Quadrics in arithmetic statistics.

Levent Alpöge

Εις μνήμην της Χρυσής Νότσκα.

Abstract.

We (re)introduce the circle method into arithmetic statistics. More specifically, we combine the circle method with Bhargava’s counting technique in order to give a general method that allows one to treat arithmetic statistical problems in which one is trying to count orbits on a subvariety of affine space defined by the vanishing of a quadratic invariant.

We explain this method by way of example by sketching how to compute the average size of 2-Selmer groups in each of the families $y^2 = x^3 + B^k$, giving full details in the case $k \equiv 1 \pmod{6}$.

In the course of the argument we introduce a smoothed form of Bhargava’s aforementioned method, as well as a trick which allows us to compute the above averages from knowledge of the averages over “unconstrained” families.

1 Introduction.

The field of arithmetic statistics lacks a general technique to count orbits on nontrivial invariant subvarieties. Essentially every application of the counting technique invented by Bhargava in his Ph.D. thesis has reduced such an orbit counting problem to one of counting lattice points in an expanding region and then dealt with the latter by invoking a standard lemma of Davenport, which we will think of as a jazzed up version of the usual count of integral points of bounded height in affine space.

The purpose of this paper is to give the first general technique going beyond this regime.

The technique arises from a substantial simplification we give of the argument in Samuel Ruth’s Ph.D. thesis [12], in which he bounds the average size of 2-Selmer groups in the family of Mordell curves $y^2 = x^3 + B$. The reason his

---

1 We know of exactly one work that faces point counting on a nontrivial subvariety “directly”, namely Samuel Ruth’s Ph.D. thesis [12]. However, there are certainly others, e.g. Yao Xiao’s [13], that change such counting problems into more tractable ones by using particular features of their situations. While quite clever, we believe they are not relevant here because our goal is to produce a general technique (awkward as it is to formulate general statements in this context).

2 We are of course abbreviating significantly.
argument needs simplification is that his treatment of what we will call the "tail" is quite hard-going and in any case particular to his problem. That said, his idea of combining the circle method with Bhargava’s counting method makes what we will call the "bulk" contribution trivial to deal with, and this is our starting point. We then replace his treatment of the "tail" with a trick. Finally we forego any calculation of local densities with another trick — by formally deducing that the 2-Selmer average in question must match the 2-Selmer average over the "unconstrained" family of elliptic curves \( y^2 = x^3 + Ax + B \).

A slightly less informal summary of the technique is as follows. We follow Ruth in using the circle method to count said points in the bulk (namely when not "polynomially high in the cusp") by using the smoothed delta symbol method, and then we provide an a priori upper bound for the point count in the tail (namely when polynomially high in the cusp) via a divisor bound. As usual the divisor bound produces an estimate which is subpolynomially worse than sharp — however the key point is that this loss is more than compensated for by the "overconvergence" of the volume integral (and the fact that we only apply said estimate when polynomially high in the cusp). As for how we avoid any calculation of constants, note that the circle method already outputs a product of local densities, which are informally calculated by "thickening" an equality to a congruence modulo a highly divisible integer \( N \) (the singular integral has the analogous property) and taking \( N \to \infty \). Thus our 2-Selmer average, aka the 2-Selmer average over the family \( y^2 = x^3 + Ax + B \) with the equality \( A = 0 \) imposed, is 3 because the 2-Selmer average over the family \( y^2 = x^3 + Ax + B \), with only the congruence \( A \equiv 0 \pmod{6} \) imposed, is 3 (independent of \( N \!\!), by work of Bhargava-Shankar.

Now let us give a precise statement of the example application.

1.1 Main theorem.

**Theorem 1.1.** Let \( k \in \mathbb{Z} \). Let \( \mathcal{B} \subseteq \mathbb{Z} - \{0\} \) be a set of positive density defined by congruence conditions. Then:

\[
\operatorname{Avg}_{B \in \mathcal{B} : |B| \leq X} \#\operatorname{Sel}_2(E_{0,B^k}/\mathbb{Q}) \leq \begin{cases} 
3 & k \not\equiv 0 \pmod{3} \\
2 & k \equiv 0 \pmod{6} \\
\infty & k \equiv 3 \pmod{6} 
\end{cases} + O_{\mathcal{B}}(o_{X \to \infty}(1)),
\]

with equality if e.g. \( \mathcal{B} \subseteq \mathbb{Z} - \{0\} \) is defined by finitely many congruence conditions.

Here a subset \( \mathcal{B} \subseteq \mathbb{Z} - \{0\} \) is defined by congruence conditions if and only if \( \mathcal{B} \cap \mathbb{Z}^+ \subseteq \mathbb{Z}^+ \) and \( (-\mathcal{B}) \cap \mathbb{Z}^+ \subseteq \mathbb{Z}^+ \) are, and a subset \( \mathcal{B} \subseteq \mathbb{Z}^+ \) is defined by congruence conditions if and only if for all \( p \) there is an open subset \( \mathcal{B}_p \subseteq \mathbb{Z}_p \) such that \( \mathcal{B} = \mathbb{Z}^+ \cap \prod_p \mathcal{B}_p \) as subsets of \( \prod_p \mathbb{Z}_p \). It is straightforward to adapt our arguments to prove equality in Theorem 1.1 for the more general class of

\footnote{We write “with equality if e.g.” to emphasize that this is not an if and only if statement.}
subsets $\emptyset \neq B \subseteq \mathbb{Z} - \{0\}$ which are "large" in a sense completely analogous to that of Bhargava-Shankar’s [9], but we will spare ourselves the notational effort.

There is of course nothing to do when $k \equiv 0 \pmod{3}$: the statement is evident if $k \equiv 0 \pmod{6}$ and follows from the fact that $\#|\text{Sel}_2(E_{0,k}/\mathbb{Q})| \geq 2^{\omega_{27,12}(B)}|\omega_{11,12}(B)}$ (consider Tamagawa factors for the evident 2-isogeny) when $k \equiv 3 \pmod{6}$, where $\omega_{a,b}(n) := \#\{p : p \equiv a \pmod{b}\}$.

In the case $k \equiv 1 \pmod{6}$ Theorem 1.1 is a slight generalization of the main theorem of Ruth’s Ph.D. thesis [12] (see Theorem 1.1.2 of his [12], which amounts to a proof of the upper bound when $B = \mathbb{Z} - \{0\}$ that was important for [2] (and thus [3]).

In our eyes it is also this case that serves as the clearest example of the method. Indeed, tracking the condition "$A = 0$" through the arguments of Bhargava-Shankar [9], we find that we must count orbits of binary quartic forms $F(X,Y) \in \mathbb{Z}[X,Y]$ with classical invariants $I(F) = 0$ and $|J(F)|$ bounded. We are counting points on a quadric because, writing $F(X,Y) = a \cdot X^4 + b \cdot X^3Y + c \cdot X^2Y^2 + d \cdot XY^3 + e \cdot Y^4$,

$$I(F) = 12ae - 3bd + c^2.$$

We will sketch the remaining cases. In the case $k \equiv 2 \pmod{6}$, as in Bhargava-Ho [8] the relevant parametrization of 2-Selmer elements is by orbits of pairs of binary cubic forms $(F_1, F_2)$ with $F_i \in \mathbb{Z}[X,Y]$, and now the invariant quadric is $0 = I(F_1, F_2) = 3a_1d_2 - b_1c_2 + c_1b_2 - 3d_1a_2$, where $F_i(X,Y) = a_i \cdot X^3 + \cdots + d_i \cdot Y^3$, but otherwise is — at a high level — similar to the case $k \equiv 1 \pmod{6}$, but is more intricate and dealt with in joint work with Bhargava and Shnidman [4]. What is perhaps different is our deduction of the cases $k \equiv 4, 5 \pmod{6}$ by rescaling our orbits suitably (the difficulty being that one must preserve integrality) and then appealing to results already proven in the corresponding case $k \equiv 3 \pmod{6}$ — an argument that ultimately works because $E_{0,B^k}$ and $E_{0,B^{k+3}}$ are quadratic twists, and so their 2-Selmer groups are both subsets of e.g. $H^1(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), E_{0,B^k}[2])$, defined by — potentially different, of course — local conditions. (Note that this trick already occurs in [4], and, though our treatment is explicit and does not manifestly involve this cohomological interpretation, we certainly learned said interpretation from that joint work.)

In the case $k \equiv 4 \pmod{6}$ we will also explain how to arrive to the same result through another explicit parametrization of $\text{Sel}_2(E_{0,B^k}/\mathbb{Q})$ by binary quartic forms. Note that the case $k \equiv 4 \pmod{6}$ is treated in yet another way — in fact by reducing to results already proven in the case $k \equiv 2 \pmod{6}$ instead — in the aforementioned joint work [4].

Note also that via the aforementioned trick one deduces the statement of Theorem 1.1 for the slightly more general families $\{y^2 = x^3 + dB^k : B \in B\}$ with $k \in \mathbb{Z}, 0 \neq d \in \mathbb{Z}$ and $B \subseteq \mathbb{Z} - \{0\}$ (quadratic twist by $d$ in the nontrivial case $k \neq 0 \pmod{3}$).
1.2 Main technique.

Having stated the main theorem, let us discuss the method of proof. In order to be specific, let us put ourselves in the case \( k \equiv 1 \pmod{6} \): we would like to count, up to \( \text{PGL}_2(\mathbb{Q}) \)-equivalence, the locally soluble binary quartic forms \( F \in \mathbb{Z}[X, Y] \) with \( I(F) = 0 \) and \( 0 \neq |J(F)| \leq X \). Write then \( V := \text{Sym}^4(2) \) for the space of binary quartic forms. It is now standard that, using a method introduced in Bhargava’s Ph.D. thesis [5] and first carried out in this context by Bhargava-Shankar [9] (let us ignore our smoothing for the sake of this discussion), the problem immediately reduces to the problem of obtaining an asymptotic of the following form:

\[
\int_{1 \ll \lambda \ll X^{1/2}} \frac{d^\infty \lambda}{|\lambda| \ll 1} \int_{1 \ll \xi \ll \lambda} t^{-2} d^\infty t \#\{ F \in \lambda \cdot n_u \cdot a_t \cdot G_0 \cdot L \cap V(\mathbb{Z})_{\text{irred}} : I(F) = 0 \} \sim \text{const.} \cdot X^{1/2}.
\]

Here \( n_u := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, a_t := \begin{pmatrix} t^{1/4} & \xi \\ 0 & \lambda \end{pmatrix} \), and \( d^\infty z := \frac{dz}{z} \), so that \( d^\infty \lambda du t^{-2} d^\infty t dk \) is the Haar measure on \( \text{GL}_2^+(\mathbb{R}) \) (the "+" denoting positive determinant) induced by Haar measures on \( \Lambda, N, A, K \) under the Iwasawa decomposition \( \text{GL}_2^+(\mathbb{R}) = \Lambda \cdot N \cdot A \cdot K \), where \( \Lambda \) are the scalars, \( N \) are the lower unipotents, \( A \) is the diagonal torus, and \( K := \text{SO}_2(\mathbb{R}) \) is the maximal compact.

For the sake of exposition, the reader should imagine \( G_0 \cdot L \subseteq V(\mathbb{R}) \) as a small ball of binary quartic forms of height \( \lambda \ll 1 \), and indeed should imagine (though also keep in mind that this is a slight oversimplification) that the set \( \lambda \cdot n_u \cdot a_t \cdot G_0 \cdot L \cap V(\mathbb{Z})_{\text{irred}} \) is simply the set \( S_{\text{approx.}} \), say, of \((a, b, c, d, e) \in \mathbb{Z}\) satisfying \( a, e \neq 0 \) and:

\[
|a| \ll \frac{\lambda^4}{t^4}, \\
|b| \ll \frac{\lambda^4}{t^2}, \\
|c| \ll \lambda^4, \\
|d| \ll t^2 \cdot \lambda^4, \\
|e| \ll t^4 \cdot \lambda^4.
\]

Examining the integral, one evidently reduces to needing to treat

\[
\#\{ F \in \lambda \cdot n_u \cdot a_t \cdot G_0 \cdot L \cap V(\mathbb{Z})_{\text{irred}} : I(F) = 0 \}.
\]

Were there no condition \( I(F) = 0 \), this would essentially be counting lattice points in a somewhat “round” set scaled by a parameter \( \lambda \), which is simple — in the standard treatment one invokes an easy lemma of Davenport, which is precisely tailored for this sort of counting problem. Indeed, examining the defining inequalities of \( S_{\text{approx.}} \), it is evidently quite simple to evaluate \( \#|S_{\text{approx.}}| \).

Thus the entire difficulty is in dealing with the condition \( I(F) = 0 \). However we are simply asking to count zeroes of a quadratic form in five variables that
lie in the scalings of a somewhat "round" set, and this is easy if the set is in fact quite "round" — in other words, if we are in the "bulk" and \(1 \ll t \ll \lambda^\varepsilon\), say, then the defining inequalities of \(S_{\text{approx}}\) are all, up to \(X^{O(\varepsilon)}\), the same, so that \(S_{\text{approx}}\) is essentially a scaled cube, and it is straightforward to count the number of solutions of a quadratic form in five variables lying in such a set, by the oldest forms of the circle method. For convenience we follow Ruth and use the smoothed delta symbol method, which trivializes the problem. In the end we obtain the asymptotic

\[
\# \{ F \in \lambda \cdot n_u \cdot \alpha t \cdot B \cdot L \cap V(\mathbb{Z})_{\text{irred}} : I(F) = 0 \} = \text{const} \cdot \lambda^{12} + O(\lambda^{12-\varepsilon})
\]

for a positive absolute constant \(\delta > 0\).

Thus the entire problem has reduced to treating the count when we are in the "tail", i.e. \(\lambda^\varepsilon \ll t \ll \lambda\). Here Ruth uses an argument that again involves the smoothed delta symbol method and is quite specific to the situation.

Our observation is that the "tail" case is also obviously trivial. Specifically, it is obvious that

\[
\# \{ F \in \lambda \cdot n_u \cdot \alpha t \cdot B \cdot L \cap V(\mathbb{Z})_{\text{irred}} : I(F) = 0 \} \ll \lambda^{12+o(1)}
\]

by the divisor bound: \(12ac - 3bd + c^2 = 0\) implies that \(a, e \mid 3bd - c^2\). Because we are dealing with elements in \(V(\mathbb{Z})_{\text{irred}}\) (and thus \(a, e \neq 0\)), it follows that \((b, c, d)\) determine \((a, e)\) up to \(\ll \lambda^{o(1)}\) choices, and the number of \((b, c, d)\) is \(4 \ll \lambda^{12}\).

Because of the \(o(1)\) in the exponent, this bound is not quite as sharp as what we obtained in the "bulk", but since we are in the "tail" we needn’t work so hard. The point is that the condition \(t \gg \lambda^\varepsilon\) combines with the \(t^{-2}d^\varepsilon t\) in the Haar measure (note that the exponent 2 is more than is needed for convergence, so to speak) to imply that this bound is enough to give a bound \(\ll \lambda^{12+\Omega(\varepsilon)}\) on the "tail" contribution after integrating over \(u\) and \(t\).

So in the end we obtain the desired asymptotic, which is equivalent to the theorem via Bhargava’s counting method. The argument in the case of pairs of binary cubic forms is the same, and indeed the above argument easily generalizes to other quadrics in at least four variables (said restriction arising in order to ensure that the circle method analysis in the "bulk" goes through easily).

2 Acknowledgments.

Chrysi Notskas’ influence on my life, mathematical and otherwise, was profound and this work is dedicated to her memory.

