Quadrics in arithmetic statistics.

Levent Alpöge

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 computing the average size of 2-Selmer groups in the families $y^2 = x^3 + B$ and $y^2 = x^3 + B^2$.

In the course of the argument we introduce a smoothed form of Bhargava's aforementioned method, as well as a trick with which we formally deduce that the above averages are 3 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[1] application of the counting technique invented by Bhargava in his Princeton Ph.D. thesis has reduced such an orbit counting problem to one of counting lattice points in an expanding region and then dealt[2] 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 Princeton Ph.D. thesis [11], 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 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,

[1] We know of exactly one work that faces point counting on a nontrivial subvariety "directly", namely Samuel Ruth's Princeton Ph.D. thesis [11]. However, there are certainly others, e.g. Yao Xiao's [12], 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.

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 $A \equiv 0 \pmod{N}$ imposed, is 3 (independent of N!), by work of Bhargava-Shankar.

Now let us give precise statements of the example applications.

1.1 Main theorems.

The first example we work out is a generalization of the main theorem of Ruth's Princeton Ph.D. thesis [11] (see Theorem 1.1.2 of his [11], which amounts to a proof of the upper bound when $\mathcal{B} = \mathbb{Z} - \{0\}$) that was important for [2] (and thus [3]).

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

$$\operatorname*{Avg}_{B \in \mathcal{B}: |B| \leq X} \#|\mathrm{Sel}_2(E_{0,B}/\mathbb{Q})| \leq 3 + O_{\mathcal{B}}(o_{X \to \infty}(1)),$$

with equality if e.g.[3] *$\mathcal{B} \subseteq \mathbb{Z} - \{0\}$ 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 subsets $\varnothing \neq \mathcal{B} \subseteq \mathbb{Z} - \{0\}$ which are "large" in a sense completely analogous

[3] We write "with equality if e.g." to emphasize that this is not an if and only if statement.

to that of Bhargava-Shankar's [8], but we will spare ourselves the notational effort.

Tracking the condition "$A = 0$" through the arguments of Bhargava-Shankar [8], 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 $F(X,Y) =: a \cdot X^4 + b \cdot X^3Y + c \cdot X^2Y^2 + d \cdot XY^3 + e \cdot Y^4$,

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

The second example concerns the yet thinner family $y^2 = x^3 + B^2$.

Theorem 1.2. *Let $\mathcal{B} \subseteq \mathbb{Z} - \{0\}$ be a set of positive upper density defined by congruence conditions. Then:*

$$\operatorname*{Avg}_{B \in \mathcal{B}: |B| \leq X} \#|\mathrm{Sel}_2(E_{0,B^2}/\mathbb{Q})| \leq 3 + O_{\mathcal{B}}(o_{X \to \infty}(1)),$$

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

Again, we could replace "defined by finitely many congruence conditions" with a more general notion of largeness analogous to that of Bhargava-Shankar's [8], but we will spare ourselves the notational effort.

As for the counting problem, as in Bhargava-Ho [7] 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 $F_i(X,Y) =: a_i \cdot X^3 + \cdots + d_i \cdot Y^3$.

1.2 Main technique.

Having stated the main theorems, let us discuss the method of proof. In order to be specific, let us put ourselves in the situation of the first theorem: we would like to count, up to $\mathrm{PGL}_2(\mathbb{Q})$-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 $V := \mathrm{Sym}^4(2)$ for the space of binary quartic forms. It is now standard that, using a method introduced in Bhargava's Princeton Ph.D. thesis [4] and first carried out in this context by Bhargava-Shankar [8] (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:

$$\int_{1 \ll \lambda \ll X^{\frac{1}{24}}} d^\times\lambda \int_{|u| \ll 1} du \int_{1 \ll t \ll \lambda} t^{-2} d^\times t \, \#|\{F \in \lambda \cdot n_u \cdot a_t \cdot G_0 \cdot L \cap V(\mathbb{Z})^{\mathrm{irred.}} : I(F) = 0\}|$$
$$\sim \mathrm{const.} \cdot X^{\frac{1}{2}}.$$

Here $n_u := \left(\begin{smallmatrix} 1 & 0 \\ u & 1 \end{smallmatrix}\right)$, $a_t := \left(\begin{smallmatrix} t^{-1} & 0 \\ 0 & t \end{smallmatrix}\right)$, and $d^\times z := \frac{dz}{z}$, so that $d^\times\lambda \, du \, t^{-2} d^\times t \, dk$ is the Haar measure on $\mathrm{GL}_2^+(\mathbb{R})$ (the "+" denoting positive determinant) induced by Haar measures on Λ, N, A, K under the Iwasawa decomposition $\mathrm{GL}_2^+(\mathbb{R}) = \Lambda \cdot N \cdot A \cdot K$, where Λ are the scalars, N are the lower unipotents, A is the diagonal torus, and $K := \mathrm{SO}_2(\mathbb{R})$ 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 $\asymp 1$, 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:

$$\begin{aligned} |a| &\ll \frac{\lambda^4}{t^4}, \\ |b| &\ll \frac{\lambda^4}{t^2}, \\ |c| &\ll \lambda^4, \\ |d| &\ll t^2 \cdot \lambda^4, \\ |e| &\ll t^4 \cdot \lambda^4. \end{aligned}$$

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 λ, 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 $1 \ll t \ll_\varepsilon \lambda^\varepsilon$, 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

$$\#|\{F \in \lambda \cdot n_u \cdot a_t \cdot B \cdot L \cap V(\mathbb{Z})^{\text{irred.}} : I(F) = 0\}| = \text{const.} \cdot \lambda^{12} + O_\varepsilon(\lambda^{12-\delta+O(\varepsilon)})$$

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. $\lambda^\varepsilon \ll_\varepsilon t \ll \lambda$. Here Ruth uses a difficult argument that again involves the smoothed delta symbol method and is quite specific to the situation — so much so that it is unclear if his argument generalizes from $\mathcal{B} = \mathbb{Z} - \{0\}$ to sets defined by finitely many congruence conditions, let alone to other arithmetic statistics problems.

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

$$\#|\{F \in \lambda \cdot n_u \cdot a_t \cdot B \cdot L \cap V(\mathbb{Z})^{\text{irred.}} : I(F) = 0\}| \ll \lambda^{12+o(1)}$$

by the divisor bound: $12ae - 3bd + c^2 = 0$ implies that $a, e \,|\, 3bd - c^2$. 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 $t \gg_\varepsilon \lambda^\varepsilon$ combines with the $t^{-2}d^\times t$ 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 $\ll_\varepsilon \lambda^{12-\Omega(\varepsilon)}$ 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.

This article is based on Chapter 3 of the author's Ph.D. thesis at Princeton University [1]. I would like to thank both my advisor Manjul Bhargava and Peter Sarnak for their patience and encouragement. I would also like to thank Jacob Tsimerman and Nina Zubrilina 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.

