Delta > 0$ when $i \in \{0, 2+, 2-\}$ and $\circ$ to be $\Delta < 0$ when $i = 1$.*

Proof: This result is an immediate consequence of Propoposition along with the following identities:

$$\mathcal{V}^{\pm}(S^{(2\mp)}) = 0; \quad \mathcal{V}^+(S^{(0)}) = \mathcal{V}^-(S^{(0)}); \quad \mathcal{V}^+(S^{(1)}) = \mathcal{V}^-(S^{(1)}).$$

The first identity is immediate since every element in $R_1^{(2+)}$ (resp. $R_1^{(2-)}$), and hence every element in $S^{(2+)} = G_0 \cdot R_1^{(2+)}$ (resp. $S^{(2-)} = G_0 \cdot R_1^{(2-)}$), is positive (resp. negative) definite. The second identity can be deduced as follows: for $i \in \{0, 1\}$, the set $\{-f : f \in R^{(i)}\}$ is also a fundamental set for the action of $\mathrm{PGL}_2(\mathbb{R})$ on $V(\mathbb{R})^{(i)}$. Hence, as in (72), we have

$$\mathcal{V}^+(S^{(i)}) + \mathcal{V}^-(S^{(i)}) = \int_{G_0\cdot R_1^{(i)}}\frac{df}{\sqrt{|a(f)|}} = \int_{G_0\cdot(-R_1^{(i)})}\frac{df}{\sqrt{|a(f)|}},$$

with the contribution from either integral to $\mathcal{V}^{\pm}(S^{(i)})$ coming from forms f with $\pm a(f) > 0$. Since $a(-f) = -a(f)$, the result follows. □

We may now compute the values of $M^{(i)}_{5/6}(\phi)$ and $M^{(i)}_{3/4}(\phi)$ for large and locally well approximated functions ϕ.

Proposition 8.7 *Let $i \in \{0, 1, 2+, 2-\}$, let $j \in \{0, 1\}$ and let ϕ be a large and locally well approximated function. Then we have*

$$\begin{aligned}M^{(i)}_{5/6}(\phi) &= \frac{2\nu(\phi)\zeta(2)}{27\sigma_i}C^{\circ}_{5/6};\\ M^{(j)}_{3/4}(\phi) &= \frac{D^+(\phi,1/2)+D^-(\phi,1/2)}{216\sigma_j\pi}C^{\circ}_{3/4};\\ M^{(2\pm)}_{3/4}(\phi) &= \frac{D^{\pm}(\phi,1/2)}{108\sigma_2\pi}C^{\Delta>0}_{3/4}.\end{aligned}$$

As before, we take $\circ$ to be $\Delta > 0$ when $i \in \{0, 2+, 2-\}$ or $j = 0$ and to be $\Delta < 0$ when $i = 1$ or $j = 1$.

Proof: The first equality above follows from Proposition 8.3, while the remaining equalities follow from Corollary 8.6. □

Theorem 7 follows from Theorem 8.1 and Proposition 8.7. Proposition 8.7 also completes the proof of Theorem 2 in §4. For Theorem 6, the only remaining pieces are the values of $\kappa_{5/6}(\sigma_p)$ and $\kappa_{5/6}(\sigma_p)$ for splitting types σ_p; these are computed in the appendix. Finally, we prove the main elliptic curve results.

Proof of Theorem 5: Let $\phi : \mathbb{Z}^2 \to \mathbb{R}$ be large and locally approximated. From Theorem 3.4 and Lemmas 4.2 and 4.5, it follows that we have

$$\begin{aligned}\sum_{E_{AB}\in\mathcal{E}(X)^+} (|\mathrm{Sel}_2(E_{AB})| - 1)\phi(A,B) &= N^{(0)}(\psi, 2^6X) + N^{(2+)}(\psi, 2^6X) + O_\epsilon(X^{2/3+\epsilon}),\\ \sum_{E_{AB}\in\mathcal{E}(X)^-} (|\mathrm{Sel}_2(E_{AB})| - 1)\phi(A,B) &= N^{(1)}(\psi, 2^6X) + O_\epsilon(X^{2/3+\epsilon}),\end{aligned}$$

where $\psi\colon V(\mathbb{Z}) \to \mathbb{R}$ is defined by $\psi(f) = \phi(A(f), B(f))\ell(f)/m(f)$. By Lemma 5.9, it follows that ψ is large and well approximated, and so we may apply Theorem 8.1 to write the left hand sides of the above equation as a sum of two main terms along up to a sufficiently small error. Finally, the fact that the main terms arising from Theorem 8.1 align with the main term claimed in Theorem 5 follows from [14, Theorem 3.1] by simply approximating each ϕ_p by a finite sum of characteristic functions. The result follows. □

Theorem 1 is an immediate consequence of Theorem 5.