\footnote{While we have seemingly used the special form of \(I(F)\) in this argument, this divisor bound argument works in general — the point is that in a box a binary quadratic form represents a nonzero integer few times, because of a divisor bound in at most a quadratic extension (and the usual proof of Dirichlet’s unit theorem to deal with units).}

\footnote{It was through her that I was introduced to higher mathematics (and also to parallel programming and computer vision)! I found her phrase "bigger and better things", regarding how to respond to the unjustifiable, unforgettable.}
3 of the author’s Ph.D. thesis at Princeton University [11]. I would like to thank both my advisor Manjul Bhargava and Peter Sarnak for their patience and encouragement. I would also like to thank Joseph Glynnias, Ari Shnidman, Odysseus Skartsis, and Jacob Tsimerman for informative discussions. Finally I thank the National Science Foundation (via their grant DMS-2002109), Columbia University, and the Society of Fellows for their support during the pandemic.

It has been indicated to me that it is worth clarifying the following. As usual I will not be giving this paper to some journal. In contrast, a complete and detailed proof of Theorem 1.1 for \( k \equiv 2 \pmod{6} \) (cf. the sketch given in Section 5) will appear (and thus be journal published along with the analytic method itself), in the joint work [4].

3 Smoothing in Bhargava’s counting method.

Let us now introduce the aforementioned technical convenience that simplifies Bhargava’s counting method, as first introduced in Chapter 5 of his Ph.D. thesis [5], though the trick of averaging over fundamental domains was first introduced in the published version (see Section 2.2 of Bhargava’s [6]). The point is that, while an average over \( G_0 \) (using the notation of the above outline) improves the situation considerably, one is still integrating a “rough” function, namely \( BD_{G_0} \), and it is wiser to instead integrate a compactly supported smooth function. We note that this is completely natural from Bhargava’s original formulation — see equation (4) in Section 2.2 of Bhargava’s [6], and note that we are taking (in his notation) \( \Phi \) to be smooth and compactly supported, rather than the indicator function of a box.

In order to be specific, and for the reader’s convenience, let us work in the setup of the proof of Theorem 1.1 for \( k \equiv 1 \pmod{6} \) (it will be clear how to modify the construction for Theorem \( k \equiv 2 \pmod{6} \), and indeed in any application of Bhargava’s counting method). That is, \( V := \text{Sym}^4(2), G := \text{GL}_2, L := L^{(0)} \prod L^{(1)} \prod L^{(2+)} \prod L^{(2-)} \) with

\[
L^{(0)} := \left\{ X^3Y - \frac{1}{3} \cdot XY^3 + \frac{J}{27} \cdot Y^4 : J \in (-2, 2) \right\},
\]

\[
L^{(1)} := \left\{ X^3Y - \frac{J}{3} \cdot XY^3 \pm \frac{2}{27} \cdot Y^4 : I \in [-1, 1] \right\} \cup \left\{ X^3Y + \frac{1}{3} \cdot XY^3 + \frac{J}{27} \cdot Y^4 : J \in (-2, 2) \right\},
\]

\[
L^{(2\pm)} := \pm \left\{ \frac{1}{16} X^4 - \sqrt{\frac{2 - J}{27}} \cdot X^3Y + \frac{1}{2} \cdot X^2Y^2 + Y^4 : J \in (-2, 2) \right\},
\]

and

\[
\mathcal{F} := \left\{ \lambda \cdot n_u \cdot a_t \cdot k : \lambda \in \mathbb{R}^+, u \in \nu(t) \subseteq \left[ -\frac{1}{2}, \frac{1}{2} \right], t \geq \sqrt{\frac{3}{2}}, k \in \text{SO}_2(\mathbb{R}) \right\} \subseteq G(\mathbb{R}),
\]

Gauss’s classical fundamental domain for \( G(\mathbb{Z}) \curvearrowright G(\mathbb{R}) \). Note that \( L \) is a fundamental domain for \( G(\mathbb{R}) \curvearrowright V(\mathbb{R}) \setminus \neq 0 \). Observe that the points with \( I = 0 \)
and \( \Delta \neq 0 \) are all in the \( G(\mathbb{R}) \)-orbit of the two forms \( F_{\pm}(X, Y) := X^3Y \pm \frac{3}{2}Y^4 \), which lie in the interior of \( L^{(1)} \) (and thus lie in small compact subintervals thereof). Thus the following setup suffices for us.

Write, for each \( v \in V(\mathbb{R}) \Delta \neq 0 \), \( v_L \in L \) for the unique element of \( L \) mapping to the image of \( v \) under \( V(\mathbb{R}) \Delta \neq 0 \to V(\mathbb{R}) \Delta \neq 0 / G(\mathbb{R}) \simeq L \).

Let \( \mu_{\pm} \in C^\infty_c(\mathbb{R}) \) be such that \( \mu_- \leq \mathbb{1}_{[-1,1]} \leq \mu_+ \). Let \( \alpha \in C^\infty_c(G(\mathbb{R})) \) and \( \beta \in C^\infty_c(L) \) be compactly supported smooth functions such that: \( \alpha \) is \( \mathrm{SO}_2(\mathbb{R}) \)-invariant, \( \int_{G(\mathbb{R})} \alpha = 1 \), \( \beta(F_{\pm}) = 1 \), and \( \supp \beta, \beta^{-1}(\{1\}) \subseteq L^{(1)} \) are both unions of two small compact intervals respectively containing \( F_{\pm} \). Let

\[
\varphi(v) := \sum_{g \cdot v_L = v} \alpha(g) \cdot \beta(v_L).
\]

Note that this is a finite sum because stabilizers of elements of \( V(\mathbb{R}) \Delta \neq 0 \) are finite.

Via \( \mu_{\pm}, \alpha, \) and \( \beta \) we get a slightly more convenient way\(^{\ref{footnote:smooth}}\) to smooth out the various integrals in Bhargava’s counting technique — instead of integrating the normalized indicator function \( \frac{1}{|J_{G_0}|} \mathbb{1}_{G_0} \) of \( G_0 \) over \( g \in G(\mathbb{R}) \) (i.e. integrating over \( g \in G_0 \)) and observing that \( g \cdot \mathcal{F} \) is also (the closure of) a fundamental domain for \( G(\mathbb{Z}) \cap G(\mathbb{R}) \) so that all counts are independent of \( g \), we instead integrate \( \alpha(g) \) over \( g \in G \) and then make the same observation.

Specifically, we observe that, since

\[
\#\{(F, \mathcal{F} g \cdot L \cap V(\mathbb{Z})^{\text{nontriv}} : I(F) = 0, 0 \neq |J(F)| \leq X}\}
\]

is constant in \( g \) outside a measure-zero subset of \( G(\mathbb{R}) \) (since \( \mathcal{F} \cdot g \) and \( \mathcal{F} \) are both fundamental domains for \( G(\mathbb{Z}) \cap G(\mathbb{R}) \) — the first main observation of Section 2.2 of Bhargava’s \([6]\)), it follows that:

\[
\int_{g \in G_0} dg \#\{(F, \mathcal{F} g \cdot L \cap V(\mathbb{Z})^{\text{nontriv}} : I(F) = 0, 0 \neq |J(F)| \leq X}\} = \int_{G(\mathbb{R})} dg \alpha(g) \cdot \#\{(F, \mathcal{F} g \cdot L \cap V(\mathbb{Z})^{\text{nontriv}} : I(F) = 0, 0 \neq |J(F)| \leq X}\}.
\]

The left-hand side is precisely Ruth’s \( N(Y^{\text{irr}}, X) \).

We then manipulate this expression just as in Section 2.3 of Bhargava-Shankar’s \([9]\) (and implicitly in Bhargava’s \([6]\)).

\(^{\ref{footnote:smooth}}\)Specifically, this insertion of a smooth weight in Bhargava’s main trick in his counting technique saves us the effort required to remove smooth weights when applying the smoothed delta symbol method. We note here that we use \( \beta \) to ensure smoothness of \( \varphi \) — note that \( L^{(0)} \bigcup L^{(1)} \) is a rectangle missing two corners, and at the other two corners a similar definition of \( \varphi \) with \( \beta = 1 \) identically would fail to be smooth. One can get around this, of course, but this choice simplifies notation.
Evidently:

\[
\int_{G(\mathbb{R})} dg \alpha(g) \cdot \#\{F \in \mathcal{F} \cdot g \cdot L \cap V(\mathbb{Z})^{\text{nontriv}.} : I(F) = 0, 0 \neq |J(F)| \leq X\} \]

= \sum_{v \in V(\mathbb{Z})^{\text{nontriv}.} : I(v) = 0, 0 \neq |J(v)| \leq X} \int_{G(\mathbb{R})} dh \alpha(h) \cdot \#\{g \in \mathcal{F} : gh \cdot v_L = v\}.

But, using that \(\beta(v_L) = 1\) if \(I(v) = 0\) and \(\Delta(v) \neq 0\) (since this implies that \(v_L = F_+\) or \(F_-\)), for \(v \in V(\mathbb{Z})^{\text{nontriv}.}\) with \(I(v) = 0\) and \(J(v) \neq 0\) each integral can be evaluated as follows:

\[
\int_{G(\mathbb{R})} dh \alpha(h) \cdot \#\{g \in \mathcal{F} : gh \cdot v_L = v\} = \sum_{\gamma \in G(\mathbb{R}) : \gamma \cdot v_L = v} \int_{G(\mathbb{R})} dh \alpha(h) \cdot \beta(v_L) \cdot \#\{g \in \mathcal{F} : gh = \gamma\} = \sum_{\gamma \in G(\mathbb{R}) : \gamma \cdot v_L = v} \int_{\mathcal{F}^{-1} \gamma} dh \alpha(h) \cdot \beta(v_L) = \sum_{\gamma \in G(\mathbb{R}) : \gamma \cdot v_L = v} \int_{\mathcal{F}^{-1} \gamma} dh \alpha(h \cdot \gamma) \cdot \beta(v_L) = \int_{\mathcal{F}} dh \sum_{\gamma \in G(\mathbb{R}) : \gamma \cdot v_L = v} \alpha(h^{-1} \cdot \gamma) \cdot \beta(v_L) = \int_{\mathcal{F}} dh \sum_{g \in G(\mathbb{R}) : g \cdot v_L = h^{-1} \cdot v} \alpha(g) \cdot \beta(v_L) = \int_{\mathcal{F}} dh \phi(h^{-1} \cdot v),
\]

by definition.

Therefore we have found that:

\[
N(Y(\mathbb{Z})^{\text{nontriv}.}, X) := \#\{F \in \mathcal{F} \cdot L \cap V(\mathbb{Z})^{\text{nontriv}.} : I(F) = 0, 0 \neq |J(F)| \leq X\} = \sum_{v \in V(\mathbb{Z})^{\text{nontriv}.} : I(v) = 0, 0 \neq |J(v)| \leq X} \int_{\mathcal{F}} dh \phi(h^{-1} \cdot v) = \int_{\mathcal{F}} dh \sum_{v \in V(\mathbb{Z})^{\text{nontriv}.} : I(v) = 0, 0 \neq |J(v)| \leq X} \phi(h^{-1} \cdot v),
\]

and so e.g. \(N_{\mu_-}(Y(\mathbb{Z})^{\text{nontriv}.}, X) \leq N(Y(\mathbb{Z})^{\text{nontriv}.}, X) \leq N_{\mu_+}(Y(\mathbb{Z})^{\text{nontriv}.}, X)\), where:

\[
N_{\mu\pm}(Y(\mathbb{Z})^{\text{nontriv}.}, X) := \int_{\mathcal{F}} dh \sum_{v \in V(\mathbb{Z})^{\text{nontriv}.} : I(v) = 0, J(v) \neq 0} \mu_{\pm}\left(\frac{J(v)}{X}\right) \cdot \phi(h^{-1} \cdot v).
\]
4 Proof of Theorem 1.1 for $k \equiv 1 \pmod{6}$.

Now let us turn to the proof of Theorem 1.1 when $k \equiv 1 \pmod{6}$.

4.1 Reduction to point counting.

For the reader’s convenience we use Ruth’s notation in what follows. Let $V := \text{Sym}^4(2)$. Let $G := \text{GL}_2$. Let, for $F \in V$,

$I(F) := 12ae - 3bd + c^2$,

$J(F) := 72ace + 9bcd - 27ad^2 - 27b^2e - 2c^3$,

so that the discriminant

$\Delta(F) = 4I^3 - J^2$,

where we have written $F(X, Y) = a \cdot X^4 + b \cdot X^3Y + c \cdot X^2Y^2 + d \cdot XY^3 + e \cdot Y^4$.

Let $G \curvearrowright V$ via $(g \cdot F)(X, Y) := F((X, Y) \cdot g)$. We note that, for $g \in G$ and $F \in V$, we have that:

$I(g \cdot F) = (\det g)^4 \cdot I(F)$,

$J(g \cdot F) = (\det g)^6 \cdot J(F)$.

Let $L := L^{(0)} \bigsqcup L^{(1)} \bigsqcup L^{(2+)} \bigsqcup L^{(2-)}$ with

$L^{(0)} := \left\{ X^3Y - \frac{1}{3} \cdot XY^3 + \frac{J}{27} \cdot Y^4 : J \in (-2, 2) \right\}$,

$L^{(1)} := \left\{ X^3Y - \frac{I}{3} \cdot XY^3 + \frac{2}{27} \cdot Y^4 : I \in [-1, 1) \right\} \cup \left\{ X^3Y + \frac{1}{3} \cdot XY^3 + \frac{J}{27} \cdot Y^4 : J \in (-2, 2) \right\}$,

$L^{(2\pm)} := \pm \left\{ \frac{1}{16} \cdot X^4 - \sqrt{\frac{2 - J}{27}} \cdot X^3Y + \frac{1}{2} \cdot X^2Y^2 + Y^4 : J \in (-2, 2) \right\}$,

a fundamental domain for $G(\mathbb{R}) \curvearrowright V(\mathbb{R})_{\Delta \neq 0}$.

Let

$F := \left\{ \lambda \cdot n_u \cdot a_t \cdot k : \lambda \in \mathbb{R}^+, u \in \nu(t), t \geq \sqrt{\frac{3}{2}}, k \in \text{SO}_2(\mathbb{R}) \right\} \subseteq G(\mathbb{R})$,

where $\nu(t) \subseteq \left[ -\frac{1}{2}, \frac{1}{2} \right]$ is $\left[ -\frac{1}{2}, \frac{1}{2} \right]$ when $t \gg 1$ or else a union of two subintervals of $\left[ -\frac{1}{2}, \frac{1}{2} \right]$ when $\sqrt{\frac{3}{2}} \leq t \ll 1$ (just imagine the usual Gauss fundamental domain for $\text{SL}_2(\mathbb{Z}) \curvearrowright \mathfrak{h}$, and note that $t^2$ corresponds to $\Im \tau$ and $u$ corresponds to $\Re \tau$). Here we have written

$n_u := \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}$,

$a_t := \begin{pmatrix} t^{-1} & 0 \\ 0 & t \end{pmatrix}$.
We note that $\mathcal{F}$ is a fundamental domain for $G(\mathbb{Z}) \lhd G(\mathbb{R})$.

Let $\delta \in \mathbb{R}^+$ be a small absolute constant (e.g., $\delta := 10^{-10^9}$ will certainly suffice). Let $\eta \in \mathbb{R}$ (we will eventually take $\eta := \pm X^{-\delta}$). Let $\mu : \mathbb{R} \to [0, 1]$ be $C^\infty$, such that $\mu = 0$ outside $[-1 - 2\eta, 1 + 2\eta]$, and such that $\mu = 1$ on $[-1 - 2\eta + |\eta|, 1 + 2\eta - |\eta|]$ — thus when $\eta < 0$ it follows that $\mu \leq 1_{[-1, 1]}$ and when $\eta > 0$ it follows that $\mu \geq 1_{[-1, 1]}$. We may of course choose $\mu$ to be the standard bump function with these properties, in which case the $C^k$ norms of $\mu$ satisfy $||\mu||_{C^k(\mathbb{R})} \ll_k \eta^{-k}$, but let us only specialize to this choice later.