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 Princeton Ph.D. thesis [4], though the trick of averaging over fundamental domains was first introduced in the published version (see Section 2.2 of Bhargava's [5]). 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 $\mathbb{1}_{G_0}$, 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 [5], and note that we are taking (in his notation) Φ 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 (it will be clear how to modify the construc-

[4]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).

tion for Theorem 1.2, and indeed in any application of Bhargava's counting method). That is, $V := \mathrm{Sym}^4(2), G := \mathrm{GL}_2$, $L := L^{(0)} \coprod L^{(1)} \coprod L^{(2+)} \coprod L^{(2-)}$ with

$$L^{(0)} := \left\{ X^3Y - \frac{1}{3} \cdot XY^3 + \frac{J}{27} \cdot Y^4 : J \in (-2, 2) \right\},$$
$$L^{(1)} := \left\{ X^3Y - \frac{I}{3} \cdot XY^3 \pm \frac{2}{27} \cdot Y^4 : I \in [-1, 1) \right\} \cup \left\{ X^3Y + \frac{1}{3} \cdot XY^3 + \frac{J}{27} \cdot Y^4 : J \in (-2, 2) \right\},$$
$$L^{(2\pm)} := \pm \left\{ \frac{1}{16} X^4 - \sqrt{\frac{2-J}{27}} \cdot X^3Y + \frac{1}{2} \cdot X^2Y^2 + Y^4 : J \in (-2, 2) \right\},$$

and

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

Gauss's classical fundamental domain for $G(\mathbb{Z}) \curvearrowright G(\mathbb{R})$. Note that L is a fundamental domain for $G(\mathbb{R}) \curvearrowright V(\mathbb{R})^{\Delta \neq 0}$. Observe that the points with $I = 0$ and $\Delta \neq 0$ are all in the $G(\mathbb{R})$-orbit of the two forms $F_\pm(X, Y) := X^3Y \pm \frac{2}{27} \cdot Y^4$, 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 $\alpha \in C_c^\infty(G(\mathbb{R}))$ and $\beta \in C_c^\infty(L)$ be compactly supported smooth functions such that: α is $\mathrm{SO}_2(\mathbb{R})$-invariant, $\int_{G(\mathbb{R})} \alpha = 1$, $\beta(F_\pm) = 1$, and $\operatorname{supp} \beta, \beta^{-1}(\{1\}) \subseteq L^{(1)}$ 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 α and β we get a slightly more convenient way[5] to smooth out the various integrals in Bhargava's counting technique — instead of integrating the normalized indicator function $\frac{1}{\int_{G_0} dg} \cdot \mathbb{1}_{G_0}$ 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 $G(\mathbb{Z}) \curvearrowright G(\mathbb{R})$ so that all counts are independent of g, we instead integrate $\alpha(g)$ over $g \in G$ and then make the same observation.

[5]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 β to ensure smoothness of φ — note that $L^{(0)} \coprod L^{(1)}$ is a rectangle missing two corners, and at the other two corners a similar definition of φ with $\beta = 1$ identically would fail to be smooth. One can get around this, of course, but this choice simplifies notation.

Specifically, we observe that, since

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

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 $G(\mathbb{Z}) \curvearrowright G(\mathbb{R})$ — the first main observation of Section 2.2 of Bhargava's [5]), it follows that:

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

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

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 [8] (and implicitly in Bhargava's [5]).

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:

$$\begin{aligned}
&\int_{G(\mathbb{R})} dh \, \alpha(h) \cdot \#|\{g \in \mathcal{F} : gh \cdot v_L = v\}| \\
&= \sum_{\gamma \in G(\mathbb{R}) : \gamma \cdot v_L = v} \int_{G(\mathbb{R})} dh \, \alpha(h) \cdot \beta(v_L) \cdot \#|\{g \in \mathcal{F} : gh = \gamma\}| \\
&= \sum_{\gamma \in G(\mathbb{R}) : \gamma \cdot v_L = v} \int_{\mathcal{F}^{-1} \cdot \gamma} dh \, \alpha(h) \cdot \beta(v_L) \\
&= \sum_{\gamma \in G(\mathbb{R}) : \gamma \cdot v_L = v} \int_{\mathcal{F}^{-1}} dh \, \alpha(h \cdot \gamma) \cdot \beta(v_L) \\
&= \int_{\mathcal{F}} dh \sum_{\gamma \in G(\mathbb{R}) : \gamma \cdot v_L = v} \alpha(h^{-1} \cdot \gamma) \cdot \beta(v_L) \\
&= \int_{\mathcal{F}} dh \sum_{g \in G(\mathbb{R}) : g \cdot v_L = h^{-1} \cdot v} \alpha(g) \cdot \beta(v_L) \\
&= \int_{\mathcal{F}} dh \, \varphi(h^{-1} \cdot v),
\end{aligned}$$

by definition.

Therefore we have found that:

$$N(Y(\mathbb{Z})^{\mathrm{nontriv.}}, X) = \sum_{v \in V(\mathbb{Z})^{\mathrm{nontriv.}} : I(v)=0, 0 \neq |J(v)| \leq X} \int_{\mathcal{F}} dh \, \varphi(h^{-1} \cdot v)$$
$$= \int_{\mathcal{F}} dh \sum_{v \in V(\mathbb{Z})^{\mathrm{nontriv.}} : I(v)=0, 0 \neq |J(v)| \leq X} \varphi(h^{-1} \cdot v).$$

4 Proof of Theorem 1.1.

Now let us turn to the proof of Theorem 1.1.

4.1 Reduction to point counting.

For the reader's convenience we use Ruth's notation in what follows. Let $V := \mathrm{Sym}^4(2)$. Let $G := \mathrm{GL}_2$. 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 $L := L^{(0)} \coprod L^{(1)} \coprod L^{(2+)} \coprod L^{(2-)}$ with

$$L^{(0)} := \left\{ X^3Y - \frac{1}{3} \cdot XY^3 + \frac{J}{27} \cdot Y^4 : J \in (-2, 2) \right\},$$
$$L^{(1)} := \left\{ X^3Y - \frac{I}{3} \cdot XY^3 \pm \frac{2}{27} \cdot Y^4 : I \in [-1, 1) \right\} \cup \left\{ X^3Y + \frac{1}{3} \cdot XY^3 + \frac{J}{27} \cdot Y^4 : J \in (-2, 2) \right\},$$
$$L^{(2\pm)} := \pm \left\{ \frac{1}{16} X^4 - \sqrt{\frac{2-J}{27}} \cdot X^3Y + \frac{1}{2} \cdot X^2Y^2 + Y^4 : J \in (-2, 2) \right\},$$

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

Let

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

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

Let $\alpha \in C_c^\infty(G(\mathbb{R}))$ and $\beta \in C_c^\infty(L)$ be compactly supported smooth functions such that: α is $\mathrm{SO}_2(\mathbb{R})$-invariant, $\int_{G(\mathbb{R})} \alpha = 1$, $\beta(F_\pm) = 1$, and $\operatorname{supp}\beta, \beta^{-1}(\{1\}) \subseteq L^{(1)}$ are both unions of two small compact intervals respectively containing $F_\pm$. Let $G_0 := \operatorname{supp}\beta$. Let

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

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

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

(Here we deviate from Ruth, and indeed Bhargava [4] and Bhargava-Shankar [8], in using the superscript $\mathrm{nontriv.}$ instead of $\mathrm{irred.}$, since in our view it is misleading to call these irreducible.) That is, $V(\mathbb{Z})^{\mathrm{nontriv.}}$ 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 $Y(\mathbb{Z})^{\mathrm{nontriv.}} := Y(\mathbb{Z}) \cap V(\mathbb{Z})^{\mathrm{nontriv.}}$. We note that our $Y(\mathbb{Z})$ and $Y(\mathbb{Z})^{\mathrm{nontriv.}}$ play the role of Ruth's Y and $Y^{\mathrm{irr.}}$, 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 $Y_{F_0 \pmod M}(\mathbb{Z})^{\mathrm{nontriv.}} := Y_{F_0 \pmod M}(\mathbb{Z}) \cap V(\mathbb{Z})^{\mathrm{nontriv.}}$.