A Computations of primary and secondary local densities

Let σ_p be a splitting type modulo p. We begin by computing the primary density $\kappa_{5/6}(\sigma_p)$.

Lemma A.1 *The values of $\kappa_{5/6}(\sigma_p)$ are as given in Table 1.*

Proof: To compute the density of integral forms with some fixed splitting type σ_p, note that this is the same as the density of elements in $V(\mathbb{F}_p)$ with splitting type σ_p, and that this latter density can be computed via a simple counting argument. For example, an element in $V(\mathbb{F}_p)$ with splitting type (1111) is determined by four distinct points in $\mathbb{P}^1_{\mathbb{F}_p}$, up to multiplication by an element in $\mathbb{F}_p^\times$. There are $\binom{p+1}{4}$ such sets of four points, and multiplying by $(p-1)/p^5$ yields the density. The cases of the other splitting types are similar. □

Next, let $\phi : V(\mathbb{Z}) \to \mathbb{R}$ be a large and locally well approximated function via the functions $\phi_p : V(\mathbb{Z}_p) \to \mathbb{R}$. We proved in Corollary 6.6 that $D^\pm(\phi, s)$ has an analytic continuation to the right of $\Re(s) = 1/3$. Assume that ϕ_p is invariant under multiplication by units in $\mathbb{Z}_p$, i.e., $\phi_p(f) = \phi_p(uf)$ for $u \in \mathbb{Z}_p^\times$. Then $D^\pm(\phi, s)$ has an Euler product expansion:

$$D^\pm(\phi, s) = \sum_{a>0} \frac{\nu_{\pm a}(\phi)}{a^s} = \sum_{a>0} \prod_{p^k \| a} \frac{\nu_{p^k}(\phi_p)}{p^{ks}} = \prod_p D_p(\phi_p, s),$$

$$\text{where} \qquad D_p(\phi_p, s) := \nu_1(\phi_p) + \frac{\nu_p(\phi_p)}{p^s} + \frac{\nu_{p^2}(\phi_p)}{p^{2s}} + \cdots = \sum_{k\geq 0} \frac{\nu_{p^k}(\phi_p)}{p^{ks}}.$$

We now compute these densities $\nu_{p^k}(\phi_p)$ for certain functions ϕ_p.

σ_p	$\kappa_{5/6}(\sigma_p)$
(1111)	$(p+1)(p-1)^2(p-2)/(24p^4)$
(112)	$(p+1)(p-1)^2/(4p^3)$
(13)	$(p+1)^2(p-1)^2/(3p^4)$
(22)	$(p-1)^2(p+1)(p-2)/(8p^4)$
(4)	$(p+1)(p-1)^2/(4p^3)$
(1^211)	$(p+1)(p-1)^2/(2p^4)$
(1^22)	$(p+1)(p-1)^2/(2p^4)$
(1^31)	$(p+1)(p-1)/p^4$
(1^21^2)	$(p+1)(p-1)/(2p^4)$
(2^2)	$(p-1)^2/(2p^4)$
(1^4)	$(p+1)(p-1)/p^5$

Table 1: Splitting type densities in $V(\mathbb{Z}_p)$

Proposition A.2 *For a splitting type σ, let $\chi_\sigma : V(\mathbb{Z}_p) \to \{0,1\}$ denote the characteristic function of the set of elements in $V(\mathbb{Z}_p)$ having splitting type σ. The values of $\nu_{p^k}(\chi_\sigma)$ are as given in Table 2.*

Proof: Since the set of elements in $V(\mathbb{Z}_p)$ with a given splitting type is defined modulo p, the densities associated to χ_σ can be computed by counting the relevant elements in $V(\mathbb{F}_p)$. So for example, the number of elements $f(x,y) \in V(\mathbb{F}_p)$ with leading coefficient 1 (resp. leading coefficient 0) and factoring into four distinct linear factors is equal to $p(p-1)(p-2)(p-3)/24$ (resp. $p(p-1)^2(p-2)/6$, leading to the listed values of $\nu_a(\chi_{(1111)})$. The computations for the other unramified splitting types are similar.

