# SMOOTHNESS OF SHORTEST ISOGENIES TO FROBENIUS CONJUGATES

YI OUYANG[1,2], WEI WEI[1] AND ZHENG XU[2]

Abstract. Let $E$ be chosen uniformly from the geometric isomorphism classes of supersingular elliptic curves over $\overline{\mathbb{F}}_p$ and $\delta_E$ be the minimum degree of isogenies from $E$ to its Frobenius conjugate $E^{(p)}$. We establish bounds for the smoothness probability of $\delta_E$ and hence affirm Wesolowski's $p^{1/3+o(1)}$ attack on the supersingular endomorphism ring problem is unconditional. We also establish a uniform asymptotic formula, with a main term expressed in terms of Dirichlet $L$-values, for the probability that a connection isogeny of degree $n$ exists between $E$ and $E^{(p)}$.

## 1. Introduction

Isogeny-based cryptography forms a distinctive family of post-quantum constructions grounded in the arithmetic of elliptic curves and quaternion algebras. The landscape of isogeny-based cryptography shifted dramatically after Castryck and Decru, Maino et al., and Robert discovered polynomial-time attacks against SIDH and SIKE [CD23, MMP+23, Rob23]. These attacks leveraged the auxiliary torsion-point information exposed by SIDH and, crucially, demonstrated the power of higher-dimensional isogeny techniques.

Following the polynomial time cryptanalytic breakthroughs against SIDH[Rob23], new constructions have branched into two primary research threads. The first seeks to protect sensitive information using new masking strategies or alternative representations, yielding M-SIDH and MD-SIDH [FMP23], as well as binSIDH and terSIDH constructed from artificial orientations [BF23]. The second reuses higher-dimensional isogeny techniques to build constructive cryptographic primitives. Examples include the trapdoor-based encryption schemes FESTA and QFESTA [BMP23, NO24], and the unknown-degree isogeny encryption scheme POKÉ [BM25].

SQIsign, first introduced in 2020, attracted substantial attention due to its compact signatures [DKL+20]. A series of subsequent variants — SQIsignHD [DLRW24], SQIsign2D-West [BDD+24], SQIsign2D-East [NOC+24], SQIsign2D$^2$ [XLZO25], and SQIsign2DPush [NO26]—employ higher-dimensional isogeny representations to improve performance and reinforce security guarantees. Collectively, these developments highlight the need for rigorous analysis of the underlying computational problems in supersingular isogeny cryptography. In May 2026, NIST moved SQIsign forward into the third round of its Additional Digital Signatures standardization effort [NIST26]. More recently, Basso et al. [BDD+26] introduced MIKE, an efficient and compact post-quantum non-interactive key exchange (NIKE) scheme.

Under their respective security models and auxiliary assumptions, the security of SQIsign and SQIsign variants can be reduced to the hardness of the supersingular endomorphism-ring problem. Wesolowski established polynomial-time reductions between this problem and the supersingular isogeny path-finding problem [Wes22], allowing cryptanalytic attacks on the underlying hardness assumption to be approached through isogeny path finding. Find a nontrivial endomorphism is important since then.

2020 *Mathematics Subject Classification.* 11G20, 11N25, 11R52, 14K02.

*Key words and phrases.* Supersingular elliptic curves, least isgeny degree, smooth number, Brandt operator, Wesolowski's $p^{1/3}$-attack.

Prior to the work of Wesolowski [Wes26], the best-known classical and quantum algorithms for the general supersingular isogeny path-finding problem had time complexities $\widetilde{O}(p^{1/2})$ [DG16] and $\widetilde{O}(p^{1/4})$ [BJS14], respectively.

For a supersingular elliptic curve $E$ in characteristic $p$, let $E^{(p)}$ denote its Frobenius conjugate, and $\delta_E$ the least isogeny degree from $E$ to $E^{(p)}$:

$$\delta_E := \min_{0\neq\varphi\in\mathrm{Hom}(E,E^{(p)})} \deg\varphi.$$

[AOV26a] tells that $\delta_E \leq \lfloor \frac{p}{2}^{\frac{1}{3}} \rfloor$, and [Wes26] uses this establish a $p^{1/3+o(1)}$ attack.

However, in isogeny-based cryptography, isogenies of smooth degree are particularly amenable to efficient computation, since they can be represented and evaluated as compositions of low-degree isogenies [Feo17, Wes26]. For an isogeny of a large prime degree $N$, given a generator of its kernel, the square-root Vélu formulas allow its codomain to be computed and the isogeny to be evaluated at a point using $\widetilde{O}(\sqrt{N})$ field operations [Vél71, BFLS20]. This naturally leads to a finer arithmetic question concerning the prime factors of $\delta_E$: how often is $\delta_E$ $B$-smooth, or equivalently, how is its largest prime factor $P^+(\delta_E)$ distributed?

The recent attack by Wesolowski[Wes26] on the supersingular endomorphism ring problem underlying the security of SQIsign searches for a $B$-smooth isogeny $\varphi\colon E \longrightarrow E^{(p)}$. By composing $\varphi$ with the Frobenius from $E^{(p)}$ back to $E$, one obtains a non-scalar endomorphism of $E$. Wesolowski gave algorithms for the one-endomorphism and endomorphism-ring problems with time and space complexity $p^{1/3+o(1)}$, under the heuristic assumption that the smoothness of $\delta_E$ behaves like that of a uniformly random integer of comparable size. The following heuristic is essential to his attack:

**Heuristic W** (Heuristic 1, [Wes26])**.** *Let $p$ be a prime number, and let $E/\mathbb{F}_{p^2}$ be a uniformly random supersingular elliptic curve. Then*

$$\Pr\big(P^+(\delta_E) < B\big) \approx \rho(u) \approx u^{-u(1+o(1))},$$

*where $u = \frac{\log(p/2)}{3\log(B)}$ (uniformly as $p\to\infty$, if $(\log p)^{\varepsilon} < u < (\log p)^{1-\varepsilon}$ for a fixed constant $\varepsilon$) and $\rho$ is the Dickman-de Bruijn function.*

Mamah [Mam26] stated another heuristic in his analysis of the effectiveness of Wesolowski's algorithm, this heuristic originally commes from [BBD+24]:

**Heuristic M** (Heuristic 1, [Mam26])**.** *For a fixed degree $n$, the probability that a supersingular curve $E$ admits an $n$-isogeny to its Frobenius conjugate $E^{(p)}$ is*

$$\Pr(A_{E,n}) \approx \sqrt{\frac{12n}{p}}.$$

Our main result in this paper is the following theorem, which implies the heuristic of Wesolowski as an application:

**Theorem 1.1.** *Fix $\varepsilon \in (0,1/2)$. There exist constants $c_1, c_2$ and $M$ depending only on $\varepsilon$, such that*

$$\frac{c_1\rho(u)}{(\log p)^3} \leq \Pr\big(P^+(\delta_E) < B\big) \leq c_2\rho(u)\log p, \tag{1.1}$$

*if $p > M$ and $\frac{\log(p/2)^{\varepsilon}}{3} < \log B < \frac{\log(p/2)^{1-\varepsilon}}{3}$ (equivalently $\log(p/2)^{\varepsilon} < u < \log(p/2)^{1-\varepsilon}$).*

We also prove that the predicted logarithmic exponent of the heuristic of Mamah is correct and demonstrate that the proposed universal leading constant does not hold in general.

We briefly mention related studies in this subject. Charles, Goren, and Lauter [CGL09] studied short paths between Frobenius-conjugate vertices in supersingular $\ell$-isogeny graphs and obtained bounds on the number of vertices admitting such paths.

Aubry and Vincent related such isogenies to inseparable endomorphisms and rank-two sublattices of the Gross lattice [AV26], while Aubry, Oyono, and Vincent proved the bound

$$\delta_E \le \left\lfloor (p/2)^{1/3} \right\rfloor$$

and showed that the exponent $1/3$ is asymptotically optimal [AOV26a]. Complementary recent work studies the extremal value $\max_E \delta_E$ [Swa26] and the analogous problem for principally polarized superspecial abelian surfaces [Pha26].

The use of isogenies to a Frobenius conjugate for constructing non-scalar endomorphisms appears in [EHL+20, FIK+25]. Building on the bound of [AOV26a], Wesolowski proposed his heuristic, which was then supported by the experiments and implementations of [Pan26, AOV26b]. An independent preprint by Delgado gives an unconditional $p^{1/3+o(1)}$ algorithm by fixing in advance a family of smooth degrees and showing that isogenies of those degrees occur on sufficiently many curves [Del26].

The paper is organized as follows. In Section 2 we introduce some necessary background about smooth numbers, class numbers of imaginary quadratic orders and divisor functions. We then study the Brandt operators and primitive connection isogenies about supersingular elliptic curves in Section 3. Section 4 and Section 5 are devoted to the proof of the upper and lower bounds respectively via number-theoretic arguments, and Section 5 is the technical heart of this paper. We then deduce the heuristic of Wesolowski in Section 6. Appendix B proves that Mamah's heuristic holds in the logarithmic sense, giving the existence probability of a connection isogeny between $E$ and $E^{(p)}$ of degree $n$. Appendix A contains detailed proofs of the technical lemmas in Section 2.

**Acknowledgement.** This wok is partially supported by NSFC (Grant No. 12371013 and 12501731) and QNMP (Grant No. 2021ZD0302902).

**Declaration of AI use.** The authors used ChatGPT-5.6 Sol to assist in developing mathematical arguments and derivations presented in this paper. All AI-assisted mathematical results and proofs have been independently verified by the authors, who take full responsibility for the accuracy and integrity of the paper.

## 2. Notation and preliminaries

2.1. **Notations and conventions.** In this work, we assume $p \ge 5$ is prime and $k = \overline{\mathbb{F}}_p$. All curves, isogenies, and homomorphisms are geometric unless a field of definition is explicitly specified. We set

$$N = \left\lfloor (p/2)^{1/3} \right\rfloor, \qquad u = \frac{\log(\frac{p}{2})}{3\log B}.$$

We let $P^+(1) = 1$ and $P^+(n)$ be the largest prime factor of integer $n > 1$. Hence $n$ is $B$-smooth is exactly the condition $P^+(n) \le B$.

For real functions $f(x)$ and $g(x)$, we denote by $f = O(g)$ (respectively $f = O_\varepsilon(g)$) if there exists an absolute constant $C > 0$ (respectively $C$ depending on $\varepsilon$) such that $f \le Cg$, by $f \asymp g$ if $f = O(g)$ and $g = O(f)$.

All probabilities and expectations below refer to the uniform probability measure on the set $\mathcal{S}_p$ of isomorphism classes of supersingular elliptic curves over $k$.

The proofs of the lemmas in this section are deferred to Appendix A.

2.2. **Counting smooth numbers.** For real numbers $x \ge 1$ and $y \ge 2$, define

$$\Psi(x, y) := \#\left\{n \in \mathbb{Z}_{\ge 1} : n \le x,\ P^+(n) \le y\right\}.$$

In other words, $\Psi(x, y)$ counts $y$-smooth positive integers $\le x$.

The Dickman-de Bruijn function $\rho\colon [0, \infty) \to \mathbb{R}$ is the unique continuous function, continuously differentiable on $(1, \infty)$, satisfying the delay differential equation

$$\begin{cases} \rho(v) = 1, & 0 \le v \le 1, \\ v\rho'(v) + \rho(v-1) = 0, & v > 1. \end{cases} \tag{2.1}$$

It is positive and non increasing on $[0, \infty)$.

For every fixed $v \ge 1$, Dickman's theorem[Dic30] (see also [HT86, (1.1)–(1.3)].) gives

$$\lim_{x\to\infty} \frac{\Psi\left(x, x^{1/v}\right)}{x} = \rho(v). \tag{2.2}$$

Moreover, de Bruijn[dB51] proved that

$$\lim_{v\to\infty} \frac{-\log \rho(v)}{v \log v} = 1. \tag{2.3}$$

Equivalently, $\rho(v) = v^{-v(1+o(1))}$.

**Lemma 2.1.** *Let $N_0 = \left(\frac{p}{2}\right)^{1/3}$, $N = \lfloor N_0 \rfloor$ and $u = \frac{\log N_0}{\log B}$. Fix $0 < \varepsilon < 1/2$ and assume $(\log p)^{\varepsilon} < u < (\log p)^{1-\varepsilon}$. When $p \to \infty$, we have*

$$\frac{\Psi(N, B)}{N} = \rho(u) \left\{ 1 + O_{\varepsilon}\left( \frac{\log(u+1)}{\log B} + \frac{1}{N} \right) \right\}. \tag{2.4}$$

*In particular, $\frac{\Psi(N,B)}{N} = \rho(u)(1 + o_{\varepsilon}(1))$.*

We define the set of squarefree $B$-smooth integers in $[\frac{X}{2}, X]$ by

$$\mathcal{A}(X, B) = \left\{ n \in \mathbb{Z}_{\ge 1} : X/2 < n \le X,\ \mu(n)^2 = 1,\ P^+(n) \le B \right\}, \tag{2.5}$$

where $\mu(n)$ is the Möbius function.

**Lemma 2.2.** *Suppose $(\log X)^3 \le B \le X$ and $\log B \ge (\log\log X)^{5/3+\delta}$ for some fixed $\delta > 0$. Put $v = \frac{\log X}{\log B}$. If $v \to \infty$, then*

$$\#\mathcal{A}(X, B) \asymp X\rho(v). \tag{2.6}$$

2.3. **Class number formula of imaginary quadratic orders.** For a negative quadratic discriminant $D \equiv 0, 1 \pmod 4$, let $K_D := \mathbb{Q}(\sqrt{D})$ and $\mathcal{O}_D := \mathbb{Z}\left[\frac{D+\sqrt{D}}{2}\right]$ be the imaginary quadratic order of discriminant $D$. Let

$$h(D) := \#\operatorname{Pic}(\mathcal{O}_D)$$

be the class number of $\mathcal{O}_D$ and

$$w(D) := [\mathcal{O}_D^{\times} : \{\pm 1\}] = \begin{cases} 3, & D = -3, \\ 2, & D = -4, \\ 1, & D < -4. \end{cases}$$

Any such discriminant $D$ admits a unique factorization

$$D = f^2 D_0, \tag{2.7}$$

where $D_0$ is the fundamental discriminant of $K_D$ and $f \ge 1$ is the conductor of $\mathcal{O}_D$. Equivalently,

$$\mathcal{O}_D = \mathbb{Z} + f\mathcal{O}_{D_0},$$

where $\mathcal{O}_{D_0}$ denotes the maximal order of $K_D$.