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

$$\#_\varphi |B(u,t,\lambda,X) \cap S| := \sum_{F \in S : |J(F)| \leq X} \varphi(a_t^{-1} \cdot n_u^{-1} \cdot (\lambda \cdot \mathrm{id})^{-1} \cdot F),$$

where we have disambiguated the action of λ (which should really be the action of $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right)$) by writing $\lambda \cdot \mathrm{id}$ for clarity. We note that this is different from Ruth's (and indeed Bhargava's) notation precisely because we have smoothed using α and β rather than simply $\mathbb{1}_{G_0}$ — though the reader may well imagine that $\varphi \sim \mathbb{1}_{G_0 \cdot L}$, in which case $\#_\varphi |B(u,t,\lambda,X) \cap S| \sim \#|\lambda \cdot n_u \cdot a_t \cdot G_0 \cdot L \cap S|$.

Just as in Section 2.2 of Ruth's [11], we find that the problem reduces to controlling $\#_\varphi|B(u,t,\lambda,X)\cap Y_{F_0 \pmod M}(\mathbb{Z})^{\mathrm{nontriv.}}|$. We do this with the following two lemmas.[6]

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

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

Then:

$$\#_\varphi|B(u,t,\lambda,X)\cap Y(\mathbb{Z})^{\mathrm{nontriv.}}| \ll_\varphi \lambda^{12+o(1)}.$$

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

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

Then:

$$\#_\varphi|B(u,t,\lambda,X)\cap Y_{F_0 \pmod M}(\mathbb{Z})| = \sigma_\infty(u,t,\lambda,X)\cdot\prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})) + O_{\varphi,M}(t^4\cdot\lambda^{8+o(1)}) + O_{\varphi,M,N}(t^N\cdot\lambda^{O(1)-N}),$$

where

$$\sigma_\infty(u,t,\lambda,X) := \lim_{\varepsilon\to 0}\frac{\int_{v\in V(\mathbb{R}):|I(v)|\leq\varepsilon,|J(v)|\leq X} dv\,\varphi(a_t^{-1}\cdot n_u^{-1}\cdot(\lambda\cdot\mathrm{id})^{-1}\cdot v)}{2\varepsilon}$$

and

$$\sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})) := \lim_{k\to\infty} p^{-4k}\cdot\#|\{F\in V(\mathbb{Z}/p^k) : I(F)\equiv 0 \pmod{p^k}, F\equiv F_0 \pmod M\}|.$$

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 [8], 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, $f\mapsto\frac{1}{m(f)}$) with M highly divisible. The following weighted version of Lemma 4.2 of course follows immediately from Lemma 4.2.

[6]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 λ and X — on page 12 of [11] he states that the condition $|J(v)|<X$ is superfluous once $\lambda\ll X^{\frac{1}{24}}$, thus one can drop it — this is false, since his definition of his $\mathcal{F}'$ depends on a parameter C' and were this to be the case the Selmer average would also grow with C', rather than be 3).

Lemma 4.3. *Let $M \in \mathbb{Z}^+$ and $\phi : V(\mathbb{Z}/M) \to \mathbb{R}$. Let*

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

Then:

$$\begin{aligned}
&\sum_{F_0 \in V(\mathbb{Z}/M)} \phi(F_0) \cdot \#_\varphi |B(u,t,\lambda,X) \cap Y_{F_0 \pmod M}(\mathbb{Z})| \\
&= \sigma_\infty(u,t,\lambda,X) \cdot \prod_{p \nmid M} \sigma_p(Y(\mathbb{Z})) \cdot \lim_{k\to\infty} M^{-4k} \cdot \sum_{F \in V(\mathbb{Z}/M^k): I(F) \equiv 0 \pmod{M^k}} \phi(F \pmod M) \\
&\quad + O_{\varphi,M}(||\phi||_1 \cdot t^4 \cdot \lambda^{8+o(1)}) + O_{\varphi,M,N}(||\phi||_1 \cdot t^N \cdot \lambda^{O(1)-N}).
\end{aligned}$$

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 $I(F) \equiv 0 \pmod{p^k}$, 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 from Lemmas 4.1 and 4.2.

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

$$\begin{aligned}
&\sum_{B \in \mathcal{B}: |B| \leq X} \#|\mathrm{Sel}_2(E_{0,B}/\mathbb{Q})| \\
&\leq \sum_{B \in \mathcal{B}_{\leq T}: |B| \leq X} \#|\mathrm{Sel}_2(E_{0,B}/\mathbb{Q})| \\
&\leq (3 + O_{\mathcal{B},T}(o_{X\to\infty}(1))) \cdot \#|\{n \in \mathcal{B}_{\leq T} : |n| \leq X\}| \\
&= (3 + O_{\mathcal{B},T}(o_{X\to\infty}(1))) \cdot \left(\frac{\#|\{n \in \mathcal{B}_{\leq T} : |n| \leq X\}|}{\#|\{n \in \mathcal{B} : |n| \leq X\}|}\right) \cdot \#|\{n \in \mathcal{B} : |n| \leq X\}| \\
&= (3 + O_{\mathcal{B}}(o_{T\to\infty}(1)) + O_{\mathcal{B},T}(o_{X\to\infty}(1))) \cdot \#|\{n \in \mathcal{B} : |n| \leq X\}|.
\end{aligned}$$

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

$$\operatorname*{Avg}_{n \in \mathcal{B}: |n| \leq X} \#|\mathrm{Sel}_2(E_{0,B}/\mathbb{Q})| \leq 3 + O_{\mathcal{B}}(o_{X\to\infty}(1)),$$

as desired.

Thus without loss of generality $\mathcal{B}$ is defined by finitely many congruence conditions, i.e. it is of the form $\mathcal{B} = \{n \in \mathbb{Z} - \{0\} : n \equiv a \pmod m\}$ 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 [11] 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 [8]: $n(F)$ is the number of $\mathrm{PGL}_2(\mathbb{Z})$-orbits in the (intersection of $V(\mathbb{Z})$ and the) $\mathrm{PGL}_2(\mathbb{Q})$-orbit of $F \in V(\mathbb{Z})^{\Delta \neq 0}$, and

$$\begin{aligned} m(F) &:= \sum_{F' \in \mathrm{PGL}_2(\mathbb{Z}) \backslash (V(\mathbb{Z}) \cap \mathrm{PGL}_2(\mathbb{Q}) \cdot F)} \frac{\#|\operatorname{Aut}_{\mathrm{PGL}_2(\mathbb{Q})}(F')|}{\#|\operatorname{Aut}_{\mathrm{PGL}_2(\mathbb{Z})}(F')|} \\ &= \prod_p \sum_{F' \in \mathrm{PGL}_2(\mathbb{Z}_p) \backslash (V(\mathbb{Z}_p) \cap \mathrm{PGL}_2(\mathbb{Q}_p) \cdot F)} \frac{\#|\operatorname{Aut}_{\mathrm{PGL}_2(\mathbb{Q}_p)}(F')|}{\#|\operatorname{Aut}_{\mathrm{PGL}_2(\mathbb{Z}_p)}(F')|} \\ &=: \prod_p m_p(F). \end{aligned}$$

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 ϕ_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 ϕ_n and take $n \to \infty$.