Next consider $\sigma = (1^211)$. An element $f(x,y) \in V(\mathbb{Z}_p)$ has splitting type σ if and only if the reduction of f modulo p has three distinct roots in $\mathbb{P}^1_{\mathbb{F}_p}$, one of which (say r) is a double root. Moreover it is easy to see that the resolvent of f is maximal if and only if $p^2 \nmid f(\tilde{r})$ for a lift $\tilde{r}$ of r. As before, the value of $\nu_1(\chi_{(1^211)})$ can be computed via counting roots in $\mathbb{F}_p$ to be $p(p-1)(p-2)/(2p^4)$, as listed in the table. Since each such element in $V(\mathbb{Z}_p)$ is maximal with probability $(p-1)/p$, the listed value of $\nu_1(\chi^{\max}_{(1^211)})$ is correct. To compute the values of $\nu_{p^k}(\chi_\sigma)$, we separate into cases depending on whether the root at infinity is the double root or a single root. In the former case, we must count elements in $V(\mathbb{F}_p)$ of the form $\alpha y^2(x-r_1y)(x-r_2y)$, where $\alpha \in \mathbb{F}_p^\times$ and $r_1, r_2 \in \mathbb{F}_p$ are distinct. There are clearly $p(p-1)^2/2$ such elements, and such a form lifts to a maximal quartic form in $V(\mathbb{Z}_p)$ if and only if $k = 1$. In the latter case, we must count elements in $V(\mathbb{F}_p)$ of the form $\alpha y(x-r_1)(x-r_2)^2$; there are $p(p-1)^2$ such elements, and they lift to maximal elements with probability $1-1/p$. Adding up the contributions from these cases gives the desired values. The computations in the other cases are similar, where we also note that a binary quartic form with splitting type (1^21^2), (2^2), or (1^4) is never strongly maximal. $\square$

ϕ	$\nu_1(\phi)$	$\nu_p(\phi)$	$\nu_{p^k}(\phi)$, $k \geq 2$
$\chi_{(1111)}$	$(p-1)(p-2)(p-3)/(24p^3)$	$(p-1)^2(p-2)/(6p^3)$	$(p-1)^2(p-2)/(6p^3)$
$\chi_{(112)}$	$(p-1)^2/(4p^2)$	$(p-1)^2/(2p^2)$	$(p-1)^2/(2p^2)$
$\chi_{(13)}$	$(p+1)(p-1)/(3p^2)$	$(p+1)(p-1)^2/(3p^3)$	$(p+1)(p-1)^2/(3p^3)$
$\chi_{(22)}$	$(p+1)(p-1)(p-2)/(8p^3)$	0	0
$\chi_{(4)}$	$(p+1)(p-1)/(4p^2)$	0	0
$\chi_{(1^2 11)}$	$(p-1)(p-2)/(2p^3)$	$3(p-1)^2/(2p^3)$	$3(p-1)^2/(2p^3)$
$\chi_{(1^2 2)}$	$(p-1)/(2p^2)$	$(p-1)^2/(2p^3)$	$(p-1)^2/(2p^3)$
$\chi_{(1^3 1)}$	$(p-1)/p^3$	$2(p-1)/p^3$	$2(p-1)/p^3$
$\chi^{\max}_{(1^2 11)}$	$(p-1)^2(p-2)/(2p^4)$	$(3p-2)(p-1)^2/(2p^4)$	$(p-1)^3/p^4$
$\chi^{\max}_{(1^2 2)}$	$(p-1)^3/(2p^4)$	$(p-1)^2/(2p^3)$	0
$\chi^{\max}_{(1^3 1)}$	$(p-1)^2/p^4$	$(2p-1)(p-1)/p^4$	$(p-1)^2/p^4$
$(1^2 1^2)$	$(p-1)/(2p^3)$	$(p-1)/p^3$	$(p-1)/p^3$
(2^2)	$(p-1)/(2p^3)$	0	0
(1^4)	$1/p^3$	$(p-1)/p^4$	$(p-1)/p^4$

Table 2: Splitting type densities in $V_a(\mathbb{Z}_p)$

References

[1] C. T. Anderson, A. Gafni, K. Hughes, R. J. Lemke Oliver, D. Lowry-Duda, F. Thorne, J. Wang, and R. Zhang. Improved bounds on number fields of small degree. *arXiv preprint 2204.01651*, 2022.