Let $\alpha \in C^\infty_c(G(\mathbb{R}))$ and $\beta \in C^\infty_c(L)$ be compactly supported smooth functions such that: $\alpha$ is $SO_2(\mathbb{R})$-invariant, $\int_{G(\mathbb{R})} \alpha = 1$, $\beta(F_\pm) = 1$, and $\text{supp} \beta, \beta^{-1}(\{1\}) \subseteq L^{(1)}$ are both unions of two small compact intervals respectively containing $F_\pm$.

Let $G_0 := \text{supp} \beta$. Let

$$\varphi(v) := \sum_{\gamma v_L = v} \alpha(g) \cdot \beta(v_L).$$

Write $V(\mathbb{Z})^{\Delta \neq 0} := \{F \in V(\mathbb{Z}) : \Delta(F) \neq 0\}$ and

$$V(\mathbb{Z})^{\text{nontriv}} := \{F \in V(\mathbb{Z})^{\Delta \neq 0} : 0 \notin F(\mathbb{P}^1(\mathbb{Q}))\}.$$

(Here we deviate from Ruth, and indeed Bhargava [5] and Bhargava-Shankar [9], in using the superscript nontriv instead of irreducible, since in our view it is misleading to call these irreducible.) That is, $V(\mathbb{Z})^{\text{nontriv}}$ is the subset of $V(\mathbb{Z})$ consisting of binary quartic forms with no root in $\mathbb{P}^1(\mathbb{Q})$ — i.e., those binary quartics that do not have a linear factor defined over $\mathbb{Q}$.

Write

$$Y(\mathbb{Z}) := \{F \in V(\mathbb{Z}) : I(F) = 0, J(F) \neq 0\}$$

and $Y(\mathbb{Z})^{\text{nontriv}} := Y(\mathbb{Z}) \cap V(\mathbb{Z})^{\text{nontriv}}$. We note that our $Y(\mathbb{Z})$ and $Y(\mathbb{Z})^{\text{nontriv}}$ play the role of Ruth’s $Y$ and $Y^{\text{irr}}$, respectively. Similarly, for $M \in \mathbb{Z}^+$ and $F_0 \in V(\mathbb{Z}/M)$, write

$$Y_{F_0 \pmod{M}}(\mathbb{Z}) := \{F \in Y(\mathbb{Z}) : F \equiv F_0 \pmod{M}\}$$

and $Y_{F_0 \pmod{M}}(\mathbb{Z})^{\text{nontriv}} := Y_{F_0 \pmod{M}}(\mathbb{Z}) \cap V(\mathbb{Z})^{\text{nontriv}}$.

Write, for $S \subseteq V(\mathbb{Z})$,

$$\#_{\varphi, \mu}|B(u, t, \lambda, X) \cap S| := \sum_{F \in S} \mu \left(\frac{J(F)}{X}\right) \cdot \varphi(a_{f_1}^{-1} \cdot n_{\mu}^{-1} \cdot (\lambda \cdot \text{id})^{-1} \cdot F),$$

where we have disambiguated the action of $\lambda$ (which should really be the action of $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$) by writing $\lambda \cdot \text{id}$ for clarity. We note that this is different from Ruth’s (and indeed Bhargava’s) notation precisely because we have smoothed using $\alpha$, $\beta$, and $\mu$ rather than simply $1_{G_0}$ — though the reader may well imagine that $\varphi \sim 1_{G_0 \cdot L}$ in which case $\#_{\varphi, \mu}|B(u, t, \lambda, X) \cap S| \sim \#|\lambda \cdot n_{\mu} \cdot a_1 \cdot G_0 \cdot L \cap S|$.

Just as in Section 2.2 of Ruth’s [12], we find that the problem reduces to controlling $\#_{\varphi, \mu}|B(u, t, \lambda, X) \cap Y_{F_0 \pmod{M}}(\mathbb{Z})^{\text{nontriv}}|$. We do this with the following
two lemmas.

Lemma 4.1 (The "tail" estimate.). Let 

\[ \lambda \in \mathbb{R}^+, \ u \in \left[-\frac{1}{2}, \frac{1}{2}\right], \ \sqrt{\frac{3}{2}} \leq t \ll \lambda. \]

Then:

\[ \#_{\phi, \mu} B(u, t, \lambda, X) \cap (Y(Z)_{\text{nontriv}}) \ll_{\phi, \lambda} \lambda^{12 + o(1)}. \]

Lemma 4.2 (The "bulk" estimate.). Let \( M \in \mathbb{Z}^+ \) and \( F_0 \in V(\mathbb{Z}/M) \). Let 

\[ \lambda \in \mathbb{R}^+, \ u \in \left[-\frac{1}{2}, \frac{1}{2}\right], \ \sqrt{\frac{3}{2}} \leq t \ll \lambda. \]

Then:

\[ \#_{\phi, \mu} B(u, t, \lambda, X) \cap Y_F(\mathbb{Z}) \| = \sigma_{\infty}(u, t, \lambda, X) \cdot \prod_p \sigma_p(Y_F(\mathbb{Z})) 
+ O_{\phi, M}(t^4 \cdot \lambda^{8+o(1)}) + O_{\phi, M, N}(|\|C^0 N(\mathbb{R}) \cdot t^N \cdot \lambda^{O(1)-N}), \]

where

\[ \sigma_{\infty}(u, t, \lambda, X) := \lim_{\epsilon \to 0} \int_{v \in V(\mathbb{R}) : |I(v)| \leq \epsilon} dv \mu \left( \frac{J(v)}{X} \right) \cdot \varphi(a^{-1} \cdot n^{-1} \cdot (\lambda \cdot \text{id})^{-1} \cdot v). \]

and

\[ \sigma_p(Y_F(\mathbb{Z})) := \lim_{k \to \infty} p^{-4k} \cdot \# \{ F \in V(\mathbb{Z}/p^k) : I(F) \equiv 0 (\mod p^k), F \equiv F_0 (\mod M) \}. \]

We have written \( \sigma_{\infty}(u, t, \lambda, X) \) despite the function being independent of \( u \) and \( t \) (via \( v \mapsto n_u \cdot a_t \cdot v \)) for notational convenience.

Just as in e.g. Theorems 2.12 and 2.21 of Bhargava-Shankar’s [9], we must introduce a weight function \( \phi : V(\mathbb{Z}/M) \to \mathbb{R} \) (which will eventually be taken to be a majorant of, in their notation, \( f \mapsto 1_{m}(f) \)) with \( M \) highly divisible. The following weighted version of Lemma 4.2 of course follows immediately from Lemma 4.2.

\footnote{Note that Lemma 4.2 matches Ruth’s Proposition 2.3.1 in form, except that we have included a congruence condition in order to sieve to locally soluble forms — Ruth overlooks doing this in his circle method analysis. We note also that Ruth overlooks a factor of the form \( \sigma_{\infty}(u, t, \lambda, X) \) (specifically he overlooks the dependence on both \( \lambda \) and \( X \); on page 12 of [12] he states that the condition \( |J(v)| < X \) is superfluous once \( \lambda \ll X^{1/4} \), thus one can drop it — this is false, since his definition of his \( F^* \) depends on a parameter \( C' \) and were this to be the case the Selmer average would also grow with \( C' \), rather than be 3 — one can think of our smoothing as solving this issue. Let us however emphasize here, lest our phrasing suggest otherwise, that such oversights are not out of the ordinary in original works like Ruth’s, and that we hope we have made clear the accuracy of the idea contained therein.)}
Lemma 4.3. Let $M \in \mathbb{Z}^+$ and $\phi : V(\mathbb{Z}/M) \to \mathbb{R}$. Let

$$\lambda \in \mathbb{R}^+, \ u \in \left[ -\frac{1}{2}, \frac{1}{2} \right], \ \sqrt{\frac{3}{2}} \leq t \ll \lambda.$$ 

Then:

$$\sum_{F_0 \in V(\mathbb{Z}/M)} \phi(F_0) \cdot \#_{\varphi,u}|B(u, t, \lambda, X) \cap Y_{F_0} \,(\text{mod} \ M)(\mathbb{Z})|$$

$$= \sigma_\infty(u, t, \lambda, X) \cdot \prod_{p|M} \sigma_p(Y(\mathbb{Z})) \cdot \lim_{k \to \infty} M^{-4k} \cdot \sum_{F \in V(\mathbb{Z}/M^k): I(F) \equiv 0 \,(\text{mod} \ M^k)} \phi(F \,(\text{mod} \ M)) + O_{\phi, M, N} (||\varphi||_{C^0(N)}(B) \cdot t^N \cdot \lambda^{O(1)-N}).$$

We note that the $p$-adic local densities are of course exact analogues of the singular integral at infinity — one thickens $p$-adically by forcing only $I(F) \equiv 0 \,(\text{mod} \ p^k)$, and then one takes a limit. It is because of this thickening that we easily reduce the calculation of the constants to results of Bhargava-Shankar.

### 4.2 Avoiding a calculation of local densities.

Let us first deduce Theorem 1.1 for $k \equiv 1 \,(\text{mod} \ 6)$ from Lemmas 4.1 and 4.2.

**Proof of Theorem 1.1 for $k \equiv 1 \,(\text{mod} \ 6)$ assuming Lemmas 4.1 and 4.2.** We reduce immediately to the case of $B$ defined by finitely many congruence conditions. Indeed, assume Theorem 1.1 for $k \equiv 1 \,(\text{mod} \ 6)$ for nonempty subsets of $\mathbb{Z} - \{0\}$ defined by finitely many congruence conditions. Writing $B_p$ for the closure of $B$ in $\mathbb{Z}_p$, and $B_{\leq T} := \{n \in \mathbb{Z} - \{0\} : \forall p \leq T, n \in B_p\}$ (thus $B \subseteq B_{\leq T}$), we find that, assuming Theorem 3.1.1 for sets defined by finitely many congruence conditions (and thus for $B_{\leq T}$):

$$\sum_{B \in B : |B| \leq X} \#|\text{Sel}_2(E_0, B/\mathbb{Q})| \leq (3 + O_{B, T}(o_{X \to \infty}(1)) \cdot \#|\{n \in B_{\leq T} : |n| \leq X\}|$$

$$= (3 + O_{B, T}(o_{X \to \infty}(1)) \cdot \left( \frac{\#|\{n \in B_{\leq T} : |n| \leq X\}|}{\#|\{n \in B : |n| \leq X\}|} \right) \cdot \#|\{n \in B : |n| \leq X\}|$$

Taking $X \to \infty$ and then $T \to \infty$ gives that

$$\text{Avg}_{n \in B : |n| \leq X} \#|\text{Sel}_2(E_0, B/\mathbb{Q})| \leq 3 + O_{B}(o_{X \to \infty}(1)),$$

as desired.

12
Thus without loss of generality $B$ is defined by finitely many congruence conditions, i.e. it is of the form $B = \{ n \in \mathbb{Z} - \{0\} : n \equiv a \pmod{m} \}$ for $m \in \mathbb{Z}^+$ and $a \in \mathbb{Z}/m$.

We will appeal to Ruth’s Lemma 2.2.3 (proven in Section 4.3 of his [12] using an analysis of the (monogenized) cubic resolvent ring arising from a binary quartic form) to use the equality $n(F) = m(F)$ outside a negligible set. Here we use the notation of Section 3.2 of Bhargava-Shankar’s [9]; $n(F)$ is the number of $PGL_2(\mathbb{Q})$-orbits in the (intersection of $V(\mathbb{Z})$ and the $) PGL_2(\mathbb{Q})$-orbit of $F \in V(\mathbb{Z})^{\Delta \neq 0}$, and

$$m(F) := \sum_{F' \in PGL_2(\mathbb{Z}) \setminus (V(\mathbb{Z}) \cap PGL_2(\mathbb{Z}))} \frac{\# \text{Aut}_{PGL_2(\mathbb{Q})} (F')} {\# \text{Aut}_{PGL_2(\mathbb{Z})} (F')}$$

$$= \prod_p \sum_{F' \in PGL_2(\mathbb{Z}) \setminus (V(\mathbb{Z}) \cap PGL_2(\mathbb{Z}))} \frac{\# \text{Aut}_{PGL_2(\mathbb{Q})} (F')} {\# \text{Aut}_{PGL_2(\mathbb{Z})} (F')}$$

$$=: \prod_p m_p(F).$$

Let $M \in \mathbb{Z}^+$ with $m|M$. Let $\phi : V(\mathbb{Z}/M) \to \mathbb{R}$. We will eventually take a sequence of majorants $\phi_n$ of $F \mapsto \frac{1}{m(F)}$ — i.e. $\phi_1(F) \geq \cdots \geq \phi_n(F) \geq \cdots \geq \frac{1}{m(F)}$ for all $F \in V(\mathbb{Z})^{\Delta \neq 0}$ — and apply the below to $\phi_n$ and take $n \to \infty$.

As we have seen in e.g. Section 3

$$N_\mu(Y_{F_0}(\text{mod } M))(\mathbb{Z})^{\text{nontriv}}, X)$$

$$= \int_F \frac{dh}{v \in V(\mathbb{Z})^{\text{nontriv}} : I(v) = 0, F \equiv F_0 \pmod{M}}\sum_{v \in V(\mathbb{Z})^{\text{nontriv}} : I(v) = 0, F \equiv F_0 \pmod{M}} \mu \left( \frac{J(v)}{X} \right) \cdot \varphi(h^{-1} \cdot v)$$

$$= \int_{1 < \lambda < X^{\frac{1}{12}}} d \lambda \int_{u \in (\lambda / t)} du \int_{t \leq t \leq \lambda} t^{-2} d^x t \#_{\varphi, \mu} |B(u, t, \lambda, X) \cap Y_{F_0}(\text{mod } M)(\mathbb{Z})^{\text{nontriv}}|$$

$$= \int_{1 < \lambda < X^{\frac{1}{12}}} d \lambda \int_{u \in (\lambda / t)} du \int_{t \leq t \leq \lambda^d / t} t^{-2} d^x t \#_{\varphi, \mu} |B(u, t, \lambda, X) \cap Y_{F_0}(\text{mod } M)(\mathbb{Z})^{\text{nontriv}}|,$$

$$= \int_{1 < \lambda < X^{\frac{1}{4}}} d \lambda \int_{u \in (\lambda / t)} du \int_{t \leq t \leq \lambda^d / t} t^{-2} d^x t \#_{\varphi, \mu} |B(u, t, \lambda, X) \cap Y_{F_0}(\text{mod } M)(\mathbb{Z})^{\text{nontriv}}|,$$

$$= \int_{1 < \lambda < X^{\frac{1}{4}}} d \lambda \int_{u \in (\lambda / t)} du \int_{t \leq t \leq \lambda^d / t} t^{-2} d^x t \sigma_\infty(u, t, \lambda, X) \cdot \prod_{p} \sigma_p(Y_{F_0}(\text{mod } M)(\mathbb{Z}))$$

$$+ O_{\varphi, M}(\|v\|_{C^0(\mathbb{Z})} \cdot t^4 \cdot \lambda^{8 + o(1)})$$

by definition, by splitting the integral, and then by Lemmas 4.1 and 4.2.