As we have seen in e.g. Section 3,

$$\begin{aligned} &N(Y_{F_0 \pmod M}(\mathbb{Z})^{\mathrm{nontriv.}}, X) \\ &= \int_{\mathcal{F}} dh \sum_{v \in V(\mathbb{Z})^{\mathrm{nontriv.}} : I(v) = 0, F \equiv F_0 \pmod M} \varphi(h^{-1} \cdot v) \\ &= \int_{1 \ll \lambda \ll X^{\frac{1}{24}}} d^\times \lambda \int_{u \in \nu(t)} du \int_{\sqrt{\frac{\sqrt{3}}{2}} \leq t \ll \lambda} t^{-2} d^\times t \, \#_\varphi |B(u,t,\lambda,X) \cap Y_{F_0 \pmod M}(\mathbb{Z})^{\mathrm{nontriv.}}| \\ &= \int_{1 \ll \lambda \ll X^{\frac{1}{24}}} d^\times \lambda \int_{u \in \nu(t)} du \int_{\sqrt{\frac{\sqrt{3}}{2}} \leq t \ll \lambda^\delta} t^{-2} d^\times t \, \#_\varphi |B(u,t,\lambda,X) \cap Y_{F_0 \pmod M}(\mathbb{Z})^{\mathrm{nontriv.}}| \\ &\quad + \int_{1 \ll \lambda \ll X^{\frac{1}{24}}} d^\times \lambda \int_{|u| \leq \frac{1}{2}} du \int_{\lambda^\delta \ll t \ll \lambda} t^{-2} d^\times t \, \#_\varphi |B(u,t,\lambda,X) \cap Y_{F_0 \pmod M}(\mathbb{Z})^{\mathrm{nontriv.}}|, \\ &= \int_{1 \ll \lambda \ll X^{\frac{1}{24}}} d^\times \lambda \int_{u \in \nu(t)} du \int_{\sqrt{\frac{\sqrt{3}}{2}} \leq t \ll \lambda^\delta} t^{-2} d^\times t \begin{pmatrix} \sigma_\infty(u,t,\lambda,X) \cdot \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})) \\ + O_{\varphi,M}(t^4 \cdot \lambda^{8+o(1)}) \end{pmatrix} \\ &\quad + \int_{1 \ll \lambda \ll X^{\frac{1}{24}}} d^\times \lambda \int_{|u| \leq \frac{1}{2}} du \int_{\lambda^\delta \ll t \ll \lambda} t^{-2} d^\times t \; O_\varphi(\lambda^{12+o(1)}) \end{aligned}$$

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:

$$
\begin{aligned}
&N(Y_{F_0 \pmod M}(\mathbb{Z})^{\mathrm{nontriv.}}, X)\\
&= \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})) \cdot \int_{1 \ll \lambda \ll X^{\frac{1}{24}}} d^\times \lambda \int_{u \in \nu(t)} du \int_{\sqrt{\frac{\sqrt{3}}{2}} \le t \ll \lambda^\delta} t^{-2} d^\times t \, \sigma_\infty(u,t,\lambda,X)\\
&\quad + O_{\varphi,M}(X^{1-\Omega(\delta)+o(1)})\\
&= \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})) \cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_{\mathcal{F}} d^\times \lambda \, du \, t^{-2} d^\times t \int_{v \in V(\mathbb{R}) : |I(v)| \le \varepsilon, |J(v)| \le X} dv \, \varphi((\lambda \cdot a_t \cdot n_u)^{-1} \cdot v)\\
&\quad + O_{\varphi,M}(X^{1-\Omega(\delta)+o(1)})\\
&= \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})) \cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_{v \in V(\mathbb{R}) : |I(v)| \le \varepsilon, |J(v)| \le X} dv \int_{h \in \mathcal{F}} dh \, \varphi(h^{-1} \cdot v)\\
&\quad + O_{\varphi,M}(X^{1-\Omega(\delta)+o(1)}),
\end{aligned}
$$

by definition of σ_∞ and then by letting $h := (\lambda \cdot \mathrm{id}) \cdot a_t \cdot n_u$ and recalling that φ is $\mathrm{SO}_2(\mathbb{R})$-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 $\beta(v_L) = 1$),

$$\int_{h \in \mathcal{F}} dh \, \varphi(h^{-1} \cdot v) = \int_{G(\mathbb{R})} dh \, \alpha(h) \cdot \#|\{g \in \mathcal{F} : gh \cdot v_L = v\}|.$$

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 $|I(v)| \le \varepsilon$ and $|J(v)| \gg 1$ when ε is sufficiently small. Inserting this into the above, we find that the main term is:

$$
\begin{aligned}
&(1 + O_M(X^{-1})) \cdot N(Y_{F_0 \pmod M}(\mathbb{Z})^{\mathrm{nontriv.}}, X)\\
&= \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z}))\\
&\qquad \cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_{v \in V(\mathbb{R}) : |I(v)| \le \varepsilon, 1 \ll |J(v)| \le X} dv \int_{G(\mathbb{R})} dh \, \alpha(h) \cdot \#|\{g \in \mathcal{F} : gh \cdot v_L = v\}|\\
&= \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z}))\\
&\qquad \cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_{G(\mathbb{R})} dh \, \alpha(h) \int_{v \in V(\mathbb{R}) : |I(v)| \le \varepsilon, 1 \ll |J(v)| \le X} dv \, \#|\{g \in \mathcal{F} : gh \cdot v_L = v\}|\\
&= \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})) \cdot \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \cdot \int_{G(\mathbb{R})} dh \, \alpha(h) \int_{v \in \mathcal{F} \cdot h \cdot L : |I(v)| \le \varepsilon, 1 \ll |J(v)| \le X} dv,
\end{aligned}
$$

by switching integrals and then noting that $\mathcal{F} \cdot h$ is a fundamental domain for the action of $G(\mathbb{Z}) \curvearrowright G(\mathbb{R})$ (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 $|I(v)| \le \varepsilon$ and $|J(v)| \ll 1$ in order to use the equality $\int_{h \in \mathcal{F}} dh \, \varphi(h^{-1} \cdot v) = \int_{G(\mathbb{R})} dh \, \alpha(h) \cdot \#|\{g \in \mathcal{F} : gh \cdot v_L = v\}|$.

Now the inner integral is independent of the choice of fundamental domain, whence:

$$\begin{aligned}
&(1+O_M(X^{-1}))\cdot N(Y_{F_0 \pmod M}(\mathbb{Z})^{\text{nontriv.}}, X)\\
&= \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z}))\cdot \lim_{\varepsilon\to 0}(2\varepsilon)^{-1}\cdot \int_{v\in\mathcal{F}\cdot L:|I(v)|\le\varepsilon, 1\ll|J(v)|\le X} dv\\
&= 2X\cdot \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z}))\cdot \lim_{\varepsilon\to 0}(4\varepsilon\cdot X)^{-1}\cdot \int_{v\in\mathcal{F}\cdot L:|I(v)|\le\varepsilon, 1\ll|J(v)|\le X} dv\\
&= \frac{2X}{m}\cdot\left(m\cdot\prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z}))\right)\cdot \lim_{\varepsilon\to 0}(4\varepsilon\cdot X)^{-1}\cdot \int_{v\in\mathcal{F}\cdot L:|I(v)|\le\varepsilon, 1\ll|J(v)|\le X} dv,
\end{aligned}$$

Now let us evaluate the remaining integral. By Proposition 2.8 of Bhargava-Shankar's [8] (we could avoid using this to get from the leftmost to the rightmost terms in the equality, of course), we have that

$$\begin{aligned}
(4\varepsilon\cdot X)^{-1}\cdot \int_{v\in\mathcal{F}\cdot L:|I(v)|\le\varepsilon, |J(v)|\le X} dv &= \frac{2\zeta(2)}{27}\cdot(1+O(\varepsilon\cdot X^{-1}))\\
&= (1+O(\varepsilon\cdot X^{-1}))\cdot \frac{\int_{\coprod_i \mathcal{R}^{(i)}(X)} dv}{\int_{\coprod_i R^{(i)}(X)} dI dJ},
\end{aligned}$$

using the notation of Section 2.4 of Bhargava-Shankar's [8]. 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:

$$N(Y_{F_0 \pmod M}(\mathbb{Z})^{\text{nontriv.}}, X) = (1+O_M(X^{-1}))\cdot\frac{2X}{m}\cdot\left(m\cdot\prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z}))\right)\cdot\frac{\int_{\coprod_i \mathcal{R}^{(i)}(X^2/4)} dv}{\int_{\coprod_i R^{(i)}(X^2/4)} dI dJ}.$$