We have the following class-number formula, see [Cox13, Theorem 7.24 and the discussion following Corollary 7.28] for details.

**Proposition 2.3.** *Let $D < 0$ be an order discriminant. Then*

$$\frac{h(D)}{w(D)} = \frac{\sqrt{|D|}}{\pi} L(1, \chi_D). \tag{2.8}$$

*where $L(s, \chi_D)$ is the Dirichlet L-function associated to the Kronecker character*

$$\chi_D \colon \mathbb{Z}_{\geq 1} \to \{-1, 0, 1\},\ m \mapsto \left(\frac{D}{m}\right),$$

*with $\left(\frac{D}{m}\right)$ the Jacobi symbol modulo $m$. Equivalently, by $\#\mathcal{O}_D^\times = 2w(D)$,*

$$L(1, \chi_D) = \frac{2\pi h(D)}{\#\mathcal{O}_D^\times \sqrt{|D|}}. \tag{2.9}$$

Also, we have the well-known upper bound for the Dirichlet L-function. This is a standard exercise; see [Dav00].

**Lemma 2.4.** *If $\chi$ is a nonprincipal Dirichlet character modulo $q$, then*

$$|L(1, \chi)| \leq \log q + 2.$$

**Lemma 2.5.** *Let $p$ be an odd prime and $2 < X < p/4$. For $n \in \mathcal{A}(X, B)$, define*

$$D_n = \begin{cases} -pn, & pn \equiv 3 \pmod 4, \\ -4pn, & \text{otherwise.} \end{cases} \tag{2.10}$$

*Then $D_n$ is the fundamental discriminant of $\mathbb{Q}(\sqrt{-pn})$, and the map $n \mapsto D_n$ is injective on $\mathcal{A}(X, B)$.*

*For sufficiently large $p$, except at most one $n \in \mathcal{A}(X, B)$, we have*

$$L(1, \chi_{D_n}) > \frac{0.655 e^{-2}}{\log p}. \tag{2.11}$$

*Moreover, if $pn \equiv 3 \pmod 4$, then the character $\chi_{4D_n} = \chi_{-4pn}$ satisfies*

$$L(1, \chi_{-4pn}) > \frac{0.655 e^{-2}}{2 \log p}. \tag{2.12}$$

2.4. **Divisor-Sum Estimates.** Let $\tau(k)$ denote the number of positive divisors of a positive integer $k$. We record two divisor-sum estimates that will be used in the subsequent sections.

**Lemma 2.6.** *For every integer $M \geq 1$, we have*

$$\sum_{\substack{t \in \mathbb{Z} \\ t^2 < 4M}} \tau(4M - t^2) = O\left(\sqrt{M}\big(1 + \log(2M)\big)^2\right), \tag{2.13}$$

*where the implied constant is absolute.*

**Lemma 2.7.** *For every squarefree integer $n \geq 1$,*

$$\sum_{m \leq n} \sum_{\substack{t \in \mathbb{Z} \\ t^2 < 4mn}} \tau(4mn - t^2) \leq 60 n^2 \big(1 + \log(2n)\big). \tag{2.14}$$

## 3. Counting connection isogenies

3.1. **Basic of supersingular elliptic curves.** Let $k = \overline{\mathbb{F}}_p$ and

$$\mathcal{S}_p = \{\text{supersingular elliptic curves over } k\} \,/\, \simeq_k$$

be the set of geometric isomorphism classes. By the usual Deuring class-number count [Deu41, Voi21], the cardinality $h_p$ of $\mathcal{S}_p$ is given by the formula

$$h_p = \left\lfloor \frac{p}{12} \right\rfloor + \epsilon_p, \quad \text{where} \quad \epsilon_p = \begin{cases} 0, & p \equiv 1 \pmod{12}, \\ 1, & p \equiv 5, 7 \pmod{12}, \\ 2, & p \equiv 11 \pmod{12}. \end{cases} \tag{3.1}$$

For $E \in \mathcal{S}_p$, define

$$\delta_E = \min_{0 \neq \varphi \in \mathrm{Hom}(E, E^{(p)})} \deg \varphi.$$

This minimum depends only on the geometric class. By [AOV26a, Theorem 4.2], $\delta_E \leq N := \lfloor (p/2)^{1/3} \rfloor$ for all $E \in \mathcal{S}_p$.

For $E \in \mathcal{S}_p$, put

$$\mathcal{O}_E := \mathrm{End}(E), \quad \mathrm{End}^0(E) := \mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Q}.$$

By Deuring's theorem, $\mathrm{End}^0(E)$ is isomorphic to the quaternion algebra $B_{p,\infty}$ ramified precisely at $p$ and at the real place, and $\mathcal{O}_E$ is a maximal order in $B_{p,\infty}$ [Deu41]. More precisely, after fixing a maximal order $\mathcal{O}_0 \subset B_{p,\infty}$, the geometric supersingular classes may be identified with the locally principal left $\mathcal{O}_0$-ideal classes, in such a way that the right order associated with $E$ is isomorphic to $\mathrm{End}(E)$; see [Voi21, Corollary 42.3.7].

Let $\pi_E : E \longrightarrow E^{(p)}$ denote the relative Frobenius morphism. It is purely inseparable of degree $p$. Associated with $\pi_E$ is the two-sided ideal generated by $\pi_E$.

$$\mathfrak{P}_E = \left\{ \beta \circ \pi_E \mid \beta : E^{(p)} \longrightarrow E \right\}. \tag{3.2}$$

**Lemma 3.1.** *The ideal $\mathfrak{P}_E$ is the unique two-sided prime ideal of $\mathcal{O}_E$ of reduced norm $p$. It satisfies $\mathfrak{P}_E^2 = p\mathcal{O}_E$, $\mathrm{trd}(\mathfrak{P}_E) \subseteq p\mathbb{Z}$. Moreover, the map $\mathrm{F}_E : \mathrm{Hom}(E, E^{(p)}) \longrightarrow \mathfrak{P}_E$ with $\varphi \longmapsto \widehat{\varphi} \circ \pi_E$, is an isomorphism of additive groups, and $\mathrm{nrd}\big(\mathrm{F}_E(\varphi)\big) = p \deg \varphi$. The map $\mathrm{F}_E$ also preserves scalar primitivity.*

*Proof.* From [Voi21, §42.2.4], $\mathfrak{P}_E^2 = p\mathcal{O}_E$. By [Voi21, 16.6.14] we have $\overline{\mathfrak{P}_E} = \mathfrak{P}_E$, and [Voi21, Theorem 18.3.6] tells $\mathfrak{P}_E \cap \mathbb{Z} = p\mathbb{Z}$, as a consequence, $\mathrm{trd}(\mathfrak{P}_E) \subseteq p\mathbb{Z}$.

By the factorisation property of the relative Frobenius, composition with $\pi_E$ induces an isomorphism of additive groups

$$\mathrm{Hom}(E^{(p)}, E) \to \mathfrak{P}_E, \qquad \beta \longmapsto \beta \circ \pi_E.$$

Composing with the duality isomorphism, we obtain the asserted additive isomorphism $\mathrm{F}_E(\varphi) = \widehat{\varphi} \circ \pi_E$. For more details about isomorphism, see also [Voi21, Lemma 42.2.7]. Moreover, $\mathrm{nrd}\big(\mathrm{F}_E(\varphi)\big) = \deg\big(\widehat{\varphi} \circ \pi_E\big) = \deg \widehat{\varphi} \deg \pi_E = p \deg \varphi$.

□

3.2. **Brandt operators.** Let

$$\mathcal{M}_p := \mathbb{Z}[\mathcal{S}_p] = \bigoplus_{[E] \in \mathcal{S}_p} \mathbb{Z}\, e_{[E]}.$$

For each $n \geq 1$, the degree-$n$ isogeny correspondence defines a natural $\mathbb{Z}$-linear operator on $\mathcal{M}_p$, whose matrix records the degree-$n$ isogeny correspondences among the supersingular isomorphism classes; see [Voi21, Definition 41.2.4 and Exercise 42.4(a)].

**Definition 3.2** (Brandt operator)**.** *For $n \geq 1$, $T_n \colon \mathcal{M}_p \to \mathcal{M}_p$, is the $\mathbb{Z}$-linear operator defined by*

$$T_n(e_{[E]}) = \sum_{\substack{C \subset E \\ \#C = n}} e_{[E/C]},$$

*After choosing an ordering $\mathcal{S}_p = \{[E_1], \ldots, [E_{h_p}]\}$,, we denote the matrix of $T_n$, with respect to the ordered basis $(e_{[E_1]}, \ldots, e_{[E_{h_p}]})$, by $T(n)$ in the column-source convention, i.e.*

$$T(n)_{ij} = \#\left\{ C \subset E_j : \begin{array}{l} C \text{ is a finite subgroup of order } n, \\ E_j/C \cong E_i \end{array} \right\}.$$

When $(n, p) = 1$, every subgroup of order $n$ is the kernel of a separable isogeny of degree $n$. Consequently,

$$T(n)_{ij} = \#\left(\mathrm{Aut}(E_i) \backslash \{\varphi \colon E_j \to E_i : \deg \varphi = n\}\right), \tag{3.3}$$

where $\mathrm{Aut}(E_i)$ acts by postcomposition.

At the ramified prime $p$, a supersingular elliptic curve $E$ admits a unique isogeny of degree $p$, namely $\pi_E$. Hence $T_p(e_{[E]}) = e_{[E^{(p)}]}$. We write $\mathcal{W}_p := T_p$ for the ramified Brandt operator and $W_p := T(p)$ for its matrix. Thus $W_p$ is the permutation matrix induced by $[E] \longmapsto [E^{(p)}]$. Since $[E^{(p)}]^{(p)} = [E]$, $\mathcal{W}_p^2 = \mathrm{id}_{\mathcal{M}_p}$ which corresponds to matrix $I_{h_p}$. See [Voi21, §42.2.4 and Exercise 42.4(b)–(c)].

For $n \geq 1$, define

$$\mathcal{R}_E(n) := \mathrm{Aut}(E^{(p)}) \backslash \left\{ \varphi \colon E \to E^{(p)} \mid \deg \varphi = n \right\},$$

where $\mathrm{Aut}(E^{(p)})$ acts by postcomposition: $u \cdot \varphi := u \circ \varphi$. We refer to the elements of $\mathcal{R}_E(n)$ as *degree-$n$ connection isogenies*, and set

$$r_E(n) := \#\mathcal{R}_E(n).$$

Thus, $r_E(n)$ counts orbits of degree-$n$ connection isogeny from $E$ to $E^{(p)}$. With the column-source convention for the Brandt matrices, one has

$$r_E(n) = T(n)_{E^{(p)}, E}. \tag{3.4}$$

Consequently,

$$\sum_{E \in \mathcal{S}_p} r_E(n) = \sum_E \sum_F (W_p)_{E,F} (T(n))_{F,E} = \mathrm{tr}\big(W_p T(n)\big). \tag{3.5}$$

There is an equivalent ideal-theoretic description of the Brandt operator, obtained through the Deuring correspondence recalled in Section 3; see [Voi21, Corollary 42.3.7]. Recall the maximal order $\mathcal{O}_0 \subset B_{p,\infty}$ fixed in Section 3. Under the Deuring correspondence, the geometric supersingular classes are identified with the locally principal left $\mathcal{O}_0$-ideal classes. Choose representatives $I_1, \ldots, I_{h_p}$ indexed compatibly with $\mathcal{S}_p = \{[E_1], \ldots, [E_{h_p}]\}$, so that $\mathcal{O}_R(I_i) \simeq \mathrm{End}(E_i)$.

With this indexing, the same Brandt matrix admits the equivalent ideal-theoretic description

$$T(n)_{ij} = \#\left\{ J \subset I_j : \mathrm{nrd}(J) = n \, \mathrm{nrd}(I_j), \quad [J]_L = [I_i]_L \right\}. \tag{3.6}$$

Equation (3.6) is precisely the ideal-class realization of the geometric Brandt operator of Definition 3.2. [1]

Through the ideal-class realization above, the trace of the Brandt operator can be computed as follows.

[1] Voight formulates Brandt matrices using right ideals[Voi21]. The left-ideal formulation used here is obtained by ideal conjugation, which changes the matrix by transposition and hence does not affect its trace.

Let $\widehat{\mathbb{Z}} := \prod_{\ell<\infty} \mathbb{Z}_\ell$ be the profinite completion of $\mathbb{Z}$, where $\ell$ ranges over all finite primes. For an imaginary quadratic order $\mathcal{O}_D$, put $\widehat{\mathcal{O}_D} := \mathcal{O}_D \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}} = \prod_{\ell<\infty} \mathcal{O}_{D,\ell}$, where $\mathcal{O}_{D,\ell} := \mathcal{O}_D \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$. Similarly, $\widehat{\mathcal{O}}_0 := \mathcal{O}_0 \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}} = \prod_{\ell<\infty} \mathcal{O}_{0,\ell}$ with $\mathcal{O}_{0,\ell} := \mathcal{O}_0 \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$. The group $\widehat{\mathcal{O}}_0^\times = \prod_{\ell<\infty} \mathcal{O}_{0,\ell}^\times$ is the unit group of the finite adelic order $\widehat{\mathcal{O}}_0$.