Next we pull out the product of local densities at finite primes in order to
calculate the integral. We find that:
\[
N_\mu(Y_{F_0} \mod M)(\mathbb{Z})^{nontriv}, X)
= \prod_p \sigma_p(Y_{F_0} \mod M)(\mathbb{Z}) \cdot \int_{1 \leq \lambda \leq X} d\lambda \int_{u \in v(t)} du \int_{t \in \lambda} t^{-2} d\lambda t \sigma_\infty(u, t, \lambda, X)
+ O_{\varphi, M}(\|\mu\|_{C^0(1)}(\mathbb{R}) \cdot X^{1-\Omega(\delta) + o(1)})
\]
\[
= \prod_p \sigma_p(Y_{F_0} \mod M)(\mathbb{Z}) \cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_F d\lambda du t^{-2} \int_{v \in V(\mathbb{R}): |I(v)| \leq \varepsilon} dv \mu \left( \frac{J(v)}{X} \right) \cdot \varphi((\lambda \cdot a_t \cdot n_u)^{-1} \cdot v)
+ O_{\varphi, M}(\|\mu\|_{C^0(1)}(\mathbb{R}) \cdot X^{1-\Omega(\delta) + o(1)})
\]
\[
= \prod_p \sigma_p(Y_{F_0} \mod M)(\mathbb{Z}) \cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_{v \in V(\mathbb{R}): |I(v)| \leq \varepsilon} dv \mu \left( \frac{J(v)}{X} \right) \int_{h \in F} dh \varphi(h^{-1} \cdot v)
+ O_{\varphi, M}(\|\mu\|_{C^0(1)}(\mathbb{R}) \cdot X^{1-\Omega(\delta) + o(1)})
\]
by definition of \(\sigma_\infty\) and then by letting \(h := (\lambda \cdot id) \cdot a_t \cdot n_u\) and recalling that \(\varphi\) is \(SO_2(\mathbb{R})\)-invariant.

Recall that we saw in Section 3 that, for \(v \in V(\mathbb{R})^{\Delta \neq 0}\) with \(I(v) = 0\) (recall that then \(\beta(v_L) = 1\)),
\[
\int_{h \in F} dh \varphi(h^{-1} \cdot v) = \int_{G(\mathbb{R})} dh \alpha(h) \cdot \#\{|g \in F : gh \cdot v_L = v\}.
\]

Because we arranged that \(\beta = 1\) on sufficiently small intervals around \(F_{\pm}\), the same identity holds for \(v \in V(\mathbb{R})^{\Delta \neq 0}\) with \(|I(v)| \leq \varepsilon\) and \(|J(v)| \gg 1\) when \(\varepsilon\) is sufficiently small. Inserting this into the above, we find that the main term is:
\[
(1 + O_{M}(X^{-1})) \cdot N_\mu(Y_{F_0} \mod M)(\mathbb{Z})^{nontriv}, X)
= \prod_p \sigma_p(Y_{F_0} \mod M)(\mathbb{Z})
\]
\[
\cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_{G(\mathbb{R})} dh \alpha(h) \int_{v \in V(\mathbb{R}): |I(v)| \leq \varepsilon, |J(v)| \gg 1} dv \mu \left( \frac{J(v)}{X} \right) \cdot \#\{|g \in F : gh \cdot v_L = v\}.
\]

by switching integrals and then noting that \(F \cdot h\) is a fundamental domain for the action of \(G(\mathbb{Z}) \cap G(\mathbb{R})\) (and using \(\int_{G(\mathbb{R})} \alpha = 1\)). Note also that the factor \((1 + O_{M}(X^{-1}))\) in the first line comes from the fact that we have thrown out the \(v \in V(\mathbb{R})^{\Delta \neq 0}\) with \(|I(v)| \leq \varepsilon\) and \(|J(v)| \ll 1\) in order to use the equality \(\int_{h \in F} dh \varphi(h^{-1} \cdot v) = \int_{G(\mathbb{R})} dh \alpha(h) \cdot \#\{|g \in F : gh \cdot v_L = v\}|.\)
Now the inner integral is independent of the choice of fundamental domain, whence:

\[
\begin{align*}
(1 + O_M(X^{-1})) \cdot N_\mu(Y_{F_0 \mod M}(\mathbb{Z})^{\text{nontriv}}, X) \\
= \prod_p \sigma_p(Y_{F_0 \mod M}(\mathbb{Z})) \cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_{v \in \mathcal{F} \setminus L \mid \varepsilon, |J(v)| \gg 1} \frac{dv \mu}{X} \\
= 2X \cdot \prod_p \sigma_p(Y_{F_0 \mod M}(\mathbb{Z})) \cdot \lim_{\varepsilon \to 0} (4\varepsilon \cdot X)^{-1} \cdot \int_{v \in \mathcal{F} \setminus L \mid \varepsilon, |J(v)| \gg 1} \frac{dv \mu}{X} \\
= \frac{2X}{m} \cdot \left( m \cdot \prod_p \sigma_p(Y_{F_0 \mod M}(\mathbb{Z})) \right) \cdot \lim_{\varepsilon \to 0} (4\varepsilon \cdot X)^{-1} \cdot \int_{v \in \mathcal{F} \setminus L \mid \varepsilon, |J(v)| \gg 1} \frac{dv \mu}{X},
\end{align*}
\]

Now let us evaluate the remaining integral. By Proposition 2.8 of Bhargava-Shankar’s [9] (we could avoid using this to get from the leftmost to the rightmost terms in the equality, of course), we have that

\[
\begin{align*}
(4\varepsilon \cdot X)^{-1} \cdot \int_{v \in \mathcal{F} \setminus L \mid \varepsilon, |J(v)| \gg 1} \frac{dv \mu}{X} \\
= (1 + O(\eta)) \cdot (4\varepsilon \cdot X)^{-1} \cdot \int_{v \in \mathcal{F} \setminus L \mid \varepsilon, 1 \ll |J(v)| \ll (1 + O(\eta)) \cdot X} dv \\
= \frac{2\zeta(2)}{27} \cdot (1 + O(\eta + \varepsilon \cdot X^{-1})) \\
= (1 + O(\eta + \varepsilon \cdot X^{-1})) \cdot \frac{\int_{\|\| \mathcal{R}(\mathbb{X}) dv}{\int_{\|\| \mathcal{R}(\mathbb{X}) dIdJ},
\end{align*}
\]

using the notation of Section 2.4 of Bhargava-Shankar’s [9]. This concludes the first step in the trick of reducing the calculation of the product of local densities to the one done in Bhargava-Shankar — at the moment, we have only dealt with the Archimedean restrictions.

Thus so far we have found that:

\[
N_\mu(Y_{F_0 \mod M}(\mathbb{Z})^{\text{nontriv}}, X) = (1 + O_M(X^{-1})) \cdot \frac{2X}{m} \left( m \cdot \prod_p \sigma_p(Y_{F_0 \mod M}(\mathbb{Z})) \right) \cdot \frac{\int_{\|\| \mathcal{R}(\mathbb{X}) dIdJ}{\int_{\|\| \mathcal{R}(\mathbb{X}) dIdJ},
\]

Note that

\[
\#\{|J \in \mathcal{B} : 0 \neq |J| \leq X\} = \frac{2X}{m} \cdot \left( 1 + O\left(\frac{m}{X}\right) \right).
\]
Just as in the passage from Lemma 4.2 to Lemma 4.3, we find that:

\[
\sum_{F_0 \in V(\mathbb{Z}/M)} \phi(F_0) \cdot N_p(Y_{F_0} \pmod{M}(\mathbb{Z}))^{\text{nontriv}}, X
\]

\[
= (1 + O(\eta) + O_M(X^{-1})) \cdot \frac{2X}{m} \int_{\mathbb{I}_r} R^{(i)}(X^2/4) \frac{dv}{dIdJ}
\]

\[
\cdot \left( m \cdot \lim_{T \to \infty} n_T^{-4} \cdot \sum_{F \in V(\mathbb{Z}/nT); I(F) \equiv 0 \pmod{nT}, J(F) \equiv a \pmod{m}} \phi(F \pmod{M}) \right),
\]

where \( n_T := \prod_{p \leq T} p^T \), say, and without loss of generality \( T \geq M \) so that \( M | n_T \).

Choose now \( \eta := X^{-\delta^2} \) and \( \mu \) to be the standard bump function with the given properties, whence \( \|\mu\|_{L^{O(\delta^2)}(\mathbb{R})} \ll X^{O(\delta^2)} \).

Write \( \mathbb{1}^{(p)}_{\text{sol}} \) for the indicator function of the \( \mathbb{Q}_p \)-soluble binary quartic forms \( F \in V(\mathbb{Q}_p) \) (that is to say, those \( F \) for which \( Z^2 = F(X, Y) \) admits a nonzero solution with all coordinates in \( \mathbb{Q}_p \)). Write \( \mathbb{1}^{(\text{loc. sol.})}_{\text{sol}} := \prod_p \mathbb{1}^{(p)}_{\text{sol}} \). Write

\[
\phi_{\ast}(F) := \frac{\mathbb{1}^{\text{loc. sol.}}_{\text{sol}}(F)}{m(F)} = \prod_p \frac{\mathbb{1}^{(p)}_{\text{sol}}(F)}{m_p(F)} = \prod_p \phi_p(F)
\]

on \( V(\mathbb{Z})^{\Delta \neq 0} \). Thus when \( 1 =: \phi_0(F) \geq \phi_1(F) \geq \cdots \geq \phi_{\ast}(F) \) for all \( F \in V(\mathbb{Z})^{\Delta \neq 0 \text{ with }}(\mathbb{Z}_p)^{\Delta \neq 0} \to [0,1] \) factoring through \( V(\mathbb{Z}/p^n) \) (and not \( V(\mathbb{Z}/p^{n-1}) \)) and such that \( \phi_0(F) \to \phi_0(F) \) as \( n \to \infty \), we find that writing

\[
\phi_n := \prod_{p \leq n} \phi_0(F),
\]

(and here is where we use Lemma 2.2.3 of Ruth’s [12] to replace \( m(F) \) by \( n(F) \) for all but \( \ll X^{\frac{1}{2}} \) forms):

\[
\frac{\text{Avg}}{B \in \mathcal{B} | B| \leq X} \#|\text{Sel}_2(E_0, B/\mathbb{Q}) - \{0\}|
\leq (1 + O_T(a_{\mathbb{X} \to \infty}(1))) \cdot (1 + o_{T \to \infty}(1))
\]

\[
\cdot \frac{1}{\mathbb{I}_r} \int_{\mathbb{I}_r} R^{(i)}(X^2/4) \frac{dv}{dIdJ}
\cdot \left( \frac{m \cdot n_T^{-4} \cdot \sum_{F \in V(\mathbb{Z}/nT); I(F) \equiv 0 \pmod{nT}, J(F) \equiv a \pmod{m}} \phi_T(F) \right).
\]

\( ^8 \)Of course one always has that, for a convergent sequence \( x_k \in \mathbb{R} \), \( \lim_{k \to \infty} x_k = x_k + o_{k \to \infty}(1) \), but here we have written \( \lim_{k \to \infty} x_k = (1 + o_{k \to \infty}(1)) \cdot x_k \), which is only justified (for \( k \) sufficiently large) when \( \lim_{k \to \infty} x_k \neq 0 \). While we will see that the relevant limit is 2, technically we are not yet justified in doing this and should carry the various additive error terms \( o_{k \to \infty}(1) \) through the argument. However we hope the reader will grant us this notational simplification, since it makes no difference to the argument.
Note that we have "thickened" by changing the constraint $I(F) = 0$ to $I(F) \equiv 0 \pmod{n_T}$. Accordingly, we write
\[ \text{Inv}^{(T,B)} := \{(I, J) \in \mathbb{Z}^2 : I \equiv 0 \pmod{n_T}, J \equiv a \pmod{m}, 4I^3 - J^2 \neq 0\}, \]
and $\text{Inv}^{(T,B)}_{p_T} \subseteq \mathbb{Z}_p^2$ for its closure in $\mathbb{Z}_p^2$.

The trick is now to notice that, for all $k \in \mathbb{Z}^+$,
\[
\sum_{F \in V(\mathbb{Z}/n_T^k) : I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F) \equiv n_T^{-5k-1} \cdot \sum_{F \in V(\mathbb{Z}/n_T^k) : I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F),
\]
so that:
\[
m \cdot n_T^{-4} \cdot \sum_{F \in V(\mathbb{Z}/n_T^k) : I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F) = \frac{n_T^{-5k-1} \cdot \sum_{F \in V(\mathbb{Z}/n_T^k) : I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F)}{n_T^{-2k} \cdot \#(\{(I, J) \in V(\mathbb{Z}/n_T^k) : I \equiv 0 \pmod{n_T}, J \equiv a \pmod{m}\})}.
\]

Now we note that
\[
\lim_{k \to \infty} n_T^{-5k-1} \cdot \sum_{F \in V(\mathbb{Z}/n_T^k) : I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F) = \int_{F \in V(\prod_{p \leq T} \mathbb{Z}_p) : (I(F), J(F)) \in \prod_{p \leq T} \text{Inv}^{(T,B)}_{p_T}} dF \phi_T(F),
\]
and so, by dominated convergence and Fubini (note that we are implicitly using Proposition 3.18 of Bhargava-Shankar’s [9], and indeed arguing as in their proof of Proposition 2.21 of their [9]), it follows that
\[
n_T^{-5k} \cdot \sum_{F \in V(\mathbb{Z}/n_T) : I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F) = (1 + o_{T \to \infty}(1)) \cdot (1 + O_T(\sigma_{k \to \infty}(1))) \cdot \prod_p \int_{v \in V(\mathbb{Z}_p) : (I(v), J(v)) \in \text{Inv}^{(T,B)}_{p_T}} dv \phi^{(p)}_T(v).
\]

Similarly, we of course have:
\[
n_T^{-2k} \cdot \#(\{(I, J) \in V(\mathbb{Z}/n_T^k) : I \equiv 0 \pmod{n_T}, J \equiv a \pmod{m}\}) = \prod_{p \leq T} \int_{I,J} dI dJ = (1 + o_{T \to \infty}(1)) \cdot \prod_p \int_{I,J} dI dJ.
\]

\(^9\)See the previous footnote (about factoring out $(1 + o_{k \to \infty}(1))$ and $(1 + o_{T \to \infty}(1))$).
Combining all these, we find that:

\[
\text{Avg}_{B \in \mathcal{B} : |B| \leq X} \frac{\# \text{Sel}_2(E_0,B/\mathbb{Q}) - \{0\}}{|B|} \\
\leq (1 + O_T(o_X \to \infty(1))) \cdot (1 + o_T \to \infty(1)) \cdot (1 + O_T(o_k \to \infty(1))) \\
\cdot \frac{\int_{\mathbb{PGL}_2} \mathcal{R}((X^2)/4) \, dv}{\int_{\mathbb{PGL}_2} \mathcal{R}((X^2)/4) \, dv} \cdot \frac{\prod_p \int_{v \in V(\mathbb{Z}_p)}: (I(v),J(v)) \in \text{Inv}_p^{(T,s)} \, dv \phi^{(p)}(v)}{\prod_p \int_{(I,J) \in \text{Inv}_p^{(T,s)}} \, dIdJ}.
\]

But the calculations in Section 3.6 of Bhargava-Shankar’s [9] amount to the statement that

\[
\frac{\int_{\mathbb{PGL}_2} \mathcal{R}((X^2)/4) \, dv}{\int_{\mathbb{PGL}_2} \mathcal{R}((X^2)/4) \, dv} \cdot \frac{\prod_p \int_{v \in V(\mathbb{Z}_p)}: (I(v),J(v)) \in \text{Inv}_p^{(T,s)} \, dv \phi^{(p)}(v)}{\prod_p \int_{(I,J) \in \text{Inv}_p^{(T,s)}} \, dIdJ} = 2.
\]

Taking \( k \to \infty \), then \( X \to \infty \), and finally \( T \to \infty \), we deduce Theorem 1.1 for \( k \equiv 1 \pmod{6} \). \( \square \)

### 4.3 The uniformity estimate.

In order to prove the matching lower bound (recall that we are in the special case where \( \mathcal{B} \subseteq \mathbb{Z} - \{0\} \) is defined by finitely many congruence conditions) we simply run the above argument with minorants instead — that is to say, we instead take \( \eta := -X^{-\sigma^2} \), \( \mu \) again the standard bump function with the given properties, and \( 0 =: \phi^{(p)}(0) \leq \phi^{(p)}(F) \leq \cdots \leq \phi^{(n)}(F) \leq \cdots \leq \phi^{(p)}(F) \) for all \( F \in V(\mathbb{Z}_p)^{\Delta \not= 0} \), with \( \phi_n : V(\mathbb{Z}_p)^{\Delta \not= 0} \to [0,1] \) factoring through \( V(\mathbb{Z}/p^n) \) (and not \( V(\mathbb{Z}/p^{n-1}) \)) and such that \( \phi_n^{(p)}(F) \to \phi^{(p)}(F) \) as \( n \to \infty \). The argument is precisely the same, with the exception of the first step.