Note that

$$\#|\{J\in\mathcal{B}: 0\neq|J|\le X\}| = \frac{2X}{m}\cdot\left(1+O\left(\frac{m}{X}\right)\right).$$

Just as in the passage from Lemma 4.2 to Lemma 4.3, we find that:

$$\begin{aligned}
&\sum_{F_0\in V(\mathbb{Z}/M): J(F_0)\equiv a \pmod m} \phi(F_0)\cdot N(Y_{F_0 \pmod M}(\mathbb{Z})^{\text{nontriv.}}, X)\\
&= (1+O(X^{-1}))\cdot\frac{2X}{m}\cdot\frac{\int_{\coprod_i \mathcal{R}^{(i)}(X^2/4)} dv}{\int_{\coprod_i R^{(i)}(X^2/4)} dI dJ}\\
&\quad\cdot\left(m\cdot\lim_{T\to\infty} n_T^{-4}\cdot\sum_{F\in V(\mathbb{Z}/n_T): I(F)\equiv 0 \pmod{n_T}, J(F)\equiv a \pmod m}\phi(F \pmod M)\right),
\end{aligned}$$

where $n_T := \prod_{p\le T} p^T$, say, and without loss of generality $T \ge M$ so that $M|n_T$.

Write $\mathbb{1}_{\text{sol.}}^{(p)}$ for the indicator function of the $\mathbb{Q}_p$-soluble binary quartic forms $F \in V(\mathbb{Z}_p)$ (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 $\mathbb{1}_{\text{loc. sol.}} := \prod_p \mathbb{1}_{\text{sol.}}^{(p)}$. Write

$$\phi_*(F) := \frac{\mathbb{1}_{\text{loc. sol.}}}{m(F)} = \prod_p \frac{\mathbb{1}_{\text{sol.}}^{(p)}}{m_p(F)} =: \prod_p \phi_*^{(p)}(F)$$

on $V(\mathbb{Z})^{\Delta\neq 0}$. Thus when $1 =: \phi_0^{(p)}(F) \geq \phi_1^{(p)}(F) \geq \cdots \geq \phi_n^{(p)}(F) \geq \cdots \geq \phi_*^{(p)}(F)$ for all $F \in V(\mathbb{Z}_p)^{\Delta\neq 0}$ with $\phi_n^{(p)} : V(\mathbb{Z}_p)^{\Delta\neq 0} \to [0,1]$ factoring through $V(\mathbb{Z}/p^n)$ (and not $V(\mathbb{Z}/p^{n-1})$) and such that $\phi_n^{(p)}(F) \to \phi_*^{(p)}(F)$ as $n \to \infty$, we find that[7], writing

$$\phi_n := \prod_{p \leq n} \phi_n^{(p)},$$

(and here is where we use Lemma 2.2.3 of Ruth's [11] to replace $m(F)$ by $n(F)$ for all but $\ll X^{\frac{5}{6}}$ forms):

$$\begin{aligned}
&\operatorname*{Avg}_{B\in\mathcal{B}:|B|\leq X} \#|\mathrm{Sel}_2(E_{0,B}/\mathbb{Q}) - \{0\}| \\
&\leq (1 + O_T(o_{X\to\infty}(1))) \cdot (1 + o_{T\to\infty}(1)) \\
&\quad \cdot \frac{\int_{\coprod_i \mathcal{R}^{(i)}(X^2/4)}\, dv}{\int_{\coprod_i R^{(i)}(X^2/4)}\, dI dJ} \cdot \left(m \cdot n_T^{-4} \cdot \sum_{F\in V(\mathbb{Z}/n_T): I(F)\equiv 0 \,(\mathrm{mod}\ n_T), J(F)\equiv a \,(\mathrm{mod}\ m)} \phi_T(F) \right).
\end{aligned}$$

Note that we have "thickened" by changing the constraint $I(F) = 0$ to $I(F) \equiv 0 \,(\mathrm{mod}\ n_T)$. Accordingly, we write

$$\mathrm{Inv}^{(T,\mathcal{B})} := \{(I,J) \in \mathbb{Z}^2 : I \equiv 0 \,(\mathrm{mod}\ n_T), J \equiv a \,(\mathrm{mod}\ m), 4I^3 - J^2 \neq 0\},$$

and $\mathrm{Inv}_p^{(T,\mathcal{B})} \subseteq \mathbb{Z}_p^2$ for its closure in $\mathbb{Z}_p^2$.

The trick is now to notice that, for all $k \in \mathbb{Z}^+$,

$$\begin{aligned}
&\sum_{F\in V(\mathbb{Z}/n_T): I(F)\equiv 0 \,(\mathrm{mod}\ n_T), J(F)\equiv a \,(\mathrm{mod}\ m)} \phi_T(F) \\
&= n_T^{-5\cdot(k-1)} \cdot \sum_{F\in V(\mathbb{Z}/n_T^k): I(F)\equiv 0 \,(\mathrm{mod}\ n_T), J(F)\equiv a \,(\mathrm{mod}\ m)} \phi_T(F),
\end{aligned}$$

[7] 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.

so that:

$$m \cdot n_T^{-4} \cdot \sum_{F \in V(\mathbb{Z}/n_T): I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F)$$
$$= \frac{n_T^{-5k} \cdot \sum_{F \in V(\mathbb{Z}/n_T^k): I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F)}{n_T^{-2k} \cdot \#|\{(I,J) \in V(\mathbb{Z}/n_T^k) : I \equiv 0 \pmod{n_T}, J \equiv a \pmod{m}\}|}.$$

Now we note that

$$\lim_{k \to \infty} n_T^{-5k} \cdot \sum_{F \in V(\mathbb{Z}/n_T^k): I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F)$$
$$= \int_{F \in V\left(\prod_{p \leq T} \mathbb{Z}_p\right): (I(F), J(F)) \in \prod_{p \leq T} \mathrm{Inv}_p^{(T,\mathcal{B})}} dF \, \phi_T(F),$$

and so, by dominated convergence and Fubini (note that we are implicitly using Proposition 3.18 of Bhargava-Shankar's [8], and indeed arguing as in their proof of Proposition 2.21 of their [8]), it follows that[8]

$$n_T^{-5k} \cdot \sum_{F \in V(\mathbb{Z}/n_T): I(F) \equiv 0 \pmod{n_T}, J(F) \equiv a \pmod{m}} \phi_T(F)$$
$$= (1 + o_{T \to \infty}(1)) \cdot (1 + O_T(o_{k \to \infty}(1)))$$
$$\cdot \prod_p \int_{v \in V(\mathbb{Z}_p): (I(v), J(v)) \in \mathrm{Inv}_p^{(T,\mathcal{B})}} dv \, \phi_*^{(p)}(v).$$

Similarly, we of course have:

$$n_T^{-2k} \cdot \#|\{(I,J) \in V(\mathbb{Z}/n_T^k) : I \equiv 0 \pmod{n_T}, J \equiv a \pmod{m}\}|$$
$$= \prod_{p \leq T} \int_{(I,J) \in \mathrm{Inv}_p^{(T,\mathcal{B})}} dI dJ$$
$$= (1 + o_{T \to \infty}(1)) \cdot \prod_p \int_{(I,J) \in \mathrm{Inv}_p^{(T,\mathcal{B})}} dI dJ.$$