[2] A. M. Baily. On the density of discriminants of quartic fields. *J. Reine Angew. Math.*, 315:190–210, 1980.

[3] J. S. Balakrishnan, W. Ho, N. Kaplan, S. Spicer, W. Stein, and J. Weigandt. Databases of elliptic curves ordered by height and distributions of Selmer groups and ranks. *LMS J. Comput. Math.*, 19:351–370, 2016.

[4] B. Bektemirov, B. Mazur, W. Stein, and M. Watkins. Average ranks of elliptic curves: tension between data and conjecture. *Bull. Amer. Math. Soc. (N.S.)*, 44(2):233–254, 2007.

[5] K. Belabas, M. Bhargava, and C. Pomerance. Error estimates for the Davenport-Heilbronn theorems. *Duke Math. J.*, 153(1):173–210, 2010.

[6] J. Bergstrom, A. Diaconu, D. Peterson, and C. Westerland. Hyperelliptic curves, the scanning map, and moments of families of quadratic l-functions. *arXiv preprint 2302.07664*, 2023.

[7] M. Bhargava. Galois groups of random integer polynomials and van der Waerden's conjecture. *Ann. of Math. (2)*. to appear.

[8] M. Bhargava. Higher composition laws. III. The parametrization of quartic rings. *Ann. of Math. (2)*, 159(3):1329–1360, 2004.

[9] M. Bhargava. The density of discriminants of quartic rings and fields. *Ann. of Math. (2)*, 162(2):1031–1063, 2005.

[10] M. Bhargava. The density of discriminants of quintic rings and fields. *Ann. of Math. (2)*, 172(3):1559–1591, 2010.

[11] M. Bhargava. The geometric sieve and the density of squarefree values of invariant polynomials. *arXiv preprint arXiv:1402.0031*, 2014.

[12] M. Bhargava, J. Hanke, and A. Shankar. The mean number of 2-torsion elements in class groups of n-monogenized cubic fields. *arXiv preprint 2010.15744*, 2020.

[13] M. Bhargava and A. Shankar. The average size of the 5-Selmer group of elliptic curves is 6, and the average rank is less than 1, 2013. arXiv:1312.7859.

[14] M. Bhargava and A. Shankar. Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves. *Ann. of Math. (2)*, 181(1):191–242, 2015.

[15] M. Bhargava and A. Shankar. Ternary cubic forms having bounded invariants, and the existence of a positive proportion of elliptic curves having rank 0. *Ann. of Math. (2)*, 181(2):587–621, 2015.

[16] M. Bhargava, A. Shankar, and J. Tsimerman. On the Davenport-Heilbronn theorems and second order terms. *Invent. Math.*, 193(2):439–499, 2013.

[17] M. Bhargava, T. Taniguchi, and F. Thorne. Improved error estimates for the Davenport-Heilbronn theorems. *Math. Ann.*, 389(4):3471–3512, 2024.

[18] B. J. Birch and H. P. F. Swinnerton-Dyer. Notes on elliptic curves. I. *J. Reine Angew. Math.*, 212:7–25, 1963.

[19] K. Chang. Hurwitz spaces, Nichols algebras, and Igusa zeta functions, 2023. arXiv:2306.10446.

[20] P. J. Cho and H. H. Kim. Low lying zeros of Artin L-functions. *Math. Z.*, 279(3-4):669–688, 2015.

[21] J. E. Cremona. *Algorithms for modular elliptic curves*. Cambridge University Press, Cambridge, second edition, 1997.

[22] H. Davenport. On a principle of Lipschitz. *J. London Math. Soc.*, 26:179–183, 1951.

[23] H. Davenport. On the class-number of binary cubic forms. I. *J. London Math. Soc.*, 26:183–192, 1951.

[24] B. N. Delone and D. K. Faddeev. *The theory of irrationalities of the third degree*. Translations of Mathematical Monographs, Vol. 10. American Mathematical Society, Providence, R.I., 1964.