Specifically, when proving the upper bound we implicitly used that the binary quartics \( F \in V(\mathbb{Z})^{\text{nontriv}} \), with \( I(F) = 0 \) representing 2-Selmer classes of \( E_0, I(F)/\mathbb{Q} \) — i.e. such that \( \mathbb{Z}^2 = F(X,Y) \) is nontrivially soluble in all completions of \( \mathbb{Q} \) — are in particular locally soluble at those \( p \leq T \). However of course the converse does not hold. Similarly, in order to deduce a lower bound for the average of \( \phi \), we must have that \( \phi_T \leq \phi \), outside a set of negligible size. So, just as in Bhargava-Shankar’s [9], we need only prove that the integral with respect to Haar measure of the number of binary quartics \( F \) in \( B(u,t,\lambda,X) \cap Y(\mathbb{Z}) \) for which there is a prime \( p \) with \( p > \Pi \) such that \( F \) is not locally soluble or \( \text{PGL}_2(\mathbb{Z}_p) \cdot F \not= V(\mathbb{Z}_p) \cap \text{PGL}_2(\mathbb{Q}_p) \cdot F \) is

\[
\ll \frac{\lambda^{12+o(1)}}{\Pi \log \Pi}
\]

when e.g. \( \Pi \leq \lambda^{10^{-10}} \) and we are in the "bulk", so that \( t \ll \lambda^{o(1)} \).

Now, just as in the proof of Proposition 3.18 of Bhargava-Shankar’s [9], if \( F \in V(\mathbb{Z})^{\text{nontriv}} \) is a binary quartic that is not locally soluble at \( p \), then \( F \mod p \) has splitting type one of \( (1^2 1^2), (2^2), \) or \( (1^4) \). But if moreover \( I(F) = \)
0, and thus $I(F) \equiv 0 \pmod{p}$, one gets much more: the splitting types $(1^2 1^2)$ and $(2^2)$ are not possible, as one can see by e.g. explicit calculation. Thus $F \pmod{p}$ must be a fourth power of a linear form, which is to say that $F \pmod{p}$ lies on the codimension 3 subvariety $Z \subseteq V$ given by fourth powers of linear forms, namely the affine cone on a rational normal curve.

Similarly if $\text{PGL}_2(\mathbb{Z}_p) \cdot F \neq V(\mathbb{Z}_p) \cap \text{PGL}_2(\mathbb{Q}_p) \cdot F$, then, as in Bhargava-Shankar’s [9], $p^2 \mid \Delta(F)$, and indeed, up to $(X,Y) \mapsto (Y,X)$, there is a $\gamma \in \text{GL}_2(\mathbb{Z}_p)$ such that $\text{diag}(1,p) \cdot \gamma \cdot F \in V(\mathbb{Z}_p)$, which is to say that $(\gamma \cdot F)(X,Y) =: a \cdot X^4 + b \cdot X^3 Y + c \cdot X^2 Y^2 + d \cdot XY^3 + e \cdot Y^4$ has $p \mid d$ and $p^2 \mid e$. Since $I(\gamma \cdot F) = 0$ this means that $p \mid c$, and so using $I(\gamma \cdot F) = 0$ again we see that $p^2 \mid b \cdot d$, whence either $p \mid b$ and $\gamma \cdot F$ (and thus so too $F$) has splitting type $(1^4)$, or else $p^2 \mid d$ in which case $\text{diag}(1,p) \cdot \gamma \cdot F$ has splitting type $(1^4)$, and of course the same holds after reduction via $\text{SL}_2(\mathbb{Z})$.

So it follows that in order to obtain the desired bound on the integral with respect to Haar measure of the number of $F \in B(u,t,\lambda,X) \cap Y(\mathbb{Z})$ which are not locally soluble or have $\text{PGL}_2(\mathbb{Z}_p) \cdot F \neq V(\mathbb{Z}_p) \cap \text{PGL}_2(\mathbb{Q}_p) \cdot F$ at some $p$ with $p > P$ it suffices to show that the number of $F \in B(u,t,\lambda,X) \cap Y(\mathbb{Z})$ for which the reduction $F \pmod{p}$ is locally soluble or have $\Delta(F) \neq 0$ for some $p > P$ is $\ll \frac{12 + o(1)}{\log P}$, and then the desired bound follows from invoking Theorem 1.1 of Browning-Heath-Brown’s [10].

### 4.4 Point counting.

Let us now prove Lemmas 4.1 and 4.2. We note that we will give an essentially one-line proof of Lemma 4.1 (namely, "use the divisor bound to determine $a, e$ from $b, c, d"), which subsumes the entirety of Ruth’s Section 3.5 (pages 29–37 of 12).

**Proof of Lemma 4.1** Let $F \in \lambda \cdot n_\lambda \cdot a_0 \cdot G_0 \cdot L \cap V(\mathbb{Z})^{\text{nontriv}}$. Write $F(X,Y) =: a \cdot X^4 + b \cdot X^3 Y + c \cdot X^2 Y^2 + d \cdot XY^3 + e \cdot Y^4$. Evidently (by e.g. compactness of $G_0$ and $L$) we have that:

\[
0 \neq |a| \ll \frac{\lambda^4}{t^4},
\]

\[
|b| \ll \frac{\lambda^4}{t^2},
\]

\[
|c| \ll \lambda^4,
\]

\[
|d| \ll t^2 \cdot \lambda^4,
\]

\[
0 \neq |e| \ll t^4 \cdot \lambda^4.
\]

---

Working over $\mathbb{F}_p$ and changing variables suitably, this amounts to the assertion that

\[
I(X^2 \cdot (X - n \cdot Y)^2) = I(X^4 - 2n \cdot X^3 Y + n^2 \cdot X^2 Y^2) = n^4.
\]
Therefore the number of tuples \((b, c, d) \in \mathbb{Z}^3\) among \(F \in \lambda \cdot n_u \cdot a_t \cdot G_0 \cdot L \cap V(\mathbb{Z})_{\text{nontriv.}}\) is \(\ll \lambda^{12}\).

Moreover, by hypothesis \(12ae - 3bd + c^2 = 0\), i.e. \(0 \neq 12ae = 3bd - c^2 \ll \lambda^8\). Thus \((b, c, d)\) determine \((a, e)\) up to \(\ll \lambda^{o(1)}\) choices. In other words, the map

\[
\lambda \cdot n_u \cdot a_t \cdot G_0 \cdot L \cap V(\mathbb{Z})_{\text{nontriv.}} \to \mathbb{Z}^3
\]

via

\[
a \cdot X^4 + b \cdot X^3Y + c \cdot X^2Y^2 + d \cdot XY^3 + e \cdot Y^4 \mapsto (b, c, d)
\]

has image of size \(\ll \lambda^{12}\) and fibres of size \(\ll \lambda^{o(1)}\). The lemma follows. \(\square\)

As for Lemma 4.2, we first note that it is essentially identical to Ruth’s Proposition 3.4.1 (modulo the inaccuracies in his treatment that we have already mentioned), which he states without proof.

For the reader’s convenience we will give a full proof of Lemma 4.2 anyway.

**Proof of Lemma 4.2.** Before we begin we note once again that it is not necessary to use the smoothed delta symbol method, since we are asking about zeroes of a quadric in five variables, something easily handled by the classical form of the circle method.

We follow the notation of Heath-Brown’s [11]. Let \(w_0 \in C_0^\infty(\mathbb{R})\) via

\[
w_0(x) := \begin{cases} \exp\left(-\frac{1}{1-|x|^2}\right) & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}.
\]

Let \(\omega(x) := \frac{4}{\int_{S} w_0(t)dt} \cdot w_0(x)\). Let \(h(x, y) := \sum_{q \geq 1} \omega(qx) - \omega\left(\frac{\omega(x)}{q^s}\right)\).

Note that \(h(x, y) = 0\) when \(x \gg 1 + |y|\) and that \(h(x, y) \ll x^{-1}\).

We will first detail the argument in the case of \(M = 1\) (i.e. no congruence condition) and then comment on necessary modifications to more general \(M\) and \(F_0 \in V(\mathbb{Z}/M)\) as above.

Applying Theorem 2 of Heath-Brown’s [11] with his \(n = 5\) and his \(Q = \lambda^4\), we find that:

\[
\sum_{F \in V(\mathbb{Z})} \varphi(a_t^{-1} \cdot n_u^{-1} \cdot (\lambda \cdot \text{id})^{-1} \cdot F)
\]

\[
= (\lambda^{-8} + O_N(\lambda^{-N})) \cdot \sum_{q \geq 1} \sum_{\xi \in \mathbb{Z}^2} \sum_{u \in (\mathbb{Z}/q)^*} \sum_{G \in V(\mathbb{Z}/q)} e_q(u \cdot I(G) + G \cdot \bar{c})
\]

\[
\cdot \int_{F \in V(\mathbb{R})} dF \varphi(a_t^{-1} \cdot n_u^{-1} \cdot (\lambda \cdot \text{id})^{-1} \cdot F) \cdot h\left(\frac{q}{\lambda^3}, \frac{I(F)}{\lambda^8}\right) \cdot e_q(-F \cdot \bar{c}),
\]

where we have written \(e_q(z) := e\left(\frac{z}{q}\right) := e^{\frac{2\pi iz}{q}}\).

For us the error term will consist of those terms where \(\bar{c} \neq \bar{0}\), and the terms with \(\bar{c} = \bar{0}\) will comprise the main term (in the end we will simply
cite Heath-Brown’s [10] for the analysis of the main term, which in any case is considerably simpler).

Via the change of variable $F \mapsto n_u \cdot a_t \cdot (\lambda \cdot \text{id}) \cdot F$ (note that $(\lambda \cdot \text{id}) \cdot F = \lambda^4 \cdot F$ since $F$ is homogeneous of degree 4 --- here $(\lambda \cdot \text{id}) \cdot F$ on the left-hand side indicates the action of $\lambda \cdot \text{id} \in G$ on $F \in V$ via $G \subset V$, and the · on the right-hand side denotes multiplication), we find that:

$$\int_{F \in V(\mathbb{R})} dF \varphi(a_t^{-1} \cdot n_u^{-1} \cdot (\lambda \cdot \text{id})^{-1} \cdot F) \cdot h \left( \frac{q}{\lambda^4} \cdot \frac{I(F)}{\lambda^8} \right) \cdot e_q(-F \cdot \bar{c})$$

$$= \lambda^{20} \cdot \int_{F \in V(\mathbb{R})} dF \varphi(F) \cdot h \left( \frac{q}{\lambda^4} \cdot I(F) \right) \cdot e_q(-\lambda^4 \cdot F \cdot ((n_u \cdot a_t)^\dagger \cdot \bar{c})).$$

Therefore we see that the error term is:

$$(\lambda^{12} + O_N(\lambda^{-N})) \cdot \sum_{q \geq 1} q^{-5} \sum_{\bar{a} \neq e \in V(\mathbb{Z})^*} \left( \sum_{a \in (\mathbb{Z}/q)^*} \sum_{G \in V(\mathbb{Z}/q)} e_q(a \cdot I(G) + G \cdot \bar{c}) \right)$$

$$\cdot \int_{F \in V(\mathbb{R})} dF \varphi(F) \cdot h \left( \frac{q}{\lambda^4} \cdot I(F) \right) \cdot e_q(-\lambda^4 \cdot F \cdot ((n_u \cdot a_t)^\dagger \cdot \bar{c})).$$

Note that the $\varphi(F)$ term in the integral forces $||F||_\infty \ll \lambda 1$ if the integrand is to be nonzero, and then our observation that $h(x, y) = 0$ if $x \gg 1 + |y|$ forces $q \ll \lambda^4$ as well.

Note also that if $q \ll \lambda^{4-\varepsilon} \cdot ||(n_u \cdot a_t)^\dagger \cdot \bar{c}||_\infty$ — i.e. if

$$|| (n_u \cdot a_t)^\dagger \cdot \bar{c} ||_\infty \gg \frac{q}{\lambda^{4-\varepsilon}}$$

— we find, by repeated integration by parts, that such terms contribute $O_{\varepsilon, \varphi, N}(t^N \cdot \lambda^{O(1) - N})$.

Also the complete exponential sum, which is just

$$\sum_{u \in (\mathbb{Z}/q)^*} \sum_{G \in V(\mathbb{Z}/q)} e_q(u \cdot (12G_0G_4 - 3G_1G_3 + G_2^2) + (c_0G_0+c_1G_1+c_2G_2+c_3G_3+c_4G_4)),$$

is very easy to calculate (see e.g. page 49 of [11]). We conclude that

$$\sum_{u \in (\mathbb{Z}/q)^*} \sum_{G \in V(\mathbb{Z}/q)} e_q(u \cdot I(G) + G \cdot \bar{c}) \ll \begin{cases} 0 & \exists p > 3 : v_p(q) = 1 \\ q^{2+o(1)} & \forall p | q \text{ s.t. } p > 3, v_p(q) \geq 2, \end{cases}$$

and indeed one can sharpen the bound significantly.

Now we return to the smoothed delta symbol method. Again, the integral is nonnegligible only for

$$|| (n_u \cdot a_t)^\dagger \cdot \bar{c} ||_\infty \ll \frac{q}{\lambda^{4-\varepsilon}}$$

21
(in which case it is \(\ll \lambda^{-4}\)). Thus either \(\vec{c} = \vec{0}\), in which case the corresponding summand contributes to the main term dealt with by Ruth (and by Heath-Brown in [11]), or else \(\vec{c} \neq \vec{0}\), in which case we must have that \(q \gg \lambda^{-\varepsilon}\) since by inspection \(\|((n_u \cdot a_t)^\dagger \cdot \vec{c})\|_\infty \gg t^{-4} \cdot ||\vec{c}||_\infty \gg t^{-4}\).

The error term is therefore:
\[
\ll \lambda^{12+o(1)} \cdot \sum_{\frac{\lambda^4-\varepsilon}{t^4} \ll q \ll \lambda^4, q \text{ powerful}} q^{-\frac{3}{2}}
\]
\[
\int_{\vec{O} \neq \vec{v} \in V(\mathbb{Z})^{n-1} : ||\vec{v}||_\infty \ll \frac{\lambda^3}{t^3}} dF \cdot \varphi(F) \cdot h \left( \frac{q}{\lambda^4}, I(F) \right) \cdot e_q \left( -\lambda^4 \cdot F \cdot ((n_u \cdot a_t)^\dagger \cdot \vec{c}) \right).
\]

Bounding the integrals trivially (i.e. by \(\ll \varphi \left( \frac{\lambda^3}{t^3} \right)\)) and noting that the number of \(0 \neq \vec{c} \in \mathbb{Z}^5\) such that \(\|((n_u \cdot a_t)^\dagger \cdot \vec{c})\|_\infty \ll \frac{q}{\lambda^3}t^3\) is
\[
\ll \frac{t^4 \cdot q}{\lambda^4} \cdot \left( 1 + \frac{t^2 \cdot q}{\lambda^4} \right) \cdot \left( 1 + \frac{q}{\lambda^4} \right) \cdot \left( 1 + \frac{q}{t^2 \cdot \lambda^4} \right) \cdot \left( 1 + \frac{q}{t^4 \cdot \lambda^4} \right)
\]
when \(\frac{\lambda^4-\varepsilon}{t^4} \ll q \ll \lambda^4\), we get that the error term is:
\[
\ll t^4 \cdot \lambda^{4+\varepsilon+o(1)} \cdot \sum_{\frac{\lambda^4-\varepsilon}{t^4} \ll q \ll \lambda^4, q \text{ powerful}} q^{-\frac{3}{2}} \cdot \left( 1 + \frac{t^2 \cdot q}{\lambda^4} \right) \cdot \left( 1 + \frac{q}{\lambda^4} \right) \cdot \left( 1 + \frac{q}{t^2 \cdot \lambda^4} \right) \cdot \left( 1 + \frac{q}{t^4 \cdot \lambda^4} \right)
\]
\[
\ll t^6 \cdot \lambda^{8+5\varepsilon},
\]
as desired.