Combining all these, we find that:

$$\operatorname*{Avg}_{B \in \mathcal{B}: |B| \leq X} \#|\mathrm{Sel}_2(E_{0,B}/\mathbb{Q}) - \{0\}|$$
$$\leq (1 + O_T(o_{X \to \infty}(1))) \cdot (1 + o_{T \to \infty}(1)) \cdot (1 + O_T(o_{k \to \infty}(1)))$$
$$\cdot \frac{\int_{\coprod_i \mathcal{R}^{(i)}(X^2/4)} dv}{\int_{\coprod_i R^{(i)}(X^2/4)} dI dJ} \cdot \frac{\prod_p \int_{v \in V(\mathbb{Z}_p): (I(v), J(v)) \in \mathrm{Inv}_p^{(T,\mathcal{B})}} dv \, \phi_*^{(p)}(v)}{\prod_p \int_{(I,J) \in \mathrm{Inv}_p^{(T,\mathcal{B})}} dI dJ}.$$

But the calculations in Section 3.6 of Bhargava-Shankar's [8] amount to the statement that

$$\frac{\int_{\coprod_i \mathcal{R}^{(i)}(X^2/4)} dv}{\int_{\coprod_i R^{(i)}(X^2/4)} dI dJ} \cdot \frac{\prod_p \int_{v \in V(\mathbb{Z}_p): (I(v), J(v)) \in \mathrm{Inv}_p^{(T,\mathcal{B})}} dv \, \phi_*^{(p)}(v)}{\prod_p \int_{(I,J) \in \mathrm{Inv}_p^{(T,\mathcal{B})}} dI dJ} = 2.$$

[8] See the previous footnote (about factoring out $(1 + o_{k \to \infty}(1))$ and $(1 + o_{T \to \infty}(1))$).

Taking $k \to \infty$, then $X \to \infty$, and finally $T \to \infty$, we deduce Theorem 1.1. $\square$

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 $0 =: \phi_0^{(p)}(F) \leq \phi_1^{(p)}(F) \leq \cdots \leq \phi_n^{(p)}(F) \leq \cdots \leq \phi_*^{(p)}(F)$ for all $F \in V(\mathbb{Z}_p)^{\Delta \neq 0}$, with $\phi_n : V(\mathbb{Z}_p)^{\Delta \neq 0} \to [0,1]$ factoring through $V(\mathbb{Z}/p^n)$ (and not $V(\mathbb{Z}/p^{n-1})$) and such that $\phi_n^{(p)}(F) \to \phi_*^{(p)}(F)$ 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 $F \in V(\mathbb{Z})^{\mathrm{nontriv.}}$ with $I(F) = 0$ representing 2-Selmer classes of $E_{0,J(F)}/\mathbb{Q}$ — i.e. such that $Z^2 = F(X,Y)$ 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. So, just as in Bhargava-Shankar's [8], we need only prove that the number of binary quartics F in e.g. $B(u,t,\lambda,X) \cap Y(\mathbb{Z})$ which are *not* locally soluble at some prime p with $p > \Pi$ 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 [8], if $F \in V(\mathbb{Z})^{\mathrm{nontriv.}}$ is a binary quartic that is not locally soluble at p, then $F \pmod p$ 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.[9] 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.

So it follows that we may bound the number of $F \in B(u,t,\lambda,X) \cap Y(\mathbb{Z})$ which are not locally soluble at some p with $p > \Pi$ by the number of $F \in B(u,t,\lambda,X) \cap Y(\mathbb{Z})$ for which the reduction $F \pmod p \in Z(\mathbb{F}_p)$ for some $p > \Pi$, and then the desired bound follows from invoking Theorem 1.1 of Browning-Heath-Brown's [9].

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

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

$$I(X^2 \cdot (X - n \cdot Y)^2) = I(X^4 - 2n \cdot X^3 Y + n^2 \cdot X^2 Y^2) = n^4.$$

from b, c, d"), which subsumes the entirety of Ruth's Section 3.5 (pages $29 - 37$ of [11]).

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

$$\begin{aligned} 0 \neq |a| &\ll \frac{\lambda^4}{t^4}, \\ |b| &\ll \frac{\lambda^4}{t^2}, \\ |c| &\ll \lambda^4, \\ |d| &\ll t^2 \cdot \lambda^4, \\ 0 \neq |e| &\ll t^4 \cdot \lambda^4. \end{aligned}$$

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. □

As for Lemma 4.2, we first note that it is essentially identical to Ruth's Proposition $3.4.1$ (modulo the small 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 [10]. Let $w_0 \in C_c^\infty(\mathbb{R})$ via

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

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

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 [10] with his $n = 5$ and his $Q = \lambda^4$, we find that:

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

$$= (\lambda^{-8} + O_N(\lambda^{-N})) \cdot \sum_{q \geq 1} q^{-5} \sum_{\vec{c} \in V(\mathbb{Z})^*} \left(\sum_{u \in (\mathbb{Z}/q)^\times} \sum_{G \in V(\mathbb{Z}/q)} e_q(u \cdot I(G) + G \cdot \vec{c}) \right)$$
$$\cdot \int_{F \in V(\mathbb{R})} dF \, \varphi(a_t^{-1} \cdot n_u^{-1} \cdot (\lambda \cdot \mathrm{id})^{-1} \cdot F) \cdot h\left(\frac{q}{\lambda^4}, \frac{I(F)}{\lambda^8} \right) \cdot e_q(-F \cdot \vec{c}),$$

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

For us the error term will consist of those terms where $\vec{c} \neq \vec{0}$, and the terms with $\vec{c} = \vec{0}$ 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 $F \mapsto n_u \cdot a_t \cdot (\lambda \cdot \mathrm{id}) \cdot F$ (note that $(\lambda \cdot \mathrm{id}) \cdot F = \lambda^4 \cdot F$ since F is homogeneous of degree 4 — here $(\lambda \cdot \mathrm{id}) \cdot F$ on the left-hand side indicates the action of $\lambda \cdot \mathrm{id} \in G$ on $F \in V$ via $G \curvearrowright V$, and the $\cdot$ on the right-hand side denotes multiplication), we find that:

$$\int_{F \in V(\mathbb{R})} dF \, \varphi(a_t^{-1} \cdot n_u^{-1} \cdot (\lambda \cdot \mathrm{id})^{-1} \cdot F) \cdot h\left(\frac{q}{\lambda^4}, \frac{I(F)}{\lambda^8} \right) \cdot e_q(-F \cdot \vec{c})$$
$$= \lambda^{20} \cdot \int_{F \in V(\mathbb{R})} dF \, \varphi(F) \cdot h\left(\frac{q}{\lambda^4}, I(F) \right) \cdot e_q(-\lambda^4 \cdot F \cdot ((n_u \cdot a_t)^\dagger \cdot \vec{c})).$$

Therefore we see that the error term is:

$$(\lambda^{12} + O_N(\lambda^{-N})) \cdot \sum_{q \geq 1} q^{-5} \sum_{\vec{0} \neq \vec{c} \in V(\mathbb{Z})^*} \left(\sum_{u \in (\mathbb{Z}/q)^\times} \sum_{G \in V(\mathbb{Z}/q)} e_q(u \cdot I(G) + G \cdot \vec{c}) \right)$$
$$\cdot \int_{F \in V(\mathbb{R})} dF \, \varphi(F) \cdot h\left(\frac{q}{\lambda^4}, I(F) \right) \cdot e_q(-\lambda^4 \cdot F \cdot ((n_u \cdot a_t)^\dagger \cdot \vec{c})).$$

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

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

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

— 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

$$\sum_{u \in (\mathbb{Z}/q)^\times} \sum_{G_0, \ldots, G_4 \in \mathbb{Z}/q} e_q(u \cdot (12 G_0 G_4 - 3 G_1 G_3 + G_2^2) + (c_0 G_0 + c_1 G_1 + c_2 G_2 + c_3 G_3 + c_4 G_4)),$$

is very easy to calculate (see e.g. page 49 of [1]). We conclude that

$$\sum_{u \in (\mathbb{Z}/q)^\times} \sum_{G \in V(\mathbb{Z}/q)} e_q(u \cdot I(G) + G \cdot \vec{c}) \ll \begin{cases} 0 & \exists p > 3 : v_p(q) = 1 \\ q^{\frac{7}{2}+o(1)} & \forall p | q \text{ s.t. } p > 3, v_p(q) \geq 2, \end{cases}$$

and indeed one can sharpen the bound significantly.