[25] A. Diaconu. On the third moment of $L(\frac{1}{2}, \chi_d)$ I: The rational function field case. *J. Number Theory*, 198:1–42, 2019.

[26] A. Diaconu and H. Twiss. Secondary terms in the asymptotics of moments of L-functions. *J. Number Theory*, 252:243–297, 2023.

[27] S. r. Galatius, A. Kupers, and O. Randal-Williams. E_2-cells and mapping class groups. *Publ. Math. Inst. Hautes Études Sci.*, 130:1–61, 2019.

[28] W. T. Gan, B. Gross, and G. Savin. Fourier coefficients of modular forms on G_2. *Duke Math. J.*, 115(1):105–169, 2002.

[29] D. Goldfeld. Conjectures on elliptic curves over quadratic fields. In *Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979)*, volume 751 of *Lecture Notes in Math.*, pages 108–118. Springer, Berlin, 1979.

[30] C. Gómez-Gonzáles and J. Wolfson. Problems in arithmetic topology. *Res. Math. Sci.*, 8(2):Paper No. 23, 14, 2021.

[31] H. Heilbronn. On the 2-classgroup of cubic fields. In *Studies in Pure Mathematics (Presented to Richard Rado)*, pages 117–119. Academic Press, London, 1971.

[32] L. Hörmander. *The analysis of linear partial differential operators. I.* Classics in Mathematics. Springer-Verlag, Berlin, 2003. Distribution theory and Fourier analysis, Reprint of the second (1990) edition [Springer, Berlin; MR1065993 (91m:35001a)].

[33] Y. Ishitsuka, T. Taniguchi, F. Thorne, and S. Y. Xiao. Exponential sums over singular binary quartic forms and applications, 2024. arXiv:2404.00541.

[34] N. M. Katz and P. Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy*, volume 45 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1999.

[35] F. Levi. Kubische zahlkorper und binare kubische formenklassen. *Akad. Wiss. Leipz., Mat. Nat. Kl.*, 66(1):26–37, 1914.

[36] J. Nakagawa. Orders of a quartic field. *Mem. Amer. Math. Soc.*, 122(583):viii+75, 1996.

[37] B. Poonen and E. Rains. Random maximal isotropic subspaces and Selmer groups. *J. Amer. Math. Soc.*, 25(1):245–269, 2012.

[38] D. P. Roberts. Density of cubic field discriminants. *Math. Comp.*, 70(236):1699–1705, 2001.

[39] M. Sato and T. Shintani. On zeta functions associated with prehomogeneous vector spaces. *Ann. of Math. (2)*, 100:131–170, 1974.

[40] A. Selberg and S. Chowla. On Epstein's zeta-function. *J. Reine Angew. Math.*, 227:86–110, 1967.

[41] A. Shankar, A. Siad, and A. Swaminathan. Counting integral points on symmetric varieties with applications to arithmetic statistics. *arXiv preprint 2304.01050*, 2023.

[42] A. Shankar, A. Södergren, and N. Templier. Sato-Tate equidistribution of certain families of Artin L-functions. *Forum Math. Sigma*, 7:Paper No. e23, 62, 2019.

[43] A. Shankar and J. Tsimerman. Heuristics for the asymptotics of the number of S_n-number fields. *J. Lond. Math. Soc. (2)*, 107(5):1613–1634, 2023.

[44] T. Shintani. On Dirichlet series whose coefficients are class numbers of integral binary cubic forms. *J. Math. Soc. Japan*, 24:132–188, 1972.

[45] T. Shintani. On zeta-functions associated with the vector space of quadratic forms. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 22:25–65, 1975.

[46] T. Taniguchi and F. Thorne. Secondary terms in counting functions for cubic fields. *Duke Math. J.*, 162(13):2451–2508, 2013.

[47] T. Taniguchi and F. Thorne. Orbital exponential sums for prehomogeneous vector spaces. *Amer. J. Math.*, 142(1):177–213, 2020.

[48] M. M. Wood. *Moduli spaces for rings and ideals*. ProQuest LLC, Ann Arbor, MI, 2009. Thesis (Ph.D.)–Princeton University.

[49] A. Yukie. On Shintani zeta functions for GL(2). *Trans. Amer. Math. Soc.*, 350(12):5067–5094, 1998.