Thus we have bounded the error term suitably. The required analysis of the main term is already done in Heath-Brown’s [11] (see e.g. the bottom of page 51, i.e. the end of the proof of his Theorems 4 and 5), at least in the case \(M = 1\).

We now discuss the modifications necessary for general \(M \in \mathbb{Z}^+\) and \(F_0 \in V(\mathbb{Z}/M)\). First, in the application of the smoothed delta symbol method, instead of summing over \(F \in V(\mathbb{Z})\), we sum instead over the \(F \in V(\mathbb{Z})\) for which \(F \equiv F_0 \mod M\) by summing over \(\tilde{F} \in V(\mathbb{Z})\) and writing \(F := F_0 + M \cdot \tilde{F}\) (we implicitly choose a representative of \(F_0\) in \(V(\mathbb{Z})\) and abuse notation by writing it as \(\tilde{F} \in V(\mathbb{Z})\)). We then change variables from \(\tilde{F}\) back to \(F\) in the integral and incur a factor of \(M^{-5}\). The rest of the analysis of the error term is precisely the same (except that the primes one has to treat separately in the omitted complete exponential sum calculation are now those \(p|6M\), and the error terms now depend on \(M\)).

It remains to treat the main term, i.e. the local densities. We note that, by definition, we find local densities
\[
\sigma_p^{(F_0 \mod M)}(V(\mathbb{Z})) := \lim_{k \to \infty} p^{-4k} \cdot |\{ \tilde{F} \in V(\mathbb{Z}/p^k) : I(F_0 + M \cdot \tilde{F}) \equiv 0 \mod p^k \}|.
\]

We therefore are reduced to the claim that
\[
M^{-5} \cdot \prod_p \sigma_p^{(F_0 \mod M)}(V(\mathbb{Z})) = \prod_p \sigma_p(V_{F_0}(\mod M)(\mathbb{Z})).
\]
Of course for $p ∤ M$ we have that
$$\sigma_p^{(F_0 \mod M)}(Y(Z)) = \sigma_p(Y_{F_0 \mod M}(Z)) = \sigma_p(Y(Z)),$$
via the evident change of variables $\tilde{F} \mapsto M^{-1} \cdot (\tilde{F} - F_0)$.
However, it is also evident that
$$M^{-5} \cdot \prod_{p | M} \sigma_p^{(F_0 \mod M)}(Y(Z)) = \prod_{p | M} \sigma_p(Y_{F_0 \mod M}(Z)),$$
for the following reason. For $k \in \mathbb{Z}^+$ with $k \gg 1$ we have that:

$$M^{-5} \cdot \prod_{p | M} \sigma_p(Y_{F_0 \mod M}(Z))$$

$$= M^{-5} \cdot \prod_{p | M} \left( p^{-k \cdot \nu_p(M)} \cdot \# \{ \tilde{F} \in V(\mathbb{Z}/p^{k \cdot \nu_p(M)}) : I(F_0 + M \cdot \tilde{F}) \equiv 0 \pmod{p^{k \cdot \nu_p(M)}} \} \right)$$

$$= \left( 1 + O(e^{-\Omega_M(k)}) \right) \cdot M^{-4k-5} \cdot \# \{ \tilde{F} \in V(\mathbb{Z}/M^k) : I(F_0 + M \cdot \tilde{F}) \equiv 0 \pmod{M^k} \},$$

by the Chinese remainder theorem and that fact that all $p \geq 2$.
Because the condition $I(F_0 + M \cdot \tilde{F}) \equiv 0 \pmod{M^k}$ only depends on $\tilde{F} \pmod{M^{k-1}}$, we find that:

$$M^{-4k-5} \cdot \# \{ \tilde{F} \in V(\mathbb{Z}/M^k) : I(F_0 + M \cdot \tilde{F}) \equiv 0 \pmod{M^k} \}$$

$$= M^{-4k} \cdot \# \{ F \in V(\mathbb{Z}/M^k) : I(F) \equiv 0 \pmod{M^k}, F \equiv F_0 \pmod{M} \}$$

$$= \prod_{p | M} \left( p^{-k \cdot \nu_p(M)} \cdot \# \{ F \in V(\mathbb{Z}/p^{k \cdot \nu_p(M)}) : I(F) \equiv 0 \pmod{p^{k \cdot \nu_p(M)}}, F \equiv F_0 \pmod{M} \} \right).$$

Thus taking $k \to \infty$ we find that

$$M^{-5} \cdot \prod_{p | M} \sigma_p^{(F_0 \mod M)}(Y(Z)) = \prod_{p | M} \sigma_p(Y_{F_0 \mod M}(Z)),$$
as desired. \( \square \)

5 Proof of Theorem 1.1 for $k \equiv 2 \pmod{6}$.

Now to the sketch of the proof of Theorem 1.1 when $k \equiv 2 \pmod{6}$.

5.1 Reduction to point counting.

We run an argument similar to the above, except our notation follows Bhargava-Ho’s [8] rather than Ruth’s [12], and we appeal in the end to part (f) of Theorem 1.1 of Bhargava-Ho’s [8] (rather than Theorem 3.1 of Bhargava-Shankar’s [9]).

23
to calculate the product of local densities. Because the analytic part of the argument is essentially the same as in the previous section (in fact it is easier, since in the circle method argument we deal with a quadric in eight variables instead of five) we will be significantly more terse in this section.

Again, we follow the notation in Bhargava-Ho’s [8] (the relevant parametrization is by triply symmetric hypercubes). Let \( V := 2 \otimes \Sym_3(2) \), the space of pairs of “threes-in” binary cubic forms. Let \( G \) be the image in \( \GL(V) \) of \( \{(g, h) \in \GL_2 \times \GL_2 : \det g \cdot (\det h)^3 = 1\} \), acting in the evident way on \( 2 \otimes \Sym_3(2) \) (the first \( \GL_2 \) on the first factor via the standard representation, and the second \( \GL_2 \) on the second via the induced action on \( \Sym_3 \) of the standard representation).

Note that \( G \cong (\SL_2 \times \SL_2)/\mu_2 \).

We write, for \( v \in V \),

\[
H(v) := \max \left( |I_2(v)|^{\frac{1}{\delta}}, |I_6(v)|^{\frac{1}{\delta}} \right)^{24},
\]

where \( I_2 \) and \( I_6 \) are the invariants \( a_2 \) and \( a_6 \) of Section 6.3.2 of Bhargava-Ho’s [7] and \( a_1 \) and \( a_3 \) of line 6 (corresponding to the family \( F_1(3) \)) of Table 1 in Bhargava-Ho’s [8].

Let \( R \) be a fundamental domain for \( G(\mathbb{R}) \cap V(\mathbb{R})^{\Delta \neq 0} \) (note that Bhargava-Ho write \( V(\mathbb{R})^{\text{stab}} := V(\mathbb{R})^{\Delta \neq 0} \)), as constructed in Section 5 of Bhargava-Ho’s [8] (via, in their notation, \( R := \bigsqcup_i R^{(i)} \)). Let \( L := \{v(\vec{a}) : \vec{a} \in (\mathbb{R}^m)^{\Delta \neq 0}\} \) (here in their notation \( m = 2 \) and \( v(\vec{a}) \) is as defined in Section 4 of Bhargava-Ho’s [8]) and \( \Lambda := \{(\lambda \cdot \id, \id) \in \GL_2(\mathbb{R}) \times \GL_2(\mathbb{R}) : \lambda \in \mathbb{R}^\ast\} \subseteq \GL_2(\mathbb{R}) \times \GL_2(\mathbb{R}) \). Note that \( R = \Lambda \cdot L \). Let \( R(X) := \{v \in R : H(v) \leq X\} \). Let \( \vec{F} := v(0, \pm 1) \in L \) be the two points in \( L \) with \( I_2 = 0 \).

Note that, by construction, since \( H((\lambda, \id) \cdot v) = \lambda^{24} \cdot H(v) \), the coefficients of a \( v \in \Lambda \cdot L \subseteq R(X) \) are all \( \ll \lambda \ll X^{\frac{m}{4}} \), and hence, for \( G_0 \subseteq G(\mathbb{R}) \) compact, the coefficients of a \( v \in \Lambda \cdot G_0 \cdot L \subseteq R(X) \) are all \( \ll_{G_0} \lambda \ll X^{\frac{m}{4}} \).

Let \( \mathcal{F} \) be a fundamental domain for \( G(\mathbb{Z}) \cap G(\mathbb{R}) \), as constructed in Section 5.2 of Bhargava-Ho’s [8]. Note that \( \mathcal{F} \) lies inside the following Siegel set:

\[
\mathcal{F} \subseteq N \cdot A \cdot K,
\]

where

\[
N := \left\{(n_{u_1}, n_{u_2}) \in G(\mathbb{R}) : |u_i| \leq \frac{1}{2}\right\},
\]

\[
A := \left\{(a_{t_1}, a_{t_2}) \in G(\mathbb{R}) : t_i \geq \sqrt{\frac{3}{2}}\right\},
\]

\[
K := \SO_2(\mathbb{R}) \times \SO_2(\mathbb{R}) \subseteq G(\mathbb{R}),
\]

with notation as before: \( n_u := (\begin{smallmatrix} 1 & 0 \\ u & 1 \end{smallmatrix}) \) and \( a_t := (\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}) \).

As before, let \( \alpha \in C_c^\infty(G(\mathbb{R})) \) and \( \beta \in C_c^\infty(L) \) be compactly supported smooth functions such that: \( \alpha \) is \( K \)-invariant, \( \int_{G(\mathbb{R})} \alpha = 1 \), \( \beta(\vec{F}_\pm) = 1 \), and
supp \( \beta, \beta^{-1}(\{1\}) \subseteq L^{(1)} \) are both unions of two small compact intervals respectively containing \( \bar{F}_\pm \). Let
\[
\varphi(v) := \sum_{g \in V_\ell = v} \alpha(g) \cdot \beta(v_L).
\]
Let \( V(Z)_{\text{nontriv.}} := \{(F_1, F_2) \in V(Z) : \forall [x, y] \in \mathbb{P}^1(\mathbb{Q}), \text{disc}_{X, Y}(x \cdot F_1(X, Y) - y \cdot F_2(X, Y)) \neq 0 \} \) and \( Y(Z)_{\text{nontriv.}} := Y(Z) \cap V(Z)_{\text{nontriv.}} \).

For \( M \in \mathbb{Z}^+ \) and \( v_0 \in V(Z/M) \), let \( Y_{v_0}(\text{mod } M)(Z) := \{ v \in Y(Z) : v \equiv v_0 \text{ (mod } M) \} \) and \( Y_{v_0}(\text{mod } M)(Z)_{\text{nontriv.}} := Y_{v_0}(\text{mod } M)(Z) \cap V(Z)_{\text{nontriv.}} \).

Write \( n_{(a_1, a_2)} := (n_{a_1}, n_{a_2}) \), and similarly \( a_{(t_1, t_2)} := (a_{t_1}, a_{t_2}) \).

Let, for \( S \subseteq V(Z) \),
\[
\#_{\varphi, \mu}|B(\bar{u}, \bar{t}, \lambda, X) \cap S| := \sum_{F \in S} \mu \left( \frac{L_0(\bar{F})}{X} \right) \cdot \varphi(\alpha_f^{-1} \cdot n_u^{-1} \cdot (\lambda \cdot \text{id, id}) \cdot (F_1, F_2)).
\]

We again see that it suffices to prove the following two lemmas. Save for an easy (since one saves a nonzero constant probability at each prime) use of the Selberg upper bound sieve to control \( \#_{\varphi, \mu}|B(\bar{u}, \bar{t}, \lambda, X) - V(Z)_{\text{nontriv.}}| \), the proof that these lemmas suffice, including reducing the evaluation of the resulting product of local densities by using the same trick to reduce to the same evaluation done (for the larger family \( F_1(3) \)) in the proof of part (f) of Theorem 1.1 of Bhargava-Ho's [3], tracks that of the previous section, so we omit it — though note that full details appear in [4].

**Lemma 5.1.** Let \( \lambda \in \mathbb{R}^+, u_i \in \left[ -\frac{1}{2}, \frac{1}{2} \right], \sqrt{\frac{3}{2}} \leq t_i \ll \lambda \). Then:
\[
\#_{\varphi, \mu}|B(\bar{u}, \bar{t}, \lambda, X) \cap Y(Z)_{\text{nontriv.}}| \ll_{\varphi} \lambda^{6+o(1)} \cdot \lambda^{5+o(1)} \cdot t_1 \cdot t_2^3.
\]

Note that the second term is admissible because Haar measure on \( G \) involves \( t_1^{-2} d^* t_1 t_2^{-2} d^* t_2 \) and \( t_1 \cdot t_2 \gg \lambda \) implies \( B(\bar{u}, \bar{t}, \lambda, X) \cap Y(Z)_{\text{nontriv.}} = \emptyset \).

**Lemma 5.2.** Let \( M \in \mathbb{Z}^+ \) and \( v_0 \in V(Z/M) \). Let
\[
\lambda \in \mathbb{R}^+, u_i \in \left[ -\frac{1}{2}, \frac{1}{2} \right], \sqrt{\frac{3}{2}} \leq t_i \ll \lambda.
\]

Then:
\[
\#_{\varphi, \mu}|B(\bar{u}, \bar{t}, \lambda, X) \cap Y_{v_0}(\text{mod } M)(Z)| = \sigma_\infty(\bar{u}, \bar{t}, \lambda, X) \cdot \prod_p \sigma_p(Y_{v_0}(\text{mod } M)(Z))
\]
\[
+ O_{\varphi}(\|\bar{t}\|_1^{8} \cdot \lambda^{4+o(1)}) + O_{\varphi}(\|\bar{t}\|_\infty^N \cdot \lambda^{O(1)-N}),
\]
where
\[
\sigma_\infty(\bar{u}, \bar{t}, \lambda, X) := \lim_{\varepsilon \to 0} \frac{\int_{v \in V(Z)} |I_2(v)| \leq \varepsilon \, dv \, \mu(\frac{L_0(\bar{F})}{X}) \cdot \varphi(\alpha_f^{-1} \cdot n_u^{-1} \cdot (\lambda \cdot \text{id, id})^{-1} \cdot v)}{2\varepsilon}
\]
and
\[
\sigma_p(Y_{v_0}(\text{mod } M)(Z)) := \lim_{n \to \infty} p^{-4n} \cdot \# \{ \bar{F} \in V(Z/p^n) : I_2(\bar{F}) \equiv 0 \text{ (mod } p^n) \}, v \equiv v_0 \text{ (mod } M) \}.
\]
5.2 The uniformity estimate.

In fact the proof of the first part of the uniformity estimate reduces to the one proven in Section 4.3 for the following reason. Recall that, given a pair of binary cubic forms \( \vec{F} =: (F_1, F_2) \) with each \( F_i \in \mathbb{Z}[X,Y] \), one produces a binary quartic form via

\[
G_{\vec{F}}(x,y) := \text{disc}_{X,Y}(x \cdot F_1(X,Y) - y \cdot F_2(X,Y)) \in \mathbb{Z}[x,y].
\]

Note that, as one can see by e.g. explicit calculation, \( I_2(\vec{F}) \mid I(G_{\vec{F}}) \).