**Definition 3.3.** *A $\mathbb{Z}_\ell$-algebra embedding $\iota_\ell \colon \mathcal{O}_{D,\ell} \hookrightarrow \mathcal{O}_{0,\ell}$ is called* optimal *if $\iota_\ell(K_D \otimes_{\mathbb{Q}} \mathbb{Q}_\ell) \cap \mathcal{O}_{0,\ell} = \iota_\ell(\mathcal{O}_{D,\ell})$. Let $\mathrm{Emb}^{\mathrm{opt}}(\mathcal{O}_{D,\ell}, \mathcal{O}_{0,\ell})$ denote the set of optimal embeddings $\iota_\ell : \mathcal{O}_{D,\ell} \hookrightarrow \mathcal{O}_{0,\ell}$. The local optimal-embedding number is defined by*

$$m\Big(\mathcal{O}_{D,\ell}, \mathcal{O}_{0,\ell}; \mathcal{O}_{0,\ell}^\times\Big) := \#\left(\mathcal{O}_{0,\ell}^\times \backslash \mathrm{Emb}^{\mathrm{opt}}(\mathcal{O}_{D,\ell}, \mathcal{O}_{0,\ell})\right), \tag{3.7}$$

*where $\mathcal{O}_{0,\ell}^\times$ acts by conjugation on the target: $(u \cdot \iota)(x) := u\,\iota(x)\,u^{-1}$.*

The corresponding finite adelic optimal embedding number is

$$m\Big(\widehat{\mathcal{O}_D}, \widehat{\mathcal{O}_0}; \widehat{\mathcal{O}_0}^\times\Big) = \prod_{\ell<\infty} m\Big(\mathcal{O}_{D,\ell}, \mathcal{O}_{0,\ell}; \mathcal{O}_{0,\ell}^\times\Big). \tag{3.8}$$

**Lemma 3.4.** *For a nonsquare positive integer $N$,*

$$\operatorname{tr} T(N) = \frac{1}{2} \sum_{\substack{t\in\mathbb{Z} \\ t^2<4N}} \sum_{\substack{Df^2=t^2-4N \\ D<0}} \frac{h(D)}{w(D)} m\Big(\widehat{\mathcal{O}_D}, \widehat{\mathcal{O}_0}; \widehat{\mathcal{O}_0}^\times\Big). \tag{3.9}$$

This is the specialization to $F = \mathbb{Q}$, $R = \mathbb{Z}$, and the maximal order $\mathcal{O}_0 \subset B_{p,\infty}$ of the Eichler trace formula; see [Voi21, Theorem 41.5.2].

### 3.3. **Primitive connection isogenies.**

**Definition 3.5.** *A connection isogeny $\varphi \colon E \to E^{(p)}$ is called* primitive *if $\varphi \neq [d]_{E^{(p)}} \circ \psi$ for any $d > 1$ and $\psi : E \to E^{(p)}$.*

Clearly, by definition, a connection isogeny whose degree is squarefree or minimal must be primitive.

Let $\mathcal{R}_E^\circ(n) \subseteq \mathcal{R}_E(n)$ denote the subset of primitive connection isogenies, and put

$$r_E^\circ(n) := \#\mathcal{R}_E^\circ(n).$$

Decomposing a connection isogeny by its scalar factor gives

$$r_E(n) = \sum_{d^2|n} r_E^\circ\left(\frac{n}{d^2}\right). \tag{3.10}$$

We define the total mass of primitive Brandt–Hurwitz type of degree $n$ by

$$A_p^{\mathrm{BH}}(n) := \sum_{E\in\mathcal{S}_p} r_E^\circ(n). \tag{3.11}$$

**Proposition 3.6.** *If $1 \le n < p/4$, then*

$$A_p^{\mathrm{BH}}(n) = \frac{\sqrt{pn}}{\pi} L(1, \chi_{-4pn}) + \frac{\sqrt{pn}}{2\pi} \mathbf{1}_{\{pn\equiv 3 \mod 4\}} L(1, \chi_{-pn}). \tag{3.12}$$

*Proof.* By (3.5), the number of all degree-$n$ connection isogenies is

$$\sum_{E\in\mathcal{S}_p} r_E(n) = \sum_{E\in\mathcal{S}_p} T(n)_{E^{(p)},E} = \operatorname{tr}(W_p T(n)) = \operatorname{tr}(T(pn)),$$

where the last equality follows from $W_p T(n) = T(pn)$ [Voi21, Proposition 41.3.1(b)]. We now apply Lemma 3.4 to the Brandt operator $T(pn)$.

We next constrain the admissible reduced trace parameters. Let $\mathcal{O}_0 \subset B_{p,\infty}$ denote the maximal orders entering the Brandt trace formula as we used in Lemma 3.4. We need to sum up $\frac{h(D)}{w(D)} m(\widehat{\mathcal{O}_D}, \widehat{\mathcal{O}_0}; \widehat{\mathcal{O}_0}^\times)$ for all integer triple $(t, D, f)$ satisfying $Df^2 = t^2 - 4pn$ and $t^2 < 4pn$.

We now show that only the term with $t = 0$ can contribute. Fix a triple $(t, D, f)$ appearing in the Eichler trace formula. Since $Df^2 = t^2 - 4pn$, $\mathbb{Z}[x]/(x^2 - tx + pn) \subseteq \mathcal{O}_D$ with index $f$.

Hence, the embedding image $\alpha \in \mathcal{O}_0$ of $x$ satisfies $\mathrm{trd}(\alpha) = t$, and $\mathrm{nrd}(\alpha) = pn$. The element $\alpha$ lies in the unique maximal two-sided ideal $\mathfrak{P}_p$ of $\mathcal{O}_0$. As $\mathrm{trd}(\mathfrak{P}_p) \subseteq p\mathbb{Z}$, we obtain $p \mid t$. On the other hand, the trace formula restricts to $t^2 < 4pn < p^2$, where the last inequality follows from $n < p/4$. Thus $|t| < p$, and consequently $t = 0$.

When $t = 0$, the imaginary quadratic order generated by $\alpha$ has discriminant $t^2 - 4pn = -4pn$. $\mathcal{O}_D$ has discriminant $D = -\frac{4pn}{f^2}$ satisfying $f^2 \mid 4pn$ and $D \equiv 0, 1 \pmod 4$.

We now evaluate the local optimal-embedding numbers $m(\mathcal{O}_{D,\ell}, \mathcal{O}_{0,\ell}; \mathcal{O}_{0,\ell}^\times)$. At $\ell = p$, since $f^2 \mid 4pn$, $v_p(D) = 1$. It follows that $\mathcal{O}_{D,p}$ is the maximal order in a ramified quadratic extension of $\mathbb{Q}_p$. Moreover, $B_{p,\infty} \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is the quaternion division algebra over $\mathbb{Q}_p$, and $\mathcal{O}_{0,p}$ is its maximal order. Therefore, [Voi21, Proposition 30.5.3(b)] gives $m\Big(\mathcal{O}_{D,p}, \mathcal{O}_{0,p}; \mathcal{O}_{0,p}^\times\Big) = 1$.

For every finite prime $\ell \neq p$, the quaternion algebra splits: $B_{p,\infty} \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \simeq M_2(\mathbb{Q}_\ell)$. Moreover, $\mathcal{O}_{0,\ell}$ is a maximal order and is therefore conjugate to $M_2(\mathbb{Z}_\ell)$. Hence [Voi21, Proposition 30.5.3(a)] gives $m\Big(\mathcal{O}_{D,\ell}, \mathcal{O}_{0,\ell}; \mathcal{O}_{0,\ell}^\times\Big) = 1$. Together with the corresponding equality at $\ell = p$, this shows that every finite local optimal-embedding factor in the trace formula is equal to 1. Consequently,

$$m\Big(\widehat{\mathcal{O}_D}, \widehat{\mathcal{O}_0}; \widehat{\mathcal{O}_0}^\times\Big) = \prod_{\ell < \infty} m\Big(\mathcal{O}_{D,\ell}, \mathcal{O}_{0,\ell}; \mathcal{O}_{0,\ell}^\times\Big) = 1.$$

The trace formula therefore becomes

$$\sum_{E \in \mathcal{S}_p} r_E(n) = \operatorname{tr} T(pn) = \frac{1}{2} \sum_{\substack{d^2 \mid 4pn \\ 4pn/d^2 \equiv 0,3 \pmod 4}} \frac{h(-4pn/d^2)}{w(-4pn/d^2)}. \tag{3.13}$$

Write

$$\beta_p(n) = \frac{1}{2}\left(\frac{h(-4pn)}{w(-4pn)} + \mathbf{1}_{\{pn \equiv 3 \pmod 4\}} \frac{h(-pn)}{w(-pn)}\right).$$

Then

$$\sum_{E \in \mathcal{S}_p} r_E(n) = \sum_{d^2 \mid n} \beta_p(n/d^2).$$

On the other hand,

$$\sum_{E \in \mathcal{S}_p} r_E(n) = \sum_{d^2 \mid n} \sum_{E \in \mathcal{S}_p} r_E^\circ(n/d^2) = \sum_{d^2 \mid n} A_p^{\mathrm{BH}}(n/d^2).$$

Hence

$$A_p^{\mathrm{BH}}(n) = \beta_p(n) = \frac{1}{2}\left(\frac{h(-4pn)}{w(-4pn)} + \mathbf{1}_{\{pn \equiv 3 \pmod 4\}} \frac{h(-pn)}{w(-pn)}\right). \tag{3.14}$$

Apply the class number formula in Proposition 2.3 into (3.14) then gives (3.12). □

## 4. Proof of the upper bound

We are ready to prove the upper bound:

**Proposition 4.1.** *For every fixed $0 < \varepsilon < 1/2$, there is a constant $c_2 > 0$ such that the following holds for all sufficiently large $p$ and all $B$ satisfying $L^{\varepsilon} < u < L^{1-\varepsilon}$:*

$$\Pr\big(P^+(\delta_E) \le B\big) \le c_2 \rho(u) \log p.$$

*Proof.* A shortest connection is primitive, and $\delta_E \le N$. Hence, pointwise in $\mathcal{S}_p$,

$$\mathbf{1}_{\{P^+(\delta_E)\le B\}} \le \sum_{\substack{1\le n\le N\\ P^+(n)\le B}} r_E^{\circ}(n).$$

Averaging over $\mathcal{S}_p$ and using (3.11), we have:

$$\Pr\big(P^+(\delta_E) \le B\big) \le \frac{1}{h_p} \sum_{\substack{n\le N\\ P^+(n)\le B}} A_p^{\mathrm{BH}}(n). \tag{4.1}$$

For sufficiently large $p$, by Proposition 3.6 and Lemma 2.4,

$$A_p^{\mathrm{BH}}(n) \le \frac{3\sqrt{pn}}{2\pi}\big(\log(4pn) + 2\big). \tag{4.2}$$

Consequently, for some absolute constant $C_{upper} > 0$:

$$\begin{aligned}\Pr\big(P^+(\delta_E) \le B\big) &\le \frac{3\sqrt{pN}}{2\pi h_p}\big(\log(4pN) + 2\big)\Psi(N, B)\\ &\le C_{upper}(\log p)\frac{\Psi(N,B)}{N},\end{aligned} \tag{4.3}$$

where we used $h_p \asymp p$ and $N \asymp p^{1/3}$. Applying Lemma 2.1 proves the result. □

## 5. Proof of the lower bound

5.1. **Quaternionic pairs estimate.** For $1 \le m, n < p/4$, define the count of ordered distinct pairs

$$C_p(m,n) = \begin{cases} \displaystyle\sum_{E\in\mathcal{S}_p} r_E^{\circ}(m) r_E^{\circ}(n), & m \ne n,\\ \displaystyle\sum_{E\in\mathcal{S}_p} r_E^{\circ}(n)\big(r_E^{\circ}(n) - 1\big), & m = n. \end{cases} \tag{5.1}$$

Thus $C_p(m,n)$ counts ordered pairs of distinct primitive connection isogenies of degrees $m$ and $n$ on a common curve. We will count the number of pairs by counting corresponding maximal order.

For each $E \in \mathcal{S}_p$ and every integer $1 \le d < p/4$, we fix, once and for all, a representative $\varphi_x : E \to E^{(p)}$ for each $x \in \mathcal{R}_E^{\circ}(d)$. These representatives are kept fixed for the entirety of the counting argument.

Fix a supersingular elliptic curve $E$ and let $\varphi_m, \varphi_n : E \to E^{(p)}$ represent two distinct primitive connection-isogeny of degrees $m$ and $n$, respectively. We associate to this ordered pair a triple $(m, n, t)$ and a corresponding quaternion order $R \subset B_{p,\infty}$.

Recall the the additive isomorphism established in Lemma 3.1:

$$\mathrm{F}_E : \mathrm{Hom}(E, E^{(p)}) \longrightarrow \mathfrak{P}_E \tag{5.2}$$

where $\varphi \longmapsto \widehat{\varphi} \circ \pi_E$.

We first apply $\mathrm{F}_E$ to the ordered pair $(\varphi_m, \varphi_n)$, obtaining $i := \mathrm{F}_E(\varphi_m), j := \mathrm{F}_E(\varphi_n)$. Then $\mathrm{Nrd}(i) = pm$ and $\mathrm{Nrd}(j) = pn$. Consider the $\mathbb{Z}$-module $R := \mathbb{Z}[1, i, j, ij] \subseteq B_{p,\infty}$. We will show that $R$ is a full rank order in $B_p$.

We now associate a integer triple $(m, n, t)$ with the isogeny pair $(\varphi_m, \varphi_n)$ by the order $R$ [2]. Throughout our counting argument, we restrict to primitive connection isogenies of degrees $m, n < X < p/4$. Since $i, j \in \mathfrak{P}_E$, we have $\mathrm{trd}(i), \mathrm{trd}(j) \in p\mathbb{Z}$. Moreover, $m < X < p/4$, so $\mathrm{trd}(i)^2 \leq 4\,\mathrm{Nrd}(i) = 4pm < p^2$. Thus $\mathrm{trd}(i) = 0$, and similarly $\mathrm{trd}(j) = 0$.