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

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

(in which case it is $\ll_\varphi \frac{\lambda^4}{q}$). 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 [10]), or else $\vec{c} \neq \vec{0}$, in which case we must have that $q \gg \frac{\lambda^{4-\varepsilon}}{t^4}$ 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:

$$\ll \lambda^{12+o(1)} \cdot \sum_{\frac{\lambda^{4-\varepsilon}}{t^4} \ll q \ll \lambda^4,\, q \text{ powerful}} q^{-\frac{3}{2}} \sum_{\vec{0} \neq \vec{c} \in V(\mathbb{Z})^* : ||\vec{c}||_\infty \ll \frac{q}{\lambda^{4-\varepsilon}}} \int_{F \in V(\mathbb{R})} dF\, \varphi(F) \cdot h\left(\frac{q}{\lambda^4}, I(F)\right) \cdot e_q(-\lambda^4 \cdot F \cdot ((n_u \cdot a_t)^\dagger \cdot \vec{c})).$$

Bounding the integrals trivially (i.e. by $\ll_\varphi \frac{q}{\lambda^4}$) and noting that the number of $0 \neq \vec{c} \in \mathbb{Z}^5$ such that $||(n_u \cdot a_t)^\dagger \cdot \vec{c}||_\infty \ll \frac{q}{\lambda^{4-\varepsilon}}$ is

$$\ll \frac{t^4 \cdot q}{\lambda^{4-\varepsilon}} \cdot \left(1 + \frac{t^2 \cdot q}{\lambda^{4-\varepsilon}}\right) \cdot \left(1 + \frac{q}{\lambda^{4-\varepsilon}}\right) \cdot \left(1 + \frac{q}{t^2 \cdot \lambda^{4-\varepsilon}}\right) \cdot \left(1 + \frac{q}{t^4 \cdot \lambda^{4-\varepsilon}}\right)$$

when $\frac{\lambda^{4-\varepsilon}}{t^4} \ll q \ll \lambda^4$, we get that the error term is:

$$\ll t^4 \cdot \lambda^{4+\varepsilon+o(1)} \cdot \sum_{\frac{\lambda^{4-\varepsilon}}{t^4} \ll q \ll \lambda^4,\, q \text{ powerful}} q^{\frac{1}{2}} \cdot \left(1 + \frac{t^2 \cdot q}{\lambda^{4-\varepsilon}}\right) \cdot \left(1 + \frac{q}{\lambda^{4-\varepsilon}}\right) \cdot \left(1 + \frac{t^{-2} \cdot q}{\lambda^{4-\varepsilon}}\right) \cdot \left(1 + \frac{t^{-4} \cdot q}{\lambda^{4-\varepsilon}}\right)$$

$$\ll t^6 \cdot \lambda^{8+5\varepsilon},$$

as desired.

Thus we have bounded the error term suitably. The required analysis of the main term is already done in Heath-Brown's [10] (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 $F \equiv F_0 \pmod{M}$ 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 $F_0 \in V(\mathbb{Z})$). 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

$$\sigma_p^{(F_0 \pmod M)}(Y(\mathbb{Z})) := \lim_{k\to\infty} p^{-4k}\cdot \#|\{\tilde{F} \in V(\mathbb{Z}/p^k) : I(F_0+M\cdot\tilde{F}) \equiv 0 \pmod{p^k}\}|.$$

We therefore are reduced to the claim that

$$M^{-5}\cdot\prod_p \sigma_p^{(F_0 \pmod M)}(Y(\mathbb{Z})) = \prod_p \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})).$$

Of course for $p \nmid M$ we have that

$$\sigma_p^{(F_0 \pmod M)}(Y(\mathbb{Z})) = \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})) = \sigma_p(Y(\mathbb{Z})),$$

via the evident change of variables $\tilde{F} \mapsto M^{-1}\cdot(\tilde{F}-F_0)$.

However, it is also evident that

$$M^{-5}\cdot\prod_{p|M} \sigma_p^{(F_0 \pmod M)}(Y(\mathbb{Z})) = \prod_{p|M} \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})),$$

for the following reason. For $k \in \mathbb{Z}^+$ with $k \gg 1$ we have that:

$$\begin{aligned}
&M^{-5}\cdot\prod_{p|M} \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z}))\\
&= M^{-5}\cdot\prod_{p|M}\begin{pmatrix} p^{-4k\cdot v_p(M)}\cdot\#|\{\tilde{F} \in V(\mathbb{Z}/p^{k\cdot v_p(M)}) : I(F_0+M\cdot\tilde{F}) \equiv 0 \pmod{p^{k\cdot v_p(M)}}\}| \\ +O(p^{-k\cdot v_p(M)})\end{pmatrix}\\
&= \Big(1+O(e^{-\Omega_M(k)})\Big)\cdot M^{-4k-5}\cdot\#|\{\tilde{F} \in V(\mathbb{Z}/M^k) : I(F_0+M\cdot\tilde{F}) \equiv 0 \pmod{M^k}\}|,
\end{aligned}$$

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:

$$\begin{aligned}
&M^{-4k-5}\cdot\#|\{\tilde{F} \in V(\mathbb{Z}/M^k) : I(F_0+M\cdot\tilde{F}) \equiv 0 \pmod{M^k}\}|\\
&= M^{-4k}\cdot\#|\{F \in V(\mathbb{Z}/M^k) : I(F) \equiv 0 \pmod{M^k}, F \equiv F_0 \pmod M\}|\\
&= \prod_{p|M}\begin{pmatrix} p^{-4k\cdot v_p(M)}\cdot\#|\{F \in V(\mathbb{Z}/p^{k\cdot v_p(M)}) : I(F) \equiv 0 \pmod{p^{k\cdot v_p(M)}}, F \equiv F_0 \pmod M\}| \\ +O(p^{-k\cdot v_p(M)})\end{pmatrix}.
\end{aligned}$$

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

$$M^{-5}\cdot\prod_{p|M} \sigma_p^{(F_0 \pmod M)}(Y(\mathbb{Z})) = \prod_{p|M} \sigma_p(Y_{F_0 \pmod M}(\mathbb{Z})),$$

as desired. □

5 Proof of Theorem 1.2.

Now to the proof of Theorem 1.2.

5.1 Reduction to point counting.

We run the same argument as above, except our notation follows Bhargava-Ho's [7] rather than Ruth's [11], and we appeal in the end to part (f) of Theorem 1.1 of Bhargava-Ho's [7] (rather than Theorem 3.1 of Bhargava-Shankar's [8]) to calculate the product of local densities. Because otherwise 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 [7] (the relevant parametrization is by triply symmetric hypercubes). Let $V := 2 \otimes \mathrm{Sym}_3(2)$, the space of pairs of "threes-in" binary cubic forms. Let G be the image in $\mathrm{GL}(V)$ of $\{(g,h) \in \mathrm{GL}_2 \times \mathrm{GL}_2 : \det g \cdot (\det h)^3 = 1\}$, acting in the evident way on $2 \otimes \mathrm{Sym}_3(2)$ (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 $G \cong (\mathrm{SL}_2 \times \mathrm{SL}_2)/\mu_2$.