Now in fact one has by definition that the pair \( \vec{F} = (F_1, F_2) \) is locally soluble at \( p \) if and only if \( -\frac{1}{3} \cdot G_{\vec{F}} \) is locally soluble at \( p \). Therefore to bound the number of \( \vec{F} \in B(\vec{u}, \vec{t}, \lambda, X) \cap Y(\mathbb{Z}) \) which are not locally soluble at a \( p \) with \( p > \Pi \), it suffices to observe that the statement that \( \vec{F} \) is not locally soluble at \( p \) implies that \( G_{\vec{F}}(\text{mod} \ p) \) lies on a codimension 3 subvariety of the space of binary quartics, so that (after checking the independence of the resulting three equations in the coefficients of \( \vec{F} \), which in fact imply that either \( F_2 \) is proportional to \( F_1 \) or else \( F_1 = 0 \)) \( \vec{F} \) (mod \( p \)) lies on a codimension 3 subvariety of the space of pairs of binary cubics, in which case we may again apply Theorem 1.1 of Browning-Heath-Brown’s [10] to conclude.

The proof of the second part of the uniformity estimate — that is, the bound on the number of \( \vec{v} \) with \( G(\mathbb{Z}_p) \cdot \vec{v} \not\in G(\mathbb{Q}_p) \cdot \vec{v} \cap V(\mathbb{Z}_p) \) for some \( p > \Pi \) — is more intricate and appears in full in [4].

5.3 Point counting.

The proof of Lemma 5.1 is very much the same as the proof of Lemma 4.1 except that nontriviality does not rule out points in a particular totally isotropic subspace, so that we must treat them separately.

Proof of Lemma 5.1  Recall that \( t_1^3 \cdot t_2^3 \ll \lambda \) (else \( B(\vec{u}, \vec{t}, \lambda, X) \cap Y(\mathbb{Z})^{\nontriv} = \emptyset \)). As mentioned, each coefficient of an \( (F_1, F_2) =: \vec{F} \in B(\vec{u}, \vec{t}, \lambda, X) \cap Y(\mathbb{Z})^{\nontriv} \) is \( \ll \lambda \). If there is an \( i \in \{0, 1, 2, 3\} \) such that \( F_i = F_{3-i} \neq 0 \) then, applying the divisor bound, we determine \( (F_i, F_{3-i}) \), up to \( \ll \lambda^{0(1)} \) choices, from \( ([F_{k,0}, F_{k,2}]_{0 \leq k \leq 3, k \neq i, i}) \) and there are \( \ll \lambda^5 \cdot t_1^3 \cdot t_2^3 \cdot \left( \frac{\lambda}{t_1 \cdot t_2} + 1 \right) = \lambda^5 + \lambda^5 \cdot t_1 \cdot t_2 \) choices for the latter.

It remains to treat those \( \vec{F} \) with \( F_i \cdot F_{3-i} = 0 \) for all \( 0 \leq i \leq 3 \). If there is a \( 1 \leq j \leq 2 \) and an \( i \in \{0, 3\} \) such that \( F_i^{(j)} = F_{i+(-1)^j} \) = 0 or else \( F_i^{(j)} = F_{i}^{(3-j)} = 0 \) then evidently \( \vec{F} \not\in V(\mathbb{Z})^{\nontriv} \). Hence there is a \( 1 \leq j \leq 2 \) such that \( F_0^{(j)} = F_3^{(j)} = F_3^{(3-j)} = F_2^{(3-j)} = 0 \). Thus the number of such \( \vec{F} \in B(\vec{u}, \vec{t}, \lambda, X) \) is \( \ll \left( \frac{\lambda}{t_1 \cdot t_2} + 1 \right) \cdot \frac{\lambda^3}{t_1} \cdot \frac{\lambda^3}{t_2} \cdot \frac{t_1^2}{t_1} \cdot \frac{t_2^2}{t_2} \cdot \frac{\lambda t_1}{t_1} \cdot \frac{\lambda t_2}{t_2} \cdot \left( \frac{\lambda t_1}{t_1} + 1 \right) \cdot \lambda t_1 \cdot t_2^3 = \lambda^4 + \lambda^4 \cdot t_1 \cdot t_2^3 \) as desired.

\( \square \)
The proof of Lemma 5.2 is also similar to the proof of Lemma 4.2. We comment on the two major differences — again, full details are given in [4].

**Proof of Lemma 5.2** The principal differences are the following. In applying the smoothed delta symbol method (i.e. Theorem 2 of Heath-Brown’s [11]), we take \( n = 8 \) and \( Q = \lambda \). In calculating the complete exponential sum, we note that, because we have an even number of variables and thus the mod-\( q \) complete exponential sums no longer vanish for \( q \) prime, we instead use the bound \( \ll q^{5+o(1)} \) for all \( q \). For the same reason we no longer reduce to a sum over only powerful \( q \), but rather we sum over all \( q \ll \lambda \) in the bounding of the error term.

Otherwise the analytic argument is mutatis mutandis the same. \( \square \)

### 6 Proof of Theorem 1.1 for \( k \equiv 5 \pmod{6} \).

Now we will deduce the sketch of the proof of Theorem 1.1 when \( k \equiv 5 \pmod{6} \). As noted we will deduce this case from the intermediate results used to deduce the case \( k \equiv 2 \pmod{6} \) as follows.

We begin with the following preliminary lemma.

**Lemma 6.1.** Let \( p \gg 1 \) be prime. Let \( A, B \in \mathbb{Z}_p \) be such that \(-4A^3 - 27B^2 \neq 0\).

Let \( E_{A,B} : y^2 = x^3 + Ax + B \) and let \( E_{A,B}/\mathbb{Z}_p \) be its scheme-theoretic closure in \( \mathbb{F}_p^2/\mathbb{Z}_p \). Let \( \Sigma \subseteq E_{A,B}(\mathbb{F}_p) \) be such that \(#|\Sigma| \ll 1\). Let \( P \in E_{A,B}(\mathbb{Q}_p) \). Then: there is an \((x, y) \in E_{A,B}(\mathbb{Z}_p)\) with \( x, y \in \mathbb{Z}_p^2 \) and \((x, y) \mod p \notin \Sigma\) such that \((x, y) \equiv P \mod 2 \cdot E_{A,B}(\mathbb{Q}_p)\) as elements of \( E_{A,B}(\mathbb{Q}_p) \).

**Proof.** Let \( S := \Sigma \cup \{ \infty \} \cup \{(x, y) \in E_{A,B}(\mathbb{F}_p) : x \cdot y = 0 \} \cup (E_{A,B} \mod p)^{\text{sing}} \subseteq E_{A,B}(\mathbb{F}_p) \). Note that \(#|S| \ll 1\). Next note that \( 0 \to E_{A,B}(\mathbb{F}_p)[2] \to E_{A,B}(\mathbb{F}_p) \to 2 \cdot E_{A,B}(\mathbb{F}_p) \to 0 \), and \(#|E_{A,B}(\mathbb{F}_p)[2]| \ll 1\), so that \(#|2 \cdot E_{A,B}(\mathbb{F}_p)| \gg p\) since \(#|E_{A,B}(\mathbb{F}_p)| = p + O(\sqrt{p}) \gg p\). Consequently for each \( \mathcal{P} \in E_{A,B}(\mathbb{F}_p) \) there is a \( \mathcal{Q} \in E_{A,B}(\mathbb{F}_p) \) such that \( \mathcal{P} + 2\mathcal{Q} \notin S \) (because \( p \gg 1\)).

So for each \( P \in E_{A,B}(\mathbb{Q}_p) \) there is a \( \mathcal{Q} \in E_{A,B}(\mathbb{F}_p) \) such that \( \mathcal{Q} \cdot (P \mod p) + 2\mathcal{Q} \notin S \), whence by Hensel’s \( \mathcal{Q} \in E_{A,B}(\mathbb{Q}_p) \) such that, writing \( P + Z \mathcal{Q} =: (x, y) \), one has \( x, y \in \mathbb{Z}_p^2 \), as desired. \( \square \)

Next let us note the following.

**Lemma 6.2.** Let \( p \gg 1 \). Let \( F \in \text{Sym}^4(2)(\mathbb{Z}) \) be such that \( I(F) = 0 \) and \( p \mid J(F) \).

Then: either \( p \mid F \) or else \( F \) has splitting type either \((1^4)\) or else \((1^31)\) modulo \( p \).

**Proof.** If \( p \mid F \) we are done. Otherwise by hypothesis \( F \) has a double root modulo \( p \), so its possible splitting types are \((1^4)\), \((1^31)\), \((1^21^2)\), \((2^2)\), and \((1^211)\).

Because \( I(X^2 \cdot (\alpha \cdot X^2 + \beta \cdot XY + \gamma \cdot Y^2)) = \gamma^2 \) all but the first two splitting types are ruled out and we are done. \( \square \)

This has the following consequence.

**Lemma 6.3.** Let \( p \gg 1 \) be a prime. Let \( F \in \text{Sym}^4(2)(\mathbb{Z}) \) with \( I(F) = 0 \) and \( p \mid J(F) \) be such that \( p \nmid F \) and such that \( F \) does not have splitting type \((1^4)\) modulo \( p \). Then: \( F \) has a root in \( \mathbb{F}^1(\mathbb{Q}_p) \).
Lemma 6.4. Let $p \gg 1$ be a prime. Let $F \in \text{Sym}^4(2)(\mathbb{Z})$ with $I(F) = 0$. Then: $d \mid J(F)$ if and only if $Z^2 = F(X, Y)$ has a $\mathbb{Q}_p$-point. Let $d \mid J(F)$. Then: $d \cdot Z^2 = F(X, Y)$ has a $\mathbb{Q}_p$-point, and so in particular if $a^3 \mid J(F)$ then it follows that there is a $g \in \text{PGL}_2(\mathbb{Q})$ such that $\frac{(g \cdot F)(X, Y)}{a} \in \mathbb{Z}_p[X, Y]$ with $(g \cdot F)(X, Y) := (\det g)^{-2} \cdot F((X, Y) \cdot g).

Proof. If $p \nmid J(F)$ we are done because $p \gg 1$. Otherwise because $Z^2 = F(X, Y)$ has a $\mathbb{Q}_p$-point $F$ is $\text{PGL}_2(\mathbb{Q}_p)$-equivalent to the Kummer image $F_{(x, y)}(X, Y) := X^4 - 6x \cdot X^2 Y^2 + 8y \cdot X^2 Y^3 - 3x^2 \cdot Y^4$ of a point $(x, y) \in E_{0 - 27, j(F)}(\mathbb{Q}_p)$ with (by Lemma 6.1) $x, y \in \mathbb{Z}_p^\times$, which is thus automatically not of mod-$p$ splitting type $(1^4)$, and so by Lemma 6.3 we conclude that $F$ has a root in $\mathbb{P}^1(\mathbb{Q}_p)$. Thus indeed $Z^2 = \frac{F(X, Y)}{a}$ has a $\mathbb{Q}_p$-point (indeed one with $Z = 0$), giving the first claim. As for the second claim, it follows in that case that there is a $g \in \text{PGL}_2(\mathbb{Q}_p)$ such that $\frac{(g \cdot F)(X, Y)}{a}$ is the Kummer image of a point $(x, y) \in E_{0 - 27, j(F)}(\mathbb{Q}_p)$ with (by Lemma 6.1) $x, y \in \mathbb{Z}_p^\times$, and so in particular that $\frac{(g \cdot F)(X, Y)}{a} \in \mathbb{Z}_p[X, Y]$ as desired (to conclude approximate $g$ sufficiently closely by an element of $\text{PGL}_2(\mathbb{Q})$).

Now let us repeat the notation of Section 5: let $V := 2 \otimes \text{Sym}^3(2)$, let $G$ be the image in $\text{GL}(V)$ of $\{(g, h) \in \text{GL}_2 \times \text{GL}_2 : \det g \cdot (\det h)^3 = 1\}$, etc.

Note that the (first) invariant binary quartic associated by Bhargava-Ho in their [2] to a $\vec{v} \in V$ — via $V \hookrightarrow 2^\otimes 4$, i.e. via regarding a pair of binary cubic forms as a triply symmetric hypercube in the usual way (i.e. $(\sum_{i=0}^3 (\vec{a}_i^3)^i \cdot x^i)^2 \mapsto (\vec{a}_{j+k+\ell-3}(1 \leq i, j, k, \ell \leq 2) \in \text{GL}(V)(X, Y) := \frac{1}{2^7} \cdot \text{disc}_{(x, y)}(X \cdot v_1(x, y) - Y \cdot v_2(x, y))$).

By construction $F = \left(\sum_{i=0}^3 (\vec{a}_i^3)^i \cdot x^i\right)^2 \in \mathbb{Z}\left[\vec{a}_0, \ldots, \vec{a}_3, X, Y\right]$.

Lemma 6.5. There is an absolute constant $\kappa \in \mathbb{Z} - \{0\}$ such that the following holds. Let $0 \neq B \in \mathbb{Z}$ be sixth-power-free. Then: there is a bijection between $\text{Sel}_2(E_{0, B^6}/\mathbb{Q}) - \{0\}$ and

$$G(\mathbb{Q}) \\setminus \left\{ \vec{v} \in \frac{1}{\kappa} \cdot V(\mathbb{Z})^{\text{nontriv}}, \cdot B \cdot 2^2 = F_{\vec{v}}(X, Y) \text{ everywhere locally soluble} \right\}.$$ 

Proof. Factor $B =: B_1 \cdot B_2$, with $B_1$ squarefree — thus $E_{0, B^6} \simeq E_{0, B_1^6, B_2^6}$ over $\mathbb{Q}$.

Of course (by e.g. Theorem 3.5 of Bhargava-Shankar’s [9]) $\text{Sel}_2(E_{0, B_1^6, B_2^6}/\mathbb{Q}) - \{0\}$ is in bijection with $\text{PGL}_2(\mathbb{Q}) \setminus \{ F \in \text{Sym}^4(2)(\mathbb{Z})^{\text{nontriv}}, I(F) = 0, J(F) = -2^3 \cdot B^2 - 3^3 \cdot B^2, F(X, Y) \text{ everywhere locally soluble} \}$. Given such an $F \in \text{Sym}^4(2)(\mathbb{Z})^{\text{nontriv}}$, we first claim that there is an absolute constant $\kappa \in \mathbb{Z} - \{0\}$ (thus $|\kappa| \ll 1$), a $\vec{v} \in \frac{1}{\kappa} \cdot V(\mathbb{Z})^{\text{nontriv}}$, and a $g \in \text{PGL}_2(\mathbb{Q})$ such that $B_1 \cdot F_{\vec{v}}(X, Y) = (g \cdot F)(X, Y)$. 

28
To see this, note that the curve $B_* \cdot Z^2 = F(X,Y)$ is a genus one curve with a degree-two line bundle (via the pullback of $O(1)$ via $[X,Y,Z] \mapsto X/Y$) whose Jacobian (namely $y^2 = x^3 + B^2$) has a marked rational 3-torsion point (namely $(0,B)$). Thus by Theorem 6.16 of Bhargava-Ho’s [Z] it follows that there is a $\vec{v} \in V(\mathbb{Q})$ such that $\frac{F(X\gamma(Y)}{B_*}$ is $\text{PGL}_2(\mathbb{Q})$-equivalent to $F_\vec{v}(X,Y)$.