Furthermore, since $\mathfrak{P}_E^2 = p\mathcal{O}_E$, we have $ij \in p\mathcal{O}_E$, and hence $\mathrm{trd}(ij) \in p\mathbb{Z}$. Using $\overline{i} = -i$ and $\overline{j} = -j$, we obtain $\mathrm{trd}(ij) = ij + \overline{ij} = ij + ji \in p\mathbb{Z}$. Consequently, there exists a unique integer $t$ such that

$$i^2 = -pm, \qquad j^2 = -pn, \qquad ij + ji = -pt. \tag{5.3}$$

These relations associate the triple $(m, n, t)$ with the ordered pair $(\varphi_m, \varphi_n)$ and determine the multiplication relations in $R = \mathbb{Z}[1, i, j, ij]$. Moreover, a direct computation gives $\mathrm{discrd}(R) = p^2(4mn - t^2)$. If $\mathrm{discrd}(R) = 0$, then $i$ and $j$ are linearly dependent over $\mathbb{Q}$. This implies $i = \pm j$. By the injectivity of $\mathrm{F}_E$, we obtain $\varphi_m = \pm\varphi_n$, contradicting the assumption that they represent distinct connection isogeny classes. Thus $\mathrm{discrd}(R) > 0$. This shows that $R = \mathbb{Z}[1, i, j, ij]$ is a full rank order in $B_{p,\infty}$, with reduced discriminant $\mathrm{discrd}(R) = p^2(4mn - t^2)$.

The bijection $\mathrm{F}_E$ associates a triple $(m, n, t)$, with $4mn - t^2 > 0$, to each pair of distinct primitive connection isogenies. However, counting such triples alone does not suffice to count the number of isogeny pairs, since multiple pairs may give rise to the same triple. We therefore need how many pairs of primitive connection isogenies corresponding to each triple $(m, n, t)$.

For $1 \leq m, n < p/4$ and $E \in \mathcal{S}_p$, define

$$\mathscr{P}_E(m, n) := \{(x, y) \in \mathcal{R}_E^\circ(m) \times \mathcal{R}_E^\circ(n) : x \neq y\}.$$

$C_p(m, n)$ counts all pairs of connecting isogenies of degrees $m$ and $n$ over all curves in $\mathcal{S}_p$, so $C_p(m, n) = \sum_{E \in \mathcal{S}_p} \#\mathscr{P}_E(m, n)$.

For each $(x, y) \in \mathscr{P}_E(m, n)$, set $t_E(x, y) := -\frac{1}{p}\,\mathrm{trd}\,(\mathrm{F}_E(x)\mathrm{F}_E(y))$. By the preceding argument, $t_E(x, y)$ is an integer satisfying $t_E(x, y)^2 < 4mn$.

For each integer $t$, set the subset of $\mathscr{P}_E(m, n)$ consists of pairs with $t_E(x, y) = t$

$$\mathscr{P}_{E,t}(m, n) := \{(x, y) \in \mathscr{P}_E(m, n) : \mathrm{trd}\,(\mathrm{F}_E(x)\mathrm{F}_E(y)) = -pt\},$$

and define

$$C_{p,t}(m, n) := \sum_{E \in \mathcal{S}_p} \#\mathscr{P}_{E,t}(m, n). \tag{5.4}$$

Since every pair of distinct connection isogenies yields a unique integer $t$, we obtain

$$C_p(m, n) = \sum_{\substack{t \in \mathbb{Z} \\ t^2 < 4mn}} C_{p,t}(m, n). \tag{5.5}$$

We next classify the isogeny pairs associated with a fixed triple $(m, n, t)$ according to the isomorphism class of $\mathrm{End}(E)$. For each maximal-order class $[\mathcal{M}]$ in $B_{p,\infty}$, let $C_{p,t}(m, n; [\mathcal{M}])$ denote the number of such pairs over all $E \in \mathcal{S}_p$ with $\mathrm{End}(E) \cong \mathcal{M}$. That is, we define

$$C_{p,t}(m, n; [\mathcal{M}]) := \sum_{\substack{E \in \mathcal{S}_p \\ \mathrm{End}(E) \simeq \mathcal{M}}} \#\mathscr{P}_{E,t}(m, n).$$

[2]Strictly speaking, the parameter $t$ depends on the choice of representatives for the primitive connection isogeny pairs. For example, replacing $\varphi_m$ by $-\varphi_m$ sends $t$ to $-t$. Throughout this counting argument, we fix one representative for each pair once and for all. With these choices, every distinct ordered pair determines a unique parameter $t$, and all subsequent upper bounds are independent of the choices.

Then

$$C_{p,t}(m,n) = \sum_{[\mathcal{M}]} C_{p,t}(m,n;[\mathcal{M}]).$$

Consequently,

$$C_p(m,n) = \sum_{\substack{t\in\mathbb{Z}\\ t^2<4mn}} \sum_{[\mathcal{M}]} C_{p,t}(m,n;[\mathcal{M}]).$$

We therefore bound $C_p(m,n)$ by estimating $C_{p,t}(m,n;[\mathcal{M}])$ for each fixed triple $(m,n,t)$ and maximal-order class $[\mathcal{M}]$.

Fix a triple $(m,n,t)$ with $D = 4mn - t^2 > 0$, and choose a corresponding pair $(i,j)$ defining the order $R = \mathbb{Z}[1,i,j,ij] \subset B_{p,\infty}$. Since the relations in (5.3) determine the multiplication table on the basis $(1,i,j,ij)$, any other pair $(i',j')$ satisfying the same relations induces a $\mathbb{Q}$-algebra automorphism of $B_{p,\infty}$ sending $i$ to $i'$ and $j$ to $j'$. By the Skolem–Noether theorem [Voi21, Theorem 7.1.3], this automorphism is given by conjugation by some $b \in B_{p,\infty}^\times$. Consequently,

$$R' = \mathbb{Z}[1,i',j',i'j'] = bRb^{-1}.$$

Now fix a maximal order $\mathcal{M} \subset B_{p,\infty}$. Every pair $(i',j')$ satisfying (5.3) and lying in $\mathcal{M}$ determines an embedding $\iota : R \hookrightarrow \mathcal{M}$, where $\iota(i) = i', \iota(j) = j'$. Thus, for a fixed curve $E$ with $\mathrm{End}(E) \cong \mathcal{M}$, the number of admissible primitive connection-isogeny pairs associated with $(m,n,t)$ is bounded by the number of such embeddings.

We next establish a bijection between the set of embeddings and a suitable quotient of the set of conjugating elements.

For a maximal order $\mathcal{M} \subset B_{p,\infty}$, let $\mathrm{Emb}^{\mathrm{lab}}(R,\mathcal{M})$ denote the set of unital $\mathbb{Z}$-algebra embeddings $\iota : R \hookrightarrow \mathcal{M}$, recorded by the ordered pair $(\iota(i),\iota(j))$. Put $N_{B_{p,\infty}^\times}(\mathcal{M})$ the normalizer of $\mathcal{M}$ in $B_{p,\infty}$. Define

$$\Omega_R([\mathcal{M}]) = \left\{\mathcal{M}' \subset B_{p,\infty} : R \subseteq \mathcal{M}',\ \mathcal{M}' \text{ maximal and conjugate to } \mathcal{M}\right\},$$

$$A(\mathcal{M}) = N_{B_{p,\infty}^\times}(\mathcal{M})/(\mathbb{Q}^\times\mathcal{M}^\times),$$

and counting them by $N_R([\mathcal{M}]) = \#\Omega_R([\mathcal{M}])$ and $a_{\mathcal{M}} = \#A(\mathcal{M})$.

**Lemma 5.1.** *There is a natural identification*

$$\mathcal{M}^\times\backslash \mathrm{Emb}^{\mathrm{lab}}(R,\mathcal{M}) \simeq \mathcal{M}^\times\backslash \left\{b \in B_{p,\infty}^\times : bRb^{-1} \subseteq \mathcal{M}\right\}/\mathbb{Q}^\times. \tag{5.6}$$

*The map* $\Gamma\colon \mathcal{M}^\times\backslash \left\{b \in B_{p,\infty}^\times : bRb^{-1} \subseteq \mathcal{M}\right\}/\mathbb{Q}^\times \to \Omega_R([\mathcal{M}])$ *with* $[b] \longmapsto b^{-1}\mathcal{M}b$ *is surjective, and every fiber is an* $A(\mathcal{M})$*-torsor. Consequently,*

$$\#\big(\mathcal{M}^\times\backslash \mathrm{Emb}^{\mathrm{lab}}(R,\mathcal{M})\big) = a_{\mathcal{M}} N_R([\mathcal{M}]). \tag{5.7}$$

*Proof.* Since $R$ is full rank so every embedding has the form $r \mapsto brb^{-1}$ (for extending to $B_{p,\infty}$ is uniquely up to automorphism of $B_{p,\infty}$). Two conjugates determine the same embedding exactly when they differ on the right by $\mathbb{Q}^\times$. So the first statement is true.

The condition $bRb^{-1} \subseteq \mathcal{M}$ is equivalent to $R \subseteq b^{-1}\mathcal{M}b$, so $\Gamma$ is well defined and surjective.

The normalizer $N(\mathcal{M})$ acts on $\mathcal{M}^\times\backslash\left\{b \in B_{p,\infty}^\times : bRb^{-1} \subseteq \mathcal{M}\right\}/\mathbb{Q}^\times$ by left multiplication: $g\cdot[b] := [gb]$. This action is well defined and preserves image of $\Gamma$. The subgroup $\mathbb{Q}^\times\mathcal{M}^\times$ acts trivially, therefore the action descends to $A(\mathcal{M}) = N(\mathcal{M})/(\mathbb{Q}^\times\cdot\mathcal{M}^\times)$.

If $\Gamma([b_1]) = \Gamma([b_2])$, then $b_1^{-1}\mathcal{M}b_1 = b_2^{-1}\mathcal{M}b_2$, which means $b_1b_2^{-1} \in N(\mathcal{M})$ and $b_1A(\mathcal{M}) = b_2A(\mathcal{M})$. Therefore, the kernel of $\Gamma$ is exactly $A(\mathcal{M})$. In particular, we have

$$\#\left(\mathcal{M}^\times\backslash\mathrm{Emb}^{\mathrm{lab}}(R,\mathcal{M})\right) = a_M N_R([\mathcal{M}]).$$

□

For a class of maximal order $[\mathcal{M}]$, consider the number of supersingular curve classes with endomorphism ring in $[\mathcal{M}]$:

$$c_{\mathcal{M}} = \#\{E \in \mathcal{S}_p : \mathrm{End}(E) \simeq \mathcal{M}\}.$$

**Lemma 5.2.** *For every maximal order $\mathcal{M} \subset B_{p,\infty}$,*

$$a_{\mathcal{M}} c_{\mathcal{M}} = 2. \tag{5.8}$$

*Consequently,*

$$c_{\mathcal{M}} \#\big(\mathcal{M}^\times \backslash \mathrm{Emb}^{\mathrm{lab}}(R, \mathcal{M})\big) = 2N_R([\mathcal{M}]). \tag{5.9}$$

*Proof.* Consider type map:

$$\mathbf{type}\colon \mathbf{Cl}(\mathcal{O}_0) \to \{[\mathcal{M}] \mid \mathcal{M} \subset B_{p,\infty} \text{ and maximal}\}$$

with $\mathbf{type}([I]) = [\mathcal{O}_R(I)]$. The fiber of the type map over $[\mathcal{M}]$ is identified with $\mathrm{Idl}(\mathcal{M})/\mathrm{PIdl}(\mathcal{M})$, the group of invertible two-sided fractional ideals modulo principal two-sided ideals; see [Voi21, Proposition 18.5.10, Corollary 42.3.7, and Lemma 42.4.1]. Hence

$$c_{\mathcal{M}} = \#\big(\mathrm{Idl}(\mathcal{M})/\mathrm{PIdl}(\mathcal{M})\big).$$

Here $\mathrm{Pic}(\mathcal{M})$ denotes the group of isomorphism classes of invertible $\mathbb{Z}$-central $\mathcal{M}$-bimodules. We then have the exact sequence

$$1 \longrightarrow A(\mathcal{M}) \longrightarrow \mathrm{Pic}(\mathcal{M}) \longrightarrow \mathrm{Idl}(\mathcal{M})/\mathrm{PIdl}(\mathcal{M}) \longrightarrow 1; \tag{5.10}$$

see [Voi21, Theorem 18.1.3 and Proposition 18.5.3]. Since $\mathbb{Z}$ has trivial ideal class group and $p$ is the only finite ramified prime of $B_{p,\infty}$, the group $\mathrm{Pic}(\mathcal{M})$ has order two, generated by the two-sided prime above $p$. Taking orders in (5.10) proves (5.8); (5.9) follows from Lemma 5.1. □

Combining the preceding results, we obtain a upper bound for $C_{p,t}(m, n; [\mathcal{M}])$.

**Lemma 5.3.** *Let $(m, n, t)$ be a signed triple arising from an isogeny pair $(\varphi_m, \varphi_n)$ with (5.3), and let $R = \mathbb{Z}[1, i, j, ij]$ be the corresponding order. Then*

$$C_{p,t}(m, n; [\mathcal{M}]) \le c_{\mathcal{M}} \#\mathcal{M}^\times \#\big(\mathcal{M}^\times \backslash \mathrm{Emb}^{\mathrm{lab}}(R, \mathcal{M})\big). \tag{5.11}$$

*Let $N_{\max}(R) = \#\{\mathcal{M}' \subset B_{p,\infty} : R \subseteq \mathcal{M}',\ \mathcal{M}' \text{ maximal}\}$, then*

$$C_{p,t}(m, n) \le 48 N_{\max}(R). \tag{5.12}$$

*Proof.* Consider supersingular elliptic curve $E$, $\iota_E\colon \mathrm{End}(E) \cong \mathcal{M}$. The isogeny pair can be represented by:

$$([\varphi_m], [\varphi_n]) \longmapsto \big(\iota_E(\mathrm{F}_E(\varphi_m)), \iota_E(\mathrm{F}_E(\varphi_n))\big).$$

The image satisfies the (5.3) and therefore gives an embedding $R \hookrightarrow \mathcal{M}$. This induces a map

$$\Phi_E\colon \mathscr{P}_{E,t}(m, n) \to \mathrm{Emb}^{\mathrm{lab}}(R, \mathcal{M})$$

For fixed $E$, $\Phi_E$ is injective. An $\mathcal{M}^\times$-orbit contains at most $\#\mathcal{M}^\times$ embeddings, proving (5.11) after summing over the $c_{\mathcal{M}}$ curves of this type.