We write, for $v \in V$,

$$H(v) := \max\left(|I_2(v)|^{\frac{1}{2}}, |I_6(v)|^{\frac{1}{6}}\right)^{24},$$

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

Let R be a fundamental domain for $G(\mathbb{R}) \curvearrowright V(\mathbb{R})^{\Delta \neq 0}$ (note that Bhargava-Ho write $V(\mathbb{R})^{\mathrm{stab}} := V(\mathbb{R})^{\Delta \neq 0}$), as constructed in Section 5 of Bhargava-Ho's [7] (via, in their notation, $R := \coprod_i R^{(i)}$). Let $L := \{v(\vec{a}) : \vec{a} \in (\mathbb{R}^m)^{\Delta \neq 0}_{H=1}\}$ (here in their notation $m = 2$ and $v(\vec{a})$ is as defined in Section 4 of Bhargava-Ho's [7]) and $\Lambda := \{(\lambda \cdot \mathrm{id}, \mathrm{id}) \in \mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R}) : \lambda \in \mathbb{R}^+\} \subseteq \mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$. Note that $R = \Lambda \cdot L$. Let $R(X) := \{v \in R : H(v) \leq X\}$. Let $\vec{F}_\pm := v((0, \pm 1)) \in L$ be the two points in L with $I_2 = 0$.

Note that, by construction, since $H((\lambda, \mathrm{id}) \cdot v) = \lambda^{24} \cdot H(v)$, the coefficients of a $v \in \lambda \cdot L \subseteq R(X)$ are all $\ll \lambda \ll X^{\frac{1}{24}}$, 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 $\ll_{G_0} \lambda \ll X^{\frac{1}{24}}$.

Let $\mathcal{F}$ be a fundamental domain for $G(\mathbb{Z}) \curvearrowright G(\mathbb{R})$, as constructed in Section 5.2 of Bhargava-Ho's [7]. Note that $\mathcal{F}$ lies inside the following Siegel set:

$$\mathcal{F} \subseteq N \cdot A \cdot K,$$

where

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

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

$$K := \mathrm{SO}_2(\mathbb{R}) \times \mathrm{SO}_2(\mathbb{R}) \subseteq G(\mathbb{R}),$$

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

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: α is K-invariant, $\int_{G(\mathbb{R})} \alpha = 1$, $\beta(\vec{F}_\pm) = 1$, and $\operatorname{supp} \beta, \beta^{-1}(\{1\}) \subseteq L^{(1)}$ are both unions of two small compact intervals respectively containing $\vec{F}_\pm$. Let

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

Let $V(\mathbb{Z})^{\mathrm{nontriv.}} := \{(F_1, F_2) \in V(\mathbb{Z}) : 0 \notin F_i(\mathbb{P}^1(\mathbb{Q}))\}$. Let $Y(\mathbb{Z}) := \{v \in V(\mathbb{Z}) : I_2(v) = 0, I_6(v) \neq 0\}$ and $Y(\mathbb{Z})^{\mathrm{nontriv.}} := Y(\mathbb{Z}) \cap V(\mathbb{Z})^{\mathrm{nontriv.}}$.

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

Write $n_{(u_1,u_2)} := (n_{u_1}, n_{u_2})$, and similarly $a_{(t_1,t_2)} := (a_{t_1}, a_{t_2})$.

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

$$\#_\varphi |B(\vec{u}, \vec{t}, \lambda, X) \cap S| := \sum_{\vec{F} \in S : ||I_6(\vec{F})||_\infty \leq X} \varphi(a_{\vec{t}}^{-1} \cdot n_{\vec{u}}^{-1} \cdot (\lambda \cdot \mathrm{id}, \mathrm{id}) \cdot (F_1, F_2)).$$

We again see that it suffices to prove the following two lemmas. 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 [7], is entirely the same as in the previous section, so we omit it.

Lemma 5.1. *Let $\lambda \in \mathbb{R}^+, u_i \in \left[-\frac{1}{2}, \frac{1}{2}\right], \sqrt{\frac{\sqrt{3}}{2}} \leq t_i \ll \lambda$. Then:*

$$\#_\varphi |B(\vec{u}, \vec{t}, \lambda, X) \cap Y(\mathbb{Z})^{\mathrm{nontriv.}}| \ll_\varphi \lambda^{6+o(1)}.$$

Lemma 5.2. *Let $M \in \mathbb{Z}^+$ and $v_0 \in V(\mathbb{Z}/M)$. Let*

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

Then:

$$\#_\varphi |B(\vec{u}, \vec{t}, \lambda, X) \cap Y_{v_0 \pmod M}(\mathbb{Z})| = \sigma_\infty(\vec{u}, \vec{t}, \lambda, X) \cdot \prod_p \sigma_p(Y_{v_0 \pmod M}(\mathbb{Z})) + O_\varphi(||\vec{t}||_\infty^8 \cdot \lambda^{4+o(1)}) + O_{\varphi,N}(||\vec{t}||_\infty^N \cdot \lambda^{O(1)-N}),$$

where

$$\sigma_\infty(\vec{u},\vec{t},\lambda,X) := \lim_{\varepsilon\to 0}\frac{\int_{v\in V(\mathbb{R}):|I_2(v)|\leq\varepsilon,|I_6(v)|\leq X} dv\,\varphi(a_{\vec{t}}^{-1}\cdot n_{\vec{u}}^{-1}\cdot(\lambda\cdot\mathrm{id},\mathrm{id})^{-1}\cdot v)}{2\varepsilon}$$

and

$$\sigma_p(Y_{v_0 \pmod M}(\mathbb{Z})) := \lim_{n\to\infty} p^{-4n}\cdot\#|\{\vec{F}\in V(\mathbb{Z}/p^n) : I_2(\vec{F})\equiv 0 \pmod{p^n}, v\equiv v_0 \pmod M\}.$$

5.2 The uniformity estimate.

In fact the proof of the uniformity estimate is identical 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

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

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

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 $G_{\vec{F}}$ 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 $G_{\vec{F}} \pmod p$ 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 $F_1=0$) $\vec{F} \pmod p$ 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 [9] to conclude.

5.3 Point counting.

The proof of Lemma 5.1 is very much the same as the proof of Lemma 4.1. We give it here anyway.

Proof of Lemma 5.1. As mentioned, each coefficient of an $(F^{(1)},F^{(2)}) =: \vec{F}\in B(\vec{u},\vec{t},\lambda,X)\cap Y(\mathbb{Z})^{\mathrm{nontriv.}}$ is $\ll_\varphi\lambda$. Applying the divisor bound, we determine $(F_0^{(1)},F_3^{(2)})$, up to $\ll_\varphi\lambda^{o(1)}$ choices, from $(F_1^{(1)},F_2^{(1)},F_3^{(1)},F_0^{(2)},F_1^{(2)},F_2^{(2)})$, and there are $\ll_\varphi\lambda^6$ choices for the latter. □

The proof of Lemma 5.2 is also much the same as the proof of Lemma 4.2. We comment on the only two differences.

Proof of Lemma 5.2. The only differences are the following. In applying the smoothed delta symbol method (i.e. Theorem 2 of Heath-Brown's [10]), 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 proof is mutatis mutandis the same. □

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

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

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

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

[8] 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

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

[10] 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

[11] 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

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