Note that this already shows that, for each $p$, there is a $\vec{w} \in G(\mathbb{Q}_p) \cdot \vec{v}$ with $\vec{w} \in \mathbb{Q}^{O(1)} \cdot V(\mathbb{Z}_p)$, since by following the procedure outlined at the bottom of page 57 of Bhargava-Ho’s [Z] (and symmetrizing suitably) one produces a universal formula for a $\vec{v} \in V \bigg( \mathbb{Z} \left[ \text{coeff.s}(\vec{F}), \sqrt{-27 J(\vec{F})}, (6 \cdot J(\vec{F}))^{-1} \right] \bigg)$ such that $\tilde{F}_{\vec{v}}$ is $\text{PGL}_2 \left( \mathbb{Q} \left[ \text{coeff.s}(\vec{F}), \sqrt{-27 J(\vec{F})} \right] \right)$-equivalent to a given $\vec{F} \in \text{Sym}^4(2)$ with $I(\vec{F}) = 0$ and $J(\vec{F}) \neq 0$.

But furthermore Lemma 6.4 implies that $\vec{v}$ is locally soluble at all $p \gg 1$, so that, just as in Bhargava-Ho’s [Z], by strong approximation for $G/\mathbb{Q}$ it follows that there is a $\vec{v} \in V(\mathbb{Q})$ such that, for all $p$, $\vec{v} \in \mathbb{Q}^{O(1)} \cdot V(\mathbb{Z}_p)$, and, for all $p \gg 1$, $\vec{v}$ is furthermore $G(\mathbb{Z}_p)$-equivalent to the Kummer image of an $(x,y) \in E_{\vec{v}}(\mathbb{Q}_p)$ with (by Lemma 6.1) $x,y \in \mathbb{Z}_p^\times$ so that in particular $\vec{v} \in V(\mathbb{Z}_p)$ for these $p$.

This proves the desired claim that for each locally soluble $F \in \text{Sym}^4(2)(\mathbb{Z})_{\text{nontriv}}$ with $I(\vec{F}) = 0$ and $J(\vec{F}) = -2^6 \cdot 3^1 \cdot B^2 \cdot B^2$ there is a $\vec{v} \in \frac{1}{\kappa} \cdot V(\mathbb{Z})_{\text{nontriv}}$ such that $F_\vec{v}(X,Y)$ and $\frac{F(X\gamma(Y)}{B_*}$ are $\text{PGL}_2(\mathbb{Q})$-equivalent. The map $\text{PGL}_2(\mathbb{Q}) \cdot \vec{F} \mapsto \text{Kummer image}$ thus defined (i.e. by appeal to Theorem 6.16 of Bhargava-Ho’s [Z]) is certainly injective because we recover $\text{PGL}_2(\mathbb{Q}) \cdot \vec{F}$ from $G(\mathbb{Q}) \cdot \vec{v}$ by the defining property of the map, but it is also of course surjective — given $\vec{v}$ just form $B_* \cdot F_\vec{v}(X,Y)$ — so we are done.

Now we may prove Theorem 1.1 for $k \equiv 5 \pmod 6$.

Proof of Theorem 1.1 for $k \equiv 5 \pmod 6$. Repeat the proof of Theorem 1.1 in the case $k \equiv 2 \pmod 6$ — including the appeal to the uniformity estimate which is proved in full in [4] — mutatis mutandis, with the exception that the local conditions (besides those modulo $\mathcal{O}(1)$) imposed be the everywhere local solubility of $2\sigma_k(\vec{v}) \cdot Z^2 = F_\vec{v}(X,Y)$ — rather than that of $Z^2 = F_\vec{v}(X,Y)$ — throughout.

(Now that the fact that these provide the same local densities ultimately reduces to Lemma 3.20 of Bhargava-Shankar’s [9].)

7 Proof of Theorem 1.1 for $k \equiv 4 \pmod 6$.

Now to the sketch of the proof of Theorem 1.1 when $k \equiv 4 \pmod 6$. The relevant inputs are as follows.

---

1Here to define the Kummer image we use the formula specified in the paragraph after (14) on page 14 of Bhargava-Ho’s [8] in this case (so crucially $v = w = 0$ in their notation).
Lemma 7.1. There is an absolute constant $\kappa \in \mathbb{Z} - \{0\}$ such that the following holds. Let $0 \neq B \in \mathbb{Z}$ be sixth-power-free. Then: there is a bijection between $\operatorname{Sel}_2(E_{0,B^4}/\mathbb{Q}) - \{0\}$ and

$$\operatorname{PGL}_2(\mathbb{Q}) \backslash \left\{ F \in \frac{1}{\kappa} \cdot V(\mathbb{Z})_{\text{nontriv}} : I(F)=0, J(F)=-2^6 \cdot 3^3 \cdot B, \quad B \cdot Z^2=F(X,Y) \text{ everywhere locally soluble} \right\}.$$ 

Proof. Factor $B = B_s \cdot B_s^2$ with $B_s$ squarefree — thus $E_{0,B^4} \simeq E_{0,B_s^4}$ over $\mathbb{Q}$. By e.g. Theorem 3.5 of Bhargava-Shankar’s [9] $\operatorname{Sel}_2(E_{0,B^4}/\mathbb{Q}) - \{0\}$ is in bijection with

$$\operatorname{PGL}_2(\mathbb{Q}) \backslash \left\{ F \in \operatorname{Sym}^4(2)(\mathbb{Z})_{\text{nontriv}} : I(F)=0, J(F)=-2^6 \cdot 3^3 \cdot B, \quad Z^2=F(X,Y) \text{ everywhere locally soluble} \right\}.$$ 

Given such an $F \in \operatorname{Sym}^4(2)(\mathbb{Z})_{\text{nontriv}}$, we claim that there is an absolute constant $\kappa \in \mathbb{Z} - \{0\}$ (thus $|\kappa| \ll 1$) and a $g \in \operatorname{PGL}_2(\mathbb{Q})$ such that $\frac{(g \cdot F)(X,Y)}{B_s} \in \frac{1}{\kappa} \cdot \mathbb{Z}[X,Y]$, where $(g \cdot F)(X,Y) := (\det g)^{-2} \cdot F((X,Y) \cdot g)$. By strong approximation for $\operatorname{PGL}_2(\mathbb{Q})$ this is a local question — for $p \gg 1$ we apply Lemma 6.4, while for the remaining $p \ll 1$ of course $\frac{F(X,Y)}{B_s} \in \frac{1}{p} \cdot \mathbb{Z}[X,Y]$ already, so we are done with the claim.

Of course at the level of $\operatorname{PGL}_2(\mathbb{Q})$-orbits this map amounts to the map $\operatorname{PGL}_2(\mathbb{Q}) \cdot F(X,Y) \mapsto \operatorname{PGL}_2(\mathbb{Q}) \cdot \frac{F(X,Y)}{B_s}$ so it is evidently a bijection, so that we are done.

Lemma 7.2. Let $p \gg 1$. Let $F \in \operatorname{Sym}^4(2)(\mathbb{Z})$ with $I(F) = 0$ be such that $-27 J(F) \cdot Z^2 = F(X,Y)$ is not soluble over $\mathbb{Q}_p$. Then: either $p \mid F$ or else $F$ has splitting type $(1^4)$ modulo $p$.

Consequently the image modulo $p$ of the set of $F \in \operatorname{Sym}^4(2)(\mathbb{Z})$ with $I(F) = 0$ such that either $\operatorname{PGL}_2(\mathbb{Z}_p) \cdot F \neq \operatorname{PGL}_2(\mathbb{Q}_p) \cdot F \cap V(\mathbb{Z}_p)$ or else $-27 J(F) \cdot Z^2 = F(X,Y)$ is not soluble over $\mathbb{Q}_p$ lies in a codimension 2 subvariety of $\{ F : I(F) = 0 \}$.

Proof. Apply Lemma 6.3 for the first claim, and then repeat the argument given in Section 4.3 for the second.

Thus we may prove Theorem 1.1 for $k \equiv 4 \pmod{6}$ in the same way as for $k \equiv 5 \pmod{6}$.

Proof of Theorem 1.1 for $k \equiv 4 \pmod{6}$. Repeat the proof of Theorem 1.1 — including the uniformity estimate (via Lemma 7.2) — in the case $k \equiv 1 \pmod{6}$ mutatis mutandis, with the exception that the local conditions (besides those modulo $\kappa^{O(1)}$) imposed be the everywhere local solubility of $-27 J(F) \cdot Z^2 = F(X,Y)$ — rather than that of $Z^2 = F(X,Y)$ — throughout. (Note that the fact that these provide the same local densities ultimately reduces to Lemma 3.20 of Bhargava-Shankar’s [9].)

Note that we could have also used Corollary 6.20 of Bhargava-Ho’s [7] and the discussion thereafter — replacing $-27 J(F) \cdot Z^2 = F(X,Y)$ by $Z^2 =
$H(F)(X,Y)$ with $H(F)(X_1, X_2) := \text{disc}_{(x,y)} \left( \det \left( x \cdot \frac{\partial F(X_1, X_2)}{\partial x_1 \partial x_j}, y \cdot \frac{\partial F(X_1, X_2)}{\partial x_2 \partial x_j} \right) \right)_{j,k=1}$

throughout. Indeed the argument is again the same save for the production of integral representatives, i.e. save for producing an $F \in \text{Sym}^2(2)(\mathbb{Z})$ for which $Z^2 = H(F)(X,Y)$ (or alternatively $\text{PGL}_2(\mathbb{Q}) \cdot H(F)$) represents a given element of $\text{Sel}_2(E_0, B^*_0)/\mathbb{Q} - \{0\}$, and we produce such integral representatives as follows.

**Lemma 7.3.** There is an absolute constant $\kappa \in \mathbb{Z} - \{0\}$ such that the following holds. Let $0 \neq B \in \mathbb{Z}$ be cubefree. Then: there is a bijection between $\text{Sel}_2(E_0, B^*_0)/\mathbb{Q} - \{0\}$ and

$$\text{PGL}_2(\mathbb{Q}) \setminus \left\{ F \in \frac{1}{\kappa} \cdot \text{Sym}_4(2)(\mathbb{Z}) \text{nontriv.} : [(F) = 0, \text{J}(F) = -2^6 \cdot 3^2 \cdot B, Z^2 = H(F)(X,Y) \text{ everywhere locally soluble}] \right\}.$$

**Proof.** Factor $B =: B_* \cdot B^3$ with $B_*$ squarefree — thus $E_{0, B^*_0} \simeq E_{0, B^*_0 \cdot B^3}$ over $\mathbb{Q}$. As usual (i.e. just as in the proofs of Lemmas 6.5 and 7.1) strong approximation for $\text{PGL}_2(\mathbb{Q})$ reduces the problem to a local question. Now from the discussion after (14) on page 14 of Bhargava-Ho’s [8] we see that given an $(x, y) \in E_{0, B^*_0 \cdot B^4}(\mathbb{Q}_p)$ with $x, y \in \mathbb{Z}_p^\times$ (which by Lemma 6.1 is the only case we must consider) there is a $\vec{v} \in (2 \otimes \text{Sym}_3(2))(\mathbb{Z})$ such that $v_1 = ((\in \mathbb{Z}_p^\times), 0, (\in \mathbb{Z}_p^\times))$ and $v_2 = (0, (\in \mathbb{Z}_p^\times), (\in \mathbb{Z}_p^\times), 0)$, and such that $F_{\vec{v}}(X,Y)$ is $\text{PGL}_2(\mathbb{Q}_p)$-equivalent to the usual Kummer image $F_{(x,y)}(X,Y) := X^4 - 6x \cdot X^2 Y^2 + 8y \cdot X Y^3 - 3x^2 \cdot Y^4$ (thus $a_6(\vec{v}) = 2 \cdot B^3 \cdot B^4$). Then by the discussion after (71) on page 65 of Bhargava-Ho’s [7] we find an explicit $g \in M_3(\mathbb{Z})$ with $\det g = a_6(\vec{v})$ such that $g \cdot \vec{v}$ is the image (under the evident map, i.e. under $\text{Sym}_4(2) \rightarrow 2 \otimes \text{Sym}_3(2)$) of a $G \in \text{Sym}_4(2)(\mathbb{Z})$, namely $G(X,Y) = x^2 \cdot X^4 + 4\tilde{\gamma} \cdot X^3 Y + 6x \cdot X^2 Y^2 + \frac{4x^2}{y} \cdot X Y^3 + Y^4$, where $\tilde{\gamma} := y - \frac{a_6(\vec{v})}{2} \cdot x^3$, which therefore has $I(G) = 0$ and $J(G) = 2^4 \cdot 3^1 \cdot a_6(\vec{v})^2 = 2^4 \cdot 3^1 \cdot B^3 \cdot B^4$. We are thus done for $p \ll 1$ or else for $p \nmid B_*$ by considering $\frac{G(X,Y)}{B_*}$. Otherwise $p \gg 1$ and $p \mid B_*$, and then writing $\tilde{G}(X,Y) := -G \left( Y, X - \frac{y^2}{x^2} \cdot y \right) =: X^4 + e \cdot X^2 Y^2 + \cdots + e \cdot Y^4$ by inspection — namely $\tilde{\gamma}^2 \equiv x^3 \left( \text{mod } a_6(\vec{v}) \right)$, so $\left( \frac{\tilde{\gamma}^2}{x^2} \cdot X + Y \right)^4 \equiv -G(X,Y) \left( \text{mod } a_6(\vec{v}) \right)$, and moreover $I(G) = 0$ — we have that $a_6(\vec{v}) \mid c, d$ and $a_6(\vec{v})^2 \mid e$, so that $\frac{G(p \cdot X, Y)}{p^2} \in \mathbb{Z}_p[X,Y].$ Thus $\tilde{G}(X,Y) := \frac{\text{disc}(G(X,Y))}{B_*} \in \mathbb{Z}_p[X,Y]$ is such that $I(\tilde{G}) = 0, J(\tilde{G}) = -2^6 \cdot 3^2 \cdot B$, and such that $H(\tilde{G})(X,Y)$ is $\text{PGL}_2(\mathbb{Q}_p)$-equivalent to $F_{(x,y)}(X,Y)$, so that $\tilde{G}$ is the desired integral representative.

We conclude as we did in the proof of Lemma 7.1.
References.

[1] Levent Hasan Ali Alpöge, Points on Curves, ProQuest LLC, Ann Arbor, MI, 2020, Thesis (Ph.D.)–Princeton University. MR 4209988

[2] Levent Alpöge, Manjul Bhargava, and Ari Shnidman, A positive proportion of cubic fields are not monogenic yet have no local obstruction to being so, (2021), arXiv:2011.01186.

[3] --------, A positive proportion of quartic fields are not monogenic yet have no local obstruction to being so, (2021), arXiv:2107.05514.

[4] --------, Integers expressible as the sum of two rational cubes, 2022, arXiv:2210.10730.

[5] Manjul Bhargava, Higher composition laws, ProQuest LLC, Ann Arbor, MI, 2001, Thesis (Ph.D.)–Princeton University. MR 2702004

[6] --------, The density of discriminants of quartic rings and fields, Ann. of Math. (2) 162 (2005), no. 2, 1031–1063. MR 2183288

[7] Manjul Bhargava and Wei Ho, Coregular spaces and genus one curves, Camb. J. Math. 4 (2016), no. 1, 1–119. MR 3472915

[8] Manjul Bhargava and Wei Ho, On average sizes of Selmer groups and ranks in families of elliptic curves having marked points, (2020).

[9] Manjul Bhargava and Arul Shankar, Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves, Ann. of Math. (2) 181 (2015), no. 1, 191–242. MR 3272925

[10] Tim Browning and Roger Heath-Brown, The geometric sieve for quadrics, (2020), arXiv:2003.09593.

[11] D. R. Heath-Brown, A new form of the circle method, and its application to quadratic forms, J. Reine Angew. Math. 481 (1996), 149–206. MR 1421949

[12] Samuel Ruth, A bound on the average rank of j-invariant zero elliptic curves, ProQuest LLC, Ann Arbor, MI, 2013, Thesis (Ph.D.)–Princeton University. MR 3211431

[13] Stanley Yao Xiao, Binary quartic forms with vanishing j-invariant, (2019), arXiv:1712.09091.