Since $\mathcal{M}^\times \simeq \mathrm{Aut}(E)$, one has $\#\mathcal{M}^\times \le 24$ [Sil09, Theorem III.10.1]. By Lemma 5.2,

$$\begin{aligned} C_{p,t}(m, n) &\le 24 \sum_{[\mathcal{M}]} c_{\mathcal{M}} a_{\mathcal{M}} N_R([\mathcal{M}]) \\ &= 48 \sum_{[\mathcal{M}]} N_R([\mathcal{M}]) = 48 N_{\max}(R), \end{aligned}$$

where the last equality partitions the maximal orders containing $R$ by type. □

We next derive an upper bound for $N_{\max}(R)$. For each finite prime $q$, we first bound the number of maximal $\mathbb{Z}_q$-orders containing $R_q := R \otimes_{\mathbb{Z}} \mathbb{Z}_q$, and then combine these local bounds to obtain a global estimate.

5.2. **Counting maximal orders containing $R$.** We next derive an upper bound for $N_{\max}(R)$. Set $D = 4mn - t^2$. For each finite prime $q$, put

$$\mathcal{B}_q(R) = \{\mathcal{M}_q \subset (B_{p,\infty})_q : \mathcal{M}_q \text{ maximal over } \mathbb{Z}_q,\ R_q \subseteq \mathcal{M}_q\}. \tag{5.13}$$

First, we factor $N_{\max}(R)$ to product of $\mathcal{B}_q(R)$. Define the restricted product of the sets $\mathcal{B}_q(R)$ by

$$\prod_q{}' \mathcal{B}_q(R) = \left\{ (\mathcal{M}_q)_q \in \prod_q \mathcal{B}_q(R) \mid \mathcal{M}_q = R_q \text{ for all but finitely many } q \right\}.$$

Since $\operatorname{discrd}(R) = p^2 D$, the local order $R_q$ is maximal for $q \nmid p^2 D$, so $\mathcal{B}_q(R) = \{R_q\}$ for all but finitely many $q$. From the local-global correspondence, we have the following lemma without proof.

**Lemma 5.4.** *There is an isomorphism:*

$$\begin{aligned} \{\mathcal{M} \subset B_{p,\infty} : \mathcal{M} \text{ is a maximal order and } R \subseteq \mathcal{M}\} &\xrightarrow{\sim} \prod_q{}' \mathcal{B}_q(R) \\ \mathcal{M} &\longmapsto \left(\mathcal{M} \otimes_{\mathbb{Z}} \mathbb{Z}_q\right)_q. \end{aligned}$$

*Consequently, we have* $N_{\max}(R) = \prod_q \#\mathcal{B}_q(R)$.

So We can get the global bound from the local case. Following computing the local case. Now we will compute the bound of $\#\mathcal{B}_q(R)$.

At $q = p$, the division algebra $(B_{p,\infty})_p$ has a unique maximal order, so $\#\mathcal{B}_p(R) = 1$ [Voi21, Proposition 13.3.4 and Theorem 13.3.11(b)].

At $q \neq p$. Since $B_{p,\infty}$ is ramified only at $p$ and $\infty$, one has $B_q \simeq M_2(\mathbb{Q}_q)$. The maximal $\mathbb{Z}_q$-orders in $B_q$ are naturally identified with the vertices of the $(q+1)$-regular Bruhat–Tits tree of $\mathbf{PGL}_2(\mathbb{Q}_q)$; see [Voi21, Definition 23.5.5 and Proposition 23.5.8]. Under this identification, the set $\mathcal{B}_q(R)$ determines a convex subtree. Set $e_q = v_q(D), g_q = \min\{v_q(m), v_q(n)\}$.

If two vertices in $\mathcal{B}_q(R)$ are at distance $d$, their intersection is an Eichler order of level $q^d$. Since $R_q$ is contained in this intersection, then $d \leq v_q(\operatorname{discrd}(R)) = e_q$. Thus $\operatorname{diam} \mathcal{B}_q(R) \leq e_q$.

Now we only consider the case $n$ or $m$ is square-free. Suppose $m$ is square-free, let $A_i = \mathbb{Z}_q[i]$, $K_i = \mathbb{Q}_q[i]$. $\mathcal{O}_{K_i}$ be the maximal $\mathbb{Z}_q$-order. Then we have $A_i = \mathbb{Z}_q + q^{f_i}\mathcal{O}_{K_i}$, for some $f_i \geq 0$. Compute the discriminant on both sides, we have $(-4pm)\mathbb{Z}_q = \operatorname{disc}(A_i) = q^{2f_i} \operatorname{disc}(\mathcal{O}_{K_i})$. Hence $2f_i \leq v_q(4pm)$, and then

$$f_i = \begin{cases} 0, & q \neq 2, p, \\ 0, 1, & q = 2. \end{cases}$$

Since $A_i \subseteq R_q$, we have $\mathcal{B}_q(R) \subseteq \mathcal{B}_q(A_i)$. Let $S_i = \mathcal{B}_q(\mathcal{O}_{K_i})$, the standard branch description of quadratic orders [ACS16, §1] gives $\mathcal{B}_q(A_i) \subseteq N_{f_i}(S_i)$.

We now translate these geometric descriptions into a uniform vertex bound. Let $T$ be a subtree of the $(q+1)$-regular Bruhat–Tits tree satisfying $\operatorname{diam}(T) \leq e$, $T \subseteq N_f(S)$, where $S$ is one of the stems described above. Let $T$ be a subtree of the $(q+1)$-regular Bruhat–Tits tree satisfying $\operatorname{diam}(T) \leq e$, $T \subseteq N_f(S)$.

If $q \neq 2, p$, $A_i = \mathcal{O}_{K_i}$. Since $S$ is convex, the nearest-point projection $\pi_S : T \to S$ does not increase distances. Thus $\#\pi_S(T) \leq e_q + 1$. Moreover, each projection fiber is contained in a ball of radius $f$, which has at most 1 vertices. Therefore, $\#T \leq e_q + 1$. If $q = 2$,

$A_i = \mathcal{O}_{K_i}$ or $\mathbb{Z}_q + 2\mathcal{O}_{K_i}$, then each projection fiber at most $1 + (q+1)\sum_{r=0}^{f-1} q^r \le 3q^f$ vertices, so $\#T \le 3(e_2+1)q^f \le 6(e_2+1)$.

**Theorem 5.5.** *Let $1 \le m, n < p/4$, and suppose that $\mu(n)^2 = 1$. For each $t$ arising from an ordered distinct pair, let $D = 4mn - t^2 > 0$. Then*

$$C_{p,t}(m,n) \le 288\,\tau(D), \tag{5.14}$$

*where $\tau$ is the divisor function. Consequently,*

$$C_p(m,n) \le 288 \sum_{\substack{t\in\mathbb{Z}\\ t^2<4mn}} \tau(4mn - t^2). \tag{5.15}$$

*Proof.* From above statement, for a fixed triple $(m,n,t)$, we have

$$\#\mathcal{B}_q(R) \le e_q + 1 \quad (q \ne 2,p); \qquad \#\mathcal{B}_2(R) \le 6(e_2+1).$$

Using the local-global product formula Lemma 5.4, we obtain

$$\begin{aligned} N_{\max}(R) &= \prod_q \#\mathcal{B}_q(R) \\ &\le 6(e_2+1) \prod_{\substack{q\mid D\\ q\ne 2,p}} (e_q+1) \le 6\tau(D). \end{aligned} \tag{5.16}$$

Here, if $p \mid D$, the factor $e_p + 1$ occurring in $\tau(D)$ only enlarges the right-hand side, since $\#\mathcal{B}_p(R) = 1$.

Finally, combining (5.16) with (5.12) gives

$$C_{p,t}(m,n) \le 48 N_{\max}(R) \le 288\tau(D),$$

which proves (5.14). Summing over $t \in \mathbb{Z}$ with $t^2 < 4mn$ and using (5.5) yields

$$C_p(m,n) \le 288 \sum_{\substack{t\in\mathbb{Z}\\ t^2<4mn}} \tau(4mn - t^2).$$

□

**Corollary 5.6.** *Let $1 \le n < p/4$ be squarefree. Then*

$$\sum_{m\le n} C_p(m,n) \le 17280\, n^2\big(1 + \log(2n)\big) \tag{5.17}$$

*In particular, $\sum_{m\le n} C_p(m,n) = O\left(n^2 \log(2n)\right)$.*

*Proof.* Combine Theorem 5.5 with Lemma 2.7. □

5.3. **Bound of $\mathrm{E}(S_E)$ and $\mathrm{E}(P_E)$.** For $E \in \mathcal{S}_p$, consider the number of all selected-isogeny:

$$S_E = \sum_{n\in\mathcal{A}(X,B)} r_E^\circ(n). \tag{5.18}$$

On each curve $E$, order all primitive connection isogenies of degree at most $X$ by: $\varphi_1 <_{\text{order}} \varphi_2$ if $\deg(\varphi_1) < \deg(\varphi_2)$ or $\deg(\varphi_1) = \deg(\varphi_2)$ and $\varphi_1 <_{\text{dic}} \varphi_2$[3]. Let $P_E$ be the number of ordered pairs $(w,z)$ such that

- $z$ is selected, meaning $\deg z \in \mathcal{A}(X,B)$;
- $w \ne z$ is a primitive connection isogeny of degree at most $X$;

[3]Since only an arbitrary total order is required, fix $E \in \mathcal{S}_p$ and an integer $n$. The set of primitive connection-isogeny of degree $n$ is finite, so we may list its elements as $[\varphi_{n,1}], \dots, [\varphi_{n,N_{E,n}}]$. We then declare $[\varphi_{n,i}] <_{\text{dic}} [\varphi_{n,j}] \iff i < j$. This ordering is arbitrary, but we hold it fixed for the rest of the argument.

- $w <_{\text{order}} z$.

Equivalently,

$$P_E = \sum_{n\in\mathcal{A}(X,B)} \left( r_E^\circ(n) \sum_{m<n} r_E^\circ(m) + \binom{r_E^\circ(n)}{2} \right). \tag{5.19}$$

Let $0 < \gamma \le 1$ be a fixed constant and set $\lambda = \gamma(\log p)^{-2}$, $X = \lfloor N\gamma^{2/3}(\log p)^{-4/3} \rfloor$ and $v = \log X/\log B$.

**Proposition 5.7.** *For every $0 < \varepsilon < 1/2$, there exists a constant $a > 0$ such that for all sufficiently large $p$, uniformly for $B$ satisfying $(\log \frac{p}{2})^\varepsilon < u < (\log \frac{p}{2})^{1-\varepsilon}$,*

$$\mathrm{E}S_E \ge \frac{a\lambda\rho(v)}{\log p} = \frac{a\gamma\rho(v)}{\log^3 p} \ge \frac{a\lambda\rho(u)}{\log p} = \frac{a\gamma\rho(u)}{\log^3 p}. \tag{5.20}$$

*Proof.* By Proposition 3.6 and Lemma 2.5, if $pn \equiv 3 \mod 4$,

$$A_p^{\mathrm{BH}}(n) = \frac{\sqrt{pn}}{\pi} L(1,\chi_{-4pn}) + \frac{\sqrt{pn}}{2\pi} L(1,\chi_{-pn}) > \frac{0.655e^{-2}\sqrt{pn}}{\pi\log p};$$

Otherwise

$$A_p^{\mathrm{BH}}(n) = \frac{\sqrt{pn}}{\pi} L(1,\chi_{-4pn}) > \frac{0.655e^{-2}}{\pi\log p}\sqrt{pn}.$$

As a consequence, outside at most one $n \in \mathcal{A}(X,B)$,

$$A_p^{\mathrm{BH}}(n) > \frac{0.655e^{-2}}{\pi\log p}\sqrt{pn}.$$

Since $\log X = \frac{\log\frac{p}{2}}{3} + O(\log\log p)$, $\log B = \frac{\log\frac{p}{2}}{3u}$, we have $v := \frac{\log X}{\log B} = u(1+o_\varepsilon(1)) \to \infty$ uniformly for $(\log\frac{p}{2})^\varepsilon < u < (\log\frac{p}{2})^{1-\varepsilon}$ as $p\to\infty$. Moreover, $\log B \ge \frac{(\log\frac{p}{2})^\varepsilon}{3}$, $\log\log X = \log\log\frac{p}{2} + O(1)$. Hence, for sufficiently large $p$, $(\log X)^3 \le B \le X$ and, for any fixed $\delta > 0$, $\log B \ge (\log\log X)^{5/3+\delta}$. Thus Lemma 2.2 implies that $\#\mathcal{A}(X,B) \asymp X\rho(v)$.

The deletion of one element is negligible because $\log(X\rho(v)) = \frac{\log\frac{p}{2}}{3} - o_\varepsilon(\log\frac{p}{2}) \longrightarrow \infty$, using $v \le (\log\frac{p}{2})^{1-\varepsilon}(1+o(1))$. We count degrees in $\mathcal{A}(X,B)$, so it satisfies $n > X/2$. Hence there exits a constant $C_{lower}$

$$\mathrm{E}S_E = \frac{1}{h_p} \sum_{n\in\mathcal{A}(X,B)} A_p^{\mathrm{BH}}(n) \ge C_{lower} \frac{\sqrt{pX}}{p\log p} X\rho(v) = C_{lower} \frac{X^{3/2}}{\sqrt{p}} \frac{\rho(v)}{\log p}.$$

Since $\rho$ is nonincreasing, the inequality $v \le u$ also gives

$$\mathrm{E}S_E \ge C_{lower} \frac{X^{3/2}}{\sqrt{p}} \frac{\rho(v)}{\log p} \ge C_{lower} \frac{X^{3/2}}{\sqrt{p}} \frac{\rho(u)}{\log p}. \tag{5.21}$$

Since $N = N_0 + O(1)$ and $X = N\lambda^{2/3} + O(1)$,

$$X^{3/2} = N_0^{3/2}\lambda(1+o_\varepsilon(1)) = \frac{\sqrt{p}}{\sqrt{2}}\lambda(1+o_\varepsilon(1)), \tag{5.22}$$

Applying (5.22) to (5.21) proves that

$$\mathrm{E}S_E \ge \frac{a\lambda\rho(v)}{\log p} = \frac{a\gamma\rho(v)}{\log^3 p} \ge a\frac{\lambda\rho(u)}{\log p} = \frac{a\gamma\rho(u)}{\log^3 p}. \qquad \square$$

**Proposition 5.8.** *Let $0 < \gamma \le 1$ be a fixed constant and set $X = \lfloor N\gamma^{2/3}(\log p)^{-4/3} \rfloor$. Let $v = \frac{\log X}{\log B}$. For every $0 < \varepsilon < 1/2$, there is a constant $b > 0$ such that for all sufficiently large $p$, uniformly for $B$ satisfying $(\log\frac{p}{2})^\varepsilon < u < (\log\frac{p}{2})^{1-\varepsilon}$,*

$$\mathrm{E}P_E \le b\gamma^2\rho(v)\log^{-3} p. \tag{5.23}$$

*Consequently,*

$$\frac{\mathrm{E}P_E}{\mathrm{E}S_E} \le \frac{b}{a}\gamma. \tag{5.24}$$

*Proof.* Summing (5.19) over $E$ and using (5.1),

$$\begin{aligned}\sum_{E\in\mathcal{S}_p} P_E &= \sum_{n\in\mathcal{A}(X,B)} \left(\sum_{m<n} C_p(m,n) + \frac{1}{2}C_p(n,n)\right)\\ &\le \sum_{n\in\mathcal{A}(X,B)} \sum_{m\le n} C_p(m,n).\end{aligned}$$

Every $n \in \mathcal{A}(X,B)$ is squarefree, so Corollary 5.6 gives

$$\begin{aligned}\sum_{E\in\mathcal{S}_p} P_E &\le 17280 \sum_{n\in\mathcal{A}(X,B)} n^2\big(1+\log(2n)\big)\\ &\le 17280X^2\big(1+\log(2X)\big)\#\mathcal{A}(X,B)\\ &\le C_{P_E}X^3\rho(v)\log p,\end{aligned}$$

where $C_{P_E} > 0$ is an absolute constant. The last line uses (2.6) and $X < p$. Since $h_p \asymp p$ and $X^3/p \asymp \lambda^2$, averaging proves (5.23). Dividing by (5.20) proves (5.24). □

5.4. **Proof of the lower bound.** We derive a lower bound for $\Pr(P^+(\delta_E) \le B)$ from estimates for $\mathrm{E}[S_E]$ and $\mathrm{E}[P_E]$.

**Lemma 5.9.** *For every $E \in \mathcal{S}_p$,*

$$\mathbf{1}_{P^+(\delta)\le B} \ge S_E - P_E. \tag{5.25}$$

*Proof.* If there is no primitive root of degree at most $X$, then $S_E = P_E = 0$. Otherwise, let $w_{\min}$ be the first root in the chosen ordering. Its degree is $\delta_E$, since a shortest connection is primitive.

If $w_{\min}$ is not selected, every selected root has $w_{\min}$ as a distinct predecessor. Hence $P_E \ge S_E$, and therefore $S_E - P_E \le 0 \le \mathbf{1}_{P^+(\delta_E)\le B}$.

If $w_{\min}$ is selected, then $\mathbf{1}_{P^+(\delta_E)} = 1$. Aside from the shortest isogeny $w_{min}$, every other selected connection isogeny can be paired with $w_{min}$ in $P_E$. Hence $P_E \ge S_E - 1$, which again gives $\mathbf{1}_{P^+(\delta_E)} \ge S_E - P_E$. □

**Proposition 5.10.** *For every fixed $0 < \varepsilon < 1/2$, there is an absolute constant $c_1 > 0$ such that, for all sufficiently large $p$ and all $B$ satisfying*

$$(\log\frac{p}{2})^{\varepsilon} < u < (\log\frac{p}{2})^{1-\varepsilon},$$

*one has*

$$\Pr\big(P^+(\delta_E) \le B\big) \ge \frac{c_1\rho(u)}{(\log p)^3}. \tag{5.26}$$

*Proof.* Let $\gamma = \min\{1, \frac{a}{2b}\}$, where $a$ and $b$ are the constants in Propositions 5.7 and 5.8. Then (5.24) gives for sufficiently large $p$,

$$\mathrm{E}P_E \le \frac{1}{2}\mathrm{E}S_E.$$

Taking expectations in (5.25) and applying (5.20),

$$\Pr\big(P^+(\delta_E) \le B\big) = \mathrm{E}[\mathbf{1}_{P^+\delta_E\le B}] \ge \mathrm{E}S_E - \mathrm{E}P_E \ge \frac{1}{2}\mathrm{E}S_E \ge \frac{a\gamma}{2}\frac{\rho(v)}{(\log p)^3}.$$

Because $v \le u$ and $\rho$ is decreasing, $\rho(v) \ge \rho(u)$. Taking $c_1 = \frac{a\gamma}{2}$ proves (5.26). □

## 6. Probabilistic Consequences and Cryptanalytic Applications

### 6.1. Proof of Wesolowski's heuristic.

**Corollary 6.1.** *Heuristic W is true.*

*Proof.* Taking logarithms in (1.1) gives

$$\log \rho(u) + \log c_1 - 3\log\log p \le \log \Pr\big(P^+(\delta_E) < B\big) \le \log \rho(u) + \log c_2 + \log\log p. \quad (6.1)$$

Since $u > \log(p/2)^{\varepsilon}$, we have:

$$\frac{\log\log p}{u \log u} < \frac{\log\log p}{\log(p/2)^{\varepsilon}} \longrightarrow 0 \text{ as } p \to \infty.$$

So for any constant $C$ there exists $p_0$, for all $p > p_0$

$$\frac{\log\log p}{u \log u} < C.$$

Combining this with (2.3) proves what we want. □

By Corollary 6.1, Wesolowski's heuristic follows from our result. Consequently, the $p^{1/3+o(1)}$ attack no longer relies on this heuristic assumption and admits an unconditional justification.

### 6.2. A higher success probability.

In the preceding analysis, $P_E$ was used to restrict attention to connecting isogenies of minimal degree $\delta_E$. The $p^{1/3}$-attack of [Wes26], however, does not require the connecting isogeny recovered by the attack to have minimal degree. The attack first enumerates all isogenies emanating from $E$ whose degrees are $B$-smooth and at most $B^{1/2}\left(\frac{p}{2}\right)^{1/6}$.

Consequently, the attack also succeeds whenever there exists a connecting isogeny $\varphi\colon E \longrightarrow E^{(p)}$ such that $\delta_E < \deg(\varphi) \le B\left(\frac{p}{2}\right)^{1/3}$ and $P^+(\deg(\varphi)) \le B$. Thus, the attack is not restricted to the case in which the minimal connecting degree $\delta_E$ is $B$-smooth.

More precisely, the relevant success event is not merely $P^+(\delta_E) \le B$, but rather

$$\left\{\exists\, \varphi\colon E \to E^{(p)} \ :\ P^+(\deg(\varphi)) \le B \text{ and } \deg(\varphi) \le B\left(\frac{p}{2}\right)^{1/3}\right\}.$$

Accordingly, the success probability is

$$\Pr\left(\exists\, \varphi\colon E \to E^{(p)} \ :\ P^+(\deg(\varphi)) \le B \text{ and } \deg(\varphi) \le B\left(\frac{p}{2}\right)^{1/3}\right).$$

The Frobenius-claw event considered in [Mam26] corresponds precisely to this broader event. Therefore, after disregarding finite $p$ deviations from the limiting distribution, the success probability of the claw attack is at least $\Pr\big(P^+(\delta_E) \le B\big)$, and may be strictly larger because the attack can also exploit non-minimal connecting isogenies.

To capture this larger success event, we may omit the restriction encoded by $P_E$ and work directly with $S_E$. Indeed, $S_E$ counts the connecting isogenies whose degrees are $B$-smooth and lie in the interval $[X/2, X]$. The previous choice of $\mathcal{A}(X,B)$ was tailored to facilitate the analysis of $P_E$. Since minimality is no longer required, we instead introduce the cumulative set

$$\mathcal{A}_{\le}(X,B) := \big\{n \in \mathbb{Z}_{\ge 2} : n \le X,\ \mu^2(n) = 1,\ P^+(n) \le B\big\}.$$

Repeating the preceding argument with $\mathcal{A}(X,B)$ replaced by $\mathcal{A}_{\le}(X,B)$ yields the following sharper lower bound for the success probability.

As a consequence, combining with Theorem 5.5, we have

$$C_p(m,n) = O\left(\sqrt{mn}(1+\log 2mn)^2\right) \quad (6.2)$$

Let $S_E^{\le} := \sum_{n \in \mathcal{A}_{\le}(X,B)} r_E^{\circ}(n)$.

**Proposition 6.2.** *Denote* $\Pr(A)$ *the probability of existing isogeny* $\varphi\colon E \to E^{(p)}$ *satisfying that* $\deg\varphi \le B(p/2)^{1/3}$ *and* $\deg\varphi$ *is* $B$*-smooth. Then*

$$\Pr(A) \ge \big(\frac{4\sqrt{2}\cdot 0.655e^{-2}}{\pi\zeta(2)} + o_\varepsilon(1)\big)\frac{\rho(u)}{\log p}. \tag{6.3}$$

*Proof.* Since $\mathbf{1}_{\{s\ge 1\}} \ge s - \binom{s}{2}$ is right for all integer,

$$\mathbb{P}(A) \ge \Pr(S_E^{\le} \ge 1) \ge \mathrm{E}[S_E^{\le}] - \mathrm{E}\left[\binom{S_E^{\le}}{2}\right].$$

Since every $n \in \mathcal{A}_{\le}(X,B)$ is squarefree, Theorem 5.5 and Lemma 2.6 give

$$C_p(m,n) = O\left(\sqrt{mn}\big(1+\log(2mn)\big)^2\right)$$

for all $m,n \in \mathcal{A}_{\le}(X,B)$. Let

$$R_{p,B,X} = \frac{\mathrm{E}\left[\binom{S_E^{\le}}{2}\right]}{\mathrm{E}[S_E^{\le}]}$$

If $R_{p,B,X} \to 0$ as $p\to\infty$, we have $\Pr(A) \ge (1-o(1))\mathrm{E}[S_E^{\ge}]$. Choose $X = N-1$ we have (6.3) .

Now we prove the limitation $R_{p,B,X}$ tends to 0:

$$\begin{aligned}
\mathrm{E}\left[\binom{S_E^{\le}}{2}\right] &= \frac{1}{h_p}\left(\sum_{\substack{m,n\in\mathcal{A}_{\le}(X,B)\\ m<n}} C_p(m,n) + \frac{1}{2}\sum_{n\in\mathcal{A}_{\le}(X,B)} C_p(n,n)\right)\\
&\le \frac{C_1(1+\log(2X))^2}{h_p}\left(\sum_{n\in\mathcal{A}_{\le}(X,B)}\sqrt{n}\right)^2\\
&\le C_2\frac{X^3}{p}\rho(v)^2(\log p)^2,
\end{aligned}$$

where $C_1, C_2 > 0$ are absolute constants.

On the other hand, since $\mathcal{A}(X,B) \subseteq \mathcal{A}_{\le}(X,B)$,

$$\begin{aligned}
\mathrm{E}[S_E^{\le}] &= \frac{1}{h_p}\sum_{n\in\mathcal{A}_{\le}(X,B)} A_p^{\mathrm{BH}}(n)\\
&\ge \frac{1}{h_p}\sum_{n\in\mathcal{A}(X,B)} A_p^{\mathrm{BH}}(n)\\
&\ge \frac{0.655e^{-2}}{\pi h_p\log p}\sqrt{\frac{pX}{2}}\big(\#\mathcal{A}(X,B)-1\big)\\
&\ge C_3\frac{X^{3/2}\rho(v)}{\sqrt{p}\log p},
\end{aligned}$$

where $C_3 > 0$ is an absolute constant. The last inequality follows from Lemma 2.2 and $h_p \asymp p$.

Combining with the upper bound of $\mathrm{E}\left[\binom{S_E^{\le}}{2}\right]$, we obtain

$$R_{p,B,X} = O\left(\frac{X^{3/2}}{\sqrt{p}}\,\rho(v)(\log p)^3\right). \tag{6.4}$$

Now choose $X = N-1$, where $N = \lfloor (p/2)^{1/3} \rfloor$. Then $\frac{X^{3/2}}{\sqrt{p}} \longrightarrow \frac{1}{\sqrt{2}}$. Since $\log \rho(v) = -(1+o(1))v \log v$ and $v \geq (\log \frac{p}{2})^{\varepsilon}/2$, we have $\rho(v)(\log p)^3 = o(1)$. Consequently, $R_{p,B,X} = o(1)$.

Therefore,

$$\begin{aligned}
\Pr(A) &\geq \Pr(S_{\overline{E}}^{\leqslant} \geq 1) \\
&\geq \mathrm{E}[S_{\overline{E}}^{\leqslant}] - \mathrm{E}\left[\binom{S_{\overline{E}}^{\leqslant}}{2}\right] \\
&= (1 - R_{p,B,X})\mathrm{E}[S_{\overline{E}}^{\leqslant}] \\
&\geq C_{\varepsilon} \frac{\rho(u)}{\log p},
\end{aligned}$$

for some constant $C_{\varepsilon} > 0$ depending only on $\varepsilon$. Here we used $\rho(v) \geq \rho(u)$, since $v \leq u$. □

As a result, the success probability satisfies $\Pr(A) \geq C_{\varepsilon} \frac{\rho(u)}{\log p}$, for some constant $C_{\varepsilon} > 0$ depending only on $\varepsilon$.

## References


[ACNL+23] Sarah Arpin, Catalina Camacho-Navarro, Kristin Lauter, Joelle Lim, Kristina Nelson, Travis Scholl, and Jana Sotáková, *Adventures in supersingularland*, Experimental Mathematics **32** (2023), no. 2, 241–268.

[ACS16] Luis Arenas-Carmona and Ignacio Saavedra, *On some branches of the Bruhat–Tits tree*, International Journal of Number Theory **12** (2016), no. 3, 813–831.

[AOV26a] Yves Aubry, Roger Oyono, and Christelle Vincent, *Minimal degree of an isogeny between a supersingular elliptic curve and its conjugate*, 2026.

[AOV26b] Yves Aubry, Roger Oyono, and Christelle Vincent, *WISDE*, 2026, GitHub repository, commit 67bb9897633f3a2c34ef1cc8e5cee269d6e49861.

[AV26] Yves Aubry and Christelle Vincent, *Inseparable endomorphisms and rank-2 sublattices of the Gross lattice*, 2026.

[BDD+24] Andrea Basso, Pierrick Dartois, Luca De Feo, Antonin Leroux, Luciano Maino, Giacomo Pope, Damien Robert, and Benjamin Wesolowski, *SQIsign2D-West: The fast, the small, and the safer*, Advances in Cryptology—ASIACRYPT 2024, Part III, LNCS, vol. 15486, Springer, 2024, pp. 339–370.

[BBD+24] Jeremy Booher, Ross Bowden, Javad Doliskani, Tako Boris Fouotsa, Steven D. Galbraith, Sabrina Kunzweiler, Simon-Philipp Merz, Christophe Petit, Benjamin Smith, Katherine E. Stange, Yan Bo Ti, Christelle Vincent, José Felipe Voloch, Charlotte Weitkämper, and Lukas Zobernig, *Failing to hash into supersingular isogeny graphs*, The Computer Journal **67** (2024), no. 8, 2702–2719. https://doi.org/10.1093/comjnl/bxae038.

[BDD+26] Andrea Basso, Pierrick Dartois, Max Duparc, Jonathan Komada Eriksen, Sabrina Kunzweiler, Michael Meyer, Giacomo Pope, Krijn Reijnders, Damien Robert, Ryan Rueger, and Sina Schaeffler, *MIKE: A Fast and Compact Post-Quantum NIKE*, Cryptology ePrint Archive, Paper 2026/2377, 2026. https://eprint.iacr.org/2026/2377.

[BF23] Andrea Basso and Tako Boris Fouotsa, *New SIDH countermeasures for a more efficient key exchange*, Advances in Cryptology—ASIACRYPT 2023, Part VIII, LNCS, vol. 14445, Springer, 2023, pp. 208–233.

[BFLS20] Daniel J. Bernstein, Luca De Feo, Antonin Leroux, and Benjamin Smith, *Faster computation of isogenies of large prime degree*, ANTS XIV—Proceedings of the Fourteenth Algorithmic Number Theory Symposium (Steven D. Galbraith, ed.), Open Book Series, vol. 4, Mathematical Sciences Publishers, Berkeley, CA, 2020, pp. 39–55.

[BJS14] Jean-François Biasse, David Jao, and Anirudh Sankar, *A quantum algorithm for computing isogenies between supersingular elliptic curves*, Progress in Cryptology—INDOCRYPT 2014, Lecture Notes in Computer Science, vol. 8885, Springer, 2014, pp. 428–442.

[BM25] Andrea Basso and Luciano Maino, *POKÉ: A compact and efficient PKE from higher-dimensional isogenies*, Advances in Cryptology—EUROCRYPT 2025, Part II, LNCS, vol. 15602, Springer, 2025, pp. 94–123.

[BMP23] Andrea Basso, Luciano Maino, and Giacomo Pope, *FESTA: Fast encryption from supersingular torsion attacks*, Advances in Cryptology—ASIACRYPT 2023, Part VII, LNCS, vol. 14444, Springer, 2023, pp. 98–126.

[CD23] Wouter Castryck and Thomas Decru, *An efficient key recovery attack on SIDH*, Advances in Cryptology—EUROCRYPT 2023, Part V, LNCS, vol. 14008, Springer, 2023, pp. 423–447.

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

[Cox13] David A. Cox, *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication*, 2 ed., Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts, John Wiley & Sons, Hoboken, NJ, 2013.

[CS22] Mathilde Chenu and Benjamin Smith, *Higher-degree supersingular group actions*, Mathematical Cryptology **1** (2022), no. 2, 85–101.

[Dav00] Harold Davenport, *Multiplicative number theory*, 3rd ed., Graduate Texts in Mathematics, vol. 74, Springer, New York, 2000, Revised by Hugh L. Montgomery.

[dB51] N. G. de Bruijn, *The asymptotic behaviour of a function occurring in the theory of primes*, Journal of the Indian Mathematical Society. New Series **15** (1951), 25–32.

[Del26] José Luis Delgado, *The supersingular isogeny problem in time and memory $p^{1/3+o(1)}$, unconditionally*, 2026.

[Deu41] Max Deuring, *Die typen der multiplikatorenringe elliptischer funktionenkörper*, Abhandlungen aus dem Mathematischen Seminar der Hansischen Universität **14** (1941), no. 1, 197–272.

[DG16] Christina Delfs and Steven D. Galbraith, *Computing isogenies between supersingular elliptic curves over $\mathbb{F}_p$*, Designs, Codes and Cryptography **78** (2016), no. 2, 425–440.

[Dic30] Karl Dickman, *On the frequency of numbers containing prime factors of a certain relative magnitude*, Arkiv för Matematik, Astronomi och Fysik **22A** (1930), no. 10, 1–14.

[DKL+20] Luca De Feo, David Kohel, Antonin Leroux, Christophe Petit, and Benjamin Wesolowski, *SQISign: Compact post-quantum signatures from quaternions and isogenies*, Advances in Cryptology—ASIACRYPT 2020, Part I, LNCS, vol. 12491, Springer, 2020, pp. 64–93.

[DLRW24] Pierrick Dartois, Antonin Leroux, Damien Robert, and Benjamin Wesolowski, *SQIsignHD: New dimensions in cryptography*, Advances in Cryptology—EUROCRYPT 2024, Part I, LNCS, vol. 14651, Springer, 2024, pp. 3–32.

[EHL+20] Kirsten Eisenträger, Sean Hallgren, Chris Leonardi, Travis Morrison, and Jennifer Park, *Computing endomorphism rings of supersingular elliptic curves and connections to path-finding in isogeny graphs*, ANTS XIV—Proceedings of the Fourteenth Algorithmic Number Theory Symposium (Berkeley, CA) (Steven D. Galbraith, ed.), Open Book Series, vol. 4, Mathematical Sciences Publishers, 2020, pp. 215–232.

[Feo17] Luca De Feo, *Mathematics of isogeny based cryptography*, 2017.

[FIK+25] Jenny Fuselier, Annamaria Iezzi, Mark Kozek, Travis Morrison, and Changningphaabi Namoijam, *Computing supersingular endomorphism rings using inseparable endomorphisms*, Journal of Algebra **668** (2025), 145–189, Version 2, 30 January 2025.

[FMP23] Tako Boris Fouotsa, Tomoki Moriya, and Christophe Petit, *M-SIDH and MD-SIDH: Countering SIDH attacks by masking information*, Advances in Cryptology—EUROCRYPT 2023, Part V, LNCS, vol. 14008, Springer, 2023, pp. 282–309.

[Hil86] Adolf Hildebrand, *On the number of positive integers $\leq x$ and free of prime factors $> y$*, Journal of Number Theory **22** (1986), no. 3, 289–307.

[HT86] Adolf Hildebrand and Gérald Tenenbaum, *On integers free of large prime factors*, Transactions of the American Mathematical Society **296** (1986), no. 1, 265–290.

[IT86] Aleksandar Ivić and Gérald Tenenbaum, *Local densities over integers free of large prime factors*, The Quarterly Journal of Mathematics **37** (1986), no. 4, 401–417.

[Mam26] Maher Mamah, *Complexity analysis and security implications of the new isogeny-path algorithm*, Cryptology ePrint Archive, Paper 2026/1821, 2026.

[MMP+23] Luciano Maino, Chloe Martindale, Lorenz Panny, Giacomo Pope, and Benjamin Wesolowski, *A direct key recovery attack on SIDH*, Advances in Cryptology—EUROCRYPT 2023, Part V, LNCS, vol. 14008, Springer, 2023, pp. 448–471.

[NIST26] Dustin Moody et al., *Status Report on the Second Round of the Additional Digital Signature Schemes for the NIST Post-Quantum Cryptography Standardization Process*, NIST Interagency/Internal Report 8610, National Institute of Standards and Technology, May 2026. doi:10.6028/NIST.IR.8610.

[NO24] Kohei Nakagawa and Hiroshi Onuki, *QFESTA: Efficient algorithms and parameters for FESTA using quaternion algebras*, Advances in Cryptology—CRYPTO 2024, Part V, LNCS, vol. 14924, Springer, 2024, pp. 75–106.

[NO26] Kohei Nakagawa and Hiroshi Onuki, *SQIsign2DPush: Faster signature scheme using 2-dimensional isogenies*, Advances in Cryptology—EUROCRYPT 2026, Part IV, LNCS, vol. 16544, Springer, 2026, pp. 511–540.

[NOC+24] Kohei Nakagawa, Hiroshi Onuki, Wouter Castryck, Mingjie Chen, Riccardo Invernizzi, Gioella Lorenzon, and Frederik Vercauteren, *SQIsign2D-East: A new signature scheme using 2-dimensional isogenies*, Advances in Cryptology—ASIACRYPT 2024, Part III, LNCS, vol. 15486, Springer, 2024, pp. 272–303.

[Pan26] Lorenz Panny, *Proof-of-concept implementation of benjamin wesolowski's 2026 attack against the OneEnd problem with time complexity* $p^{1/3+o(1)}$, Python/SageMath script, 2026.

[Pha26] Lam L. Pham, *Small degree isogenies between conjugate principally polarized superspecial abelian surfaces*, 2026.

[Rob23] Damien Robert, *Breaking SIDH in polynomial time*, Advances in Cryptology—EUROCRYPT 2023, Part V, LNCS, vol. 14008, Springer, 2023, pp. 472–503.

[Sil09] Joseph H. Silverman, *The arithmetic of elliptic curves*, 2 ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009.

[Sie35] Carl Ludwig Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arithmetica **1** (1935), no. 1, 83–86.

[Swa26] Nicolas Swanson, *Supersingular elliptic curves without inseparable small degree endomorphisms*, 2026.

[Tat51] Tikao Tatuzawa, *On a theorem of Siegel*, Japanese Journal of Mathematics **21** (1951), 163–178, Published in 1952.

[Vél71] Jacques Vélu, *Isogénies entre courbes elliptiques*, C. R. Acad. Sci. Paris Sér. A-B **273** (1971), A238–A241 (French). MR 0294345

[Voi21] John Voight, *Quaternion algebras*, Graduate Texts in Mathematics, vol. 288, Springer, Cham, 2021, Post-publication version 1.0.7u used for section and theorem numbering.

[Wes22] Benjamin Wesolowski, *The supersingular isogeny path and endomorphism ring problems are equivalent*, 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 2022, pp. 1100–1111. doi:10.1109/FOCS52979.2021.00109.

[Wes26] Benjamin Wesolowski, *The supersingular isogeny problem in time and memory* $p^{1/3+o(1)}$, Cryptology ePrint Archive, Paper 2026/1486, 2026.

[XLZO25] Zheng Xu, Kaizhan Lin, Chang-An Zhao, and Yi Ouyang, *SQIsign2D*$^2$*: New SQIsign2D variant by leveraging power smooth isogenies in dimension one*, Advances in Cryptology—ASIACRYPT 2025, Part IV, LNCS, vol. 16248, Springer, 2025, pp. 341–371.

## Appendix A. Proof of the lemmas in Section 2.

*Proof of Lemma 2.1.* Hildebrand's uniform estimate[Hil86, Theorem 1, pp. 289–290] states that for each fixed $\eta > 0$,

$$\frac{\Psi(x,y)}{x} = \rho(v)\left\{1 + O_\eta\left(\frac{\log(v+1)}{\log y}\right)\right\}$$

where $v = \frac{\log x}{\log y}$, whenever $2 \le y \le x$ and $\log y \ge (\log\log x)^{5/3+\eta}$.

We apply this result with $x = N_0, y = B, v = u, \eta = \frac{\varepsilon}{2}$. To apply correctly, we first prove $\log B \ge (\log\log N_0)^{5/3+\eta}$. Since $\log\log N_0 = \log\log p + O(1)$, and $\log B = \frac{\log N_0}{u} \ge C_0 \frac{\log p}{(\log p)^{1-\varepsilon}} = C_0(\log p)^\varepsilon$ for some $C_0$, we have that there exists constant $C_0'$

$$\frac{\log B}{(\log\log N_0)^{5/3+\varepsilon/2}} \ge \frac{C_0(\log p)^\varepsilon}{(\log\log N_0)^{5/3+\varepsilon/2}} = \frac{C_0'(\log p)^\varepsilon}{(\log\log p)^{5/3+\varepsilon/2}} \longrightarrow \infty \text{ as } p \to \infty$$

Therefore, we apply this result and obtain

$$\frac{\Psi(N_0,B)}{N_0} = \rho(u)\left\{1 + O_\varepsilon\left(\frac{\log(u+1)}{\log B}\right)\right\}.$$

Since $N = \lfloor N_0 \rfloor$ we have $\Psi(N,B) = \Psi(N_0,B)$. Then we obtain

$$\frac{\Psi(N,B)}{N} = \frac{N_0}{N}\frac{\Psi(N_0,B)}{N_0} = \rho(u)\left\{1 + O_\varepsilon\left(\frac{\log(u+1)}{\log B} + \frac{1}{N}\right)\right\}.$$

Using the upper bound $u < (\log p)^{1-\varepsilon}$ and $\log B \ge C_0(\log p)^\varepsilon$, we obtain as $p \to \infty$:

$$\frac{\log(u+1)}{\log B} \le \frac{(1-\varepsilon)\log\log p}{C_0(\log p)^\varepsilon} \longrightarrow 0.$$

□

*Proof of Lemma 2.2.* For counting $\mathcal{A}(X,B)$, we construct the following $\Psi_1$ counting $B$-smooth squarefree integers in $[1,x]$:

$$\Psi_1(x,B;\mu^2) = \sum_{\substack{n \le x \\ P^+(n) \le B}} \mu(n)^2.$$

Then $\#\mathcal{A}(X,B) = \Psi_1(X,B;\mu^2) - \Psi_1(\frac{X}{2},B;\mu^2)$.

From the theorem of Ivić and Tenenbaum[IT86, Theorem 1 and Corollary (2.12)] , we have

$$\Psi_1(x,B;\mu^2) = \zeta(2\beta(x,B))^{-1}\Psi(x,B)\left\{1 + O_\delta\left(\frac{\log\log B}{\log B}\right)\right\},$$

where $\beta(x,B) = 1 - \frac{\xi(v_x)}{\log B}$ with $v_x = \frac{\log x}{\log B}$ and $\xi(w)$ is determined by $e^{\xi(w)} = 1 + w\xi(w)$.

The standard estimate $\xi(w) = O(\log(w+1))$ gives $1 - \beta(x,B) = O\left(\frac{\log(v_x+1)}{\log B}\right)$. That is to say $\beta(x,B) = 1 + O\left(\frac{\log(v_x+1)}{\log B}\right)$. By definition $v_x = \frac{\log x}{\log B} \le \frac{\log X}{\log 2}$ and therefore $\log(v_x+1) = O(\log\log X)$. By condition $\log B \ge (\log\log X)^{5/3+\delta}$, we have

$$\frac{\log(v_x+1)}{\log B} = O\left((\log\log X)^{-2/3-\delta}\right) = o(1).$$

Thus $\beta(x,B) \to 1$ uniformly in the stated parameter range. We only need the value of $\Psi_1$ at $x = X$ and $x = \frac{X}{2}$, consider $x = X$ and $x = \frac{X}{2}$ specially, we have $\beta(X,B) \to 1$ and $\beta(\frac{X}{2},B) \to 1$.

Hence, we obtain

$$\Psi_1(X,B;\mu^2) = \frac{1}{\zeta(2)}\Psi(X,B); \quad \Psi_1\left(\frac{X}{2},B;\mu^2\right) = \frac{1}{\zeta(2)}\Psi\left(\frac{X}{2},B\right).$$

Moreover, the theorem of Hildebrand and Tenenbaum[HT86, Theorem 3] yields

$$\frac{\Psi(X/2,B)}{\Psi(X,B)} = 2^{-\alpha(X/2,B)}\left\{1+O\left(\frac{1}{v}+\frac{\log B}{B}\right)\right\} = 2^{-\alpha(X/2,B)}\left\{1+o_\delta(1)\right\}, \tag{A.1}$$

where $\alpha(x,B)$ denote the unique positive solution of

$$\sum_{\substack{\ell\le B\\ \ell \text{ prime}}} \frac{\log \ell}{\ell^{\alpha(x,B)}-1} = \log x. \tag{A.2}$$

The saddle-point estimate[HT86, Lemma 2, Equation (3.5)] gives $1-\alpha(x,B) = O\left(\frac{\log(v_x+1)}{\log B}\right)$. So $\alpha(x,B)\to 1$ uniformly as well. As a consequence,

$$\frac{\Psi(X/2,B)}{\Psi(X,B)} = \frac{1}{2}\left\{1+O\left(\frac{1}{v}+\frac{\log B}{B}\right)\right\} = \frac{1}{2}\{1+o_\delta(1)\}.$$

Therefore

$$\#\mathcal{A}(X,B) = \Psi_1(X,B;\mu^2) - \Psi_1\left(\frac{X}{2},B;\mu\right) = \left(\frac{1}{2\zeta(2)}+o_\delta(1)\right)\Psi(X,B).$$

Lemma 2.1 applies and gives

$$\Psi(X,B) = X\rho(v)\big(1+o_\delta(1)\big),$$

which complete the proof. □

*Proof of Lemma 2.5.* Since $n$ is squarefree and $n<p$ for $n\in\mathcal{A}(X,B)$, $X<p/4$, the integer $pn$ is squarefree. The quadratic discriminant formula therefore gives (2.10). The map $n\mapsto D_n$ is injective: injectivity within each case is immediate, while equality between the two cases would require one squarefree integer to be four times another.

Since $n>1$ and $X<p/4$, we have

$$p<|D_n|\le 4pX<p^2. \tag{A.3}$$

For any fixed $0<\eta<1/2$, Tatuzawa's theorem [Tat51, Theorem 2] states that, apart from at most one exceptional fundamental discriminant $D<0$, every primitive real Dirichlet character $\chi_D$ with $|D|>\max\{e^{1/\eta},e^{11.2}\}$ satisfies $L(1,\chi_D)>0.655\,\eta|D|^{-\eta}$.

Set $\eta=1/\log p$ and assume $p>e^{11.2}$. By (A.3), $|D_n|>p=e^{1/\eta}>e^{11.2}$. Hence, apart from at most one exceptional discriminant,

$$L(1,\chi_{D_n})>0.655\,\eta|D_n|^{-\eta}>\frac{0.655e^{-2}}{\log p}.$$

Since the map $n\mapsto D_n$ is injective, the exceptional discriminant corresponds to at most one $n\in\mathcal{A}(X,B)$. This proves (2.11).

If $pn\equiv 3 \pmod 4$, then $D_n=-pn$, and the character $\chi_{-4pn}=\chi_{4D_n}$ is induced from $\chi_{D_n}$. Hence

$$L(1,\chi_{-4pn}) = L(1,\chi_{D_n})\left(1-\frac{\chi_{D_n}(2)}{2}\right)\ge\frac{1}{2}L(1,\chi_{D_n})>\frac{0.655e^{-2}}{2\log p},$$

as required. □

*Proof of Lemma 2.6.* For any positive integer $k$, pairing each divisor $d\mid k$ with $k/d$ yields

$$\tau(k)\le 2\sum_{\substack{d\mid k\\ d\le\sqrt{k}}} 1.$$

When $t^2 < 4M$, we have $0 < 4M - t^2 \le 4M$. We may therefore widen the divisor range and interchange the order of summation to obtain

$$\begin{aligned}\sum_{\substack{t\in\mathbb{Z}\\ t^2<4M}} \tau(4M-t^2) &\le 2\sum_{\substack{t\in\mathbb{Z}\\ t^2<4M}} \sum_{\substack{d\mid(4M-t^2)\\ d\le\sqrt{4M}}} 1 \\ &= 2\sum_{d\le\sqrt{4M}} \#\left\{t\in\mathbb{Z} : |t|<\sqrt{4M},\ t^2\equiv 4M \pmod d\right\}. \end{aligned} \tag{A.4}$$

For each positive integer $d$, let

$$\nu_{4M}(d) := \#\{x \bmod d : x^2 \equiv 4M \pmod d\}.$$

Each residue class modulo $d$ contains at most $2\sqrt{4M}/d + 1$ integers in the interval $[-\sqrt{4M}, \sqrt{4M}]$. Since $d \le \sqrt{4M}$, we have

$$\begin{aligned}&\#\left\{t\in\mathbb{Z} : |t|<\sqrt{4M},\ t^2\equiv 4M \pmod d\right\} \\ &\quad\le \left(\frac{2\sqrt{4M}}{d}+1\right)\nu_{4M}(d) \le \frac{3\sqrt{4M}}{d}\nu_{4M}(d).\end{aligned}$$

Consequently, by (A.4),

$$\sum_{\substack{t\in\mathbb{Z}\\ t^2<4M}} \tau(4M-t^2) \le 6\sqrt{4M}\sum_{d\le\sqrt{4M}} \frac{\nu_{4M}(d)}{d}. \tag{A.5}$$

We next estimate the sum on the right hand side. By the Chinese remainder theorem, $\nu_{4M}$ is multiplicative. We therefore first derive bounds for $\nu_{4M}(q^j)$ for prime $q$, and then assemble these local estimates via an Euler product.

Fix a prime $q$, and write $e = v_q(4M)$, i.e. $4M = q^e a$ where $q \nmid a$. We estimate $\nu_{4M}(q^j)$ for each $j \ge 1$.

First, suppose that $1 \le j \le e$. Since $q^j \mid 4M$, the congruence $x^2 \equiv 4M \pmod{q^j}$ reduces to $x^2 \equiv 0 \pmod{q^j}$, which is equivalent to $q^{\lceil j/2\rceil} \mid x$. Consequently, $\nu_{4M}(q^j) = q^{\lfloor j/2\rfloor}$.

Now suppose that $j > e$. If $e$ is odd, the congruence $x^2 \equiv q^e a \pmod{q^j}$ forces $v_q(x^2) = e$, it is impossible because $v_q(x^2) = 2v_q(x)$ is even. Hence $\nu_{4M}(q^j) = 0$ if $j > e$ and $e$ is odd.

It remains to treat the case $e = 2k$. Write $j = 2k + r$, where $r \ge 1$. The congruence $x^2 \equiv q^{2k}a \pmod{q^{2k+r}}$ implies $q^k \mid x$. Setting $x = q^k y$, we obtain $y^2 \equiv a \pmod{q^r}$. Thus $\nu_{4M}(q^{2k+r}) = q^k\nu_a(q^r)$, where $\nu_a(d) = \#\{y \bmod d : y^2 \equiv a \bmod d\}$

We next bound $\nu_a(q^r)$. Since $q \nmid a$, the polynomial $x^2 - a$ has at most two roots in $\mathbb{F}_q$. For odd $q$, every such root is simple, so Hensel's lemma gives $\nu_a(q^r) \le 2$ when $q$ odd and $r \ge 1$. Therefore,

$$\sum_{r\ge1}\frac{\nu_a(q^r)}{q^r} \le 2\sum_{r\ge1} q^{-r} = \frac{2}{q-1}.$$

For $q = 2$, since $a$ is odd, we have $\nu_a(2) \le 1$ and $\nu_a(4) \le 2$. When $r \ge 3$, the congruence $x^2 \equiv 1 \pmod{2^r}$ has at most four solutions. Indeed, $2^r \mid (x-1)(x+1)$ implies $2^{r-1} \mid (x-1)$ or $2^{r-1} \mid (x+1)$, because $\gcd(x-1, x+1) = 2$ for odd $x$. Thus $x \equiv \pm 1 \pmod{2^{r-1}}$, yielding at most four residue classes modulo $2^r$.

If $x_0$ solves $x^2 \equiv a \pmod{2^r}$, then the congruence is equivalent to $(x_0^{-1}x)^2 \equiv 1 \pmod{2^r}$. Consequently, $\nu_a(2^r) \le 4$ when $r \ge 3$. It follows that

$$\sum_{r\ge1}\frac{\nu_a(2^r)}{2^r} \le \frac{1}{2} + \frac{2}{4} + 4\sum_{r\ge3} 2^{-r} = 2.$$

Combining cases above, we obtain

$$\sum_{r\geq 1}\frac{\nu_a(q^r)}{q^r}\leq\frac{2}{q-1} \tag{A.6}$$

for every prime $q$ with $q\nmid a$.

We put these estimates together. If $e=2k$, then

$$\begin{aligned}\sum_{j\geq 1}\frac{\nu_{4M}(q^j)}{q^j}&=2\sum_{h=1}^{k}\frac{1}{q^h}+q^{-k}\sum_{r\geq 1}\frac{\nu_a(q^r)}{q^r}\\&\leq\frac{2(1-q^{-k})}{q-1}+\frac{2q^{-k}}{q-1}=\frac{2}{q-1}.\end{aligned}$$

If $e=2k+1$, then

$$\begin{aligned}\sum_{j\geq 1}\frac{\nu_{4M}(q^j)}{q^j}&=2\sum_{h=1}^{k}\frac{1}{q^h}+\frac{1}{q^{k+1}}\\&\leq\frac{2}{q-1}.\end{aligned}$$

Therefore, uniformly in $M$ and $q$,

$$1+\sum_{j\geq 1}\frac{\nu_{4M}(q^j)}{q^j}\leq 1+\frac{2}{q-1}\leq(1-q^{-1})^{-2}. \tag{A.7}$$

By multiplicativity and nonnegativity of $\nu_{4M}$, we obtain

$$\begin{aligned}\sum_{d\leq\sqrt{4M}}\frac{\nu_{4M}(d)}{d}&=\sum_{d\leq\sqrt{4M}}\prod_{q\mid d}\frac{\nu_{4M}(q^{v_q(d)})}{q^{v_q(d)}}\\&\leq\prod_{q\leq\sqrt{4M}}\left(1+\sum_{j\geq 1}\frac{\nu_{4M}(q^j)}{q^j}\right)\\&\leq\prod_{q\leq\sqrt{4M}}(1-q^{-1})^{-2}\\&\leq C_{Mer}\big(1+\log\sqrt{4M}\big)^2,\end{aligned} \tag{A.8}$$

where $C_{Mer}>0$ is an absolute constant. The last inequality follows from Mertens' product estimate.

Substituting this bound into (A.5), we conclude that

$$\sum_{\substack{t\in\mathbb{Z}\\ t^2<4M}}\tau(4M-t^2)=O\left(\sqrt{M}\big(1+\log(2M)\big)^2\right),$$

with an absolute constant independent of $M$. □

*Proof of Lemma 2.7.* Put $Q(n)=\sum_{m\leq n}\sum_{t^2<4mn}\tau(4mn-t^2)$. For every positive integer $k$, there is

$$\tau(k)\leq 2\sum_{\substack{d\mid k\\ d\leq\sqrt{k}}}1.$$

Since $0<4mn-t^2\leq 4n^2$, every divisor occurring on the right is at most $2n$.

$$Q(n)\leq 2\sum_{m\leq n}\sum_{t^2<4mn}\sum_{d\mid 4mn-t^2\, d\leq\sqrt{4mn-t^2}}1\leq 2\sum_{d\leq 2n}N_d, \tag{A.9}$$

where $N_d = \#\big\{(m,t) : 1 \le m \le n,\ |t| < 2n,\ 4nm \equiv t^2 \pmod d\big\}$.

Fix $d$, and put $g = (4n, d)$. The linear congruence in $m$ is soluble if and only if $g \mid t^2$. Define

$$h(g) = \prod_{q^e \| g} q^{\lceil e/2 \rceil}.$$

Then $g \mid t^2$ if and only if $h(g) \mid t$. Since $g \mid 4n$ and $n$ is squarefree, every odd prime occurs in $g$ to exponent at most one and $v_2(g) \le 3$. Hence $\frac{g}{h(g)} \le 2$. It follows that the number of integers $t$ with $|t| < 2n$ and $g \mid t^2$ is at most $\frac{4n}{h(g)} + 1 \le \frac{8n}{g} + 1 \le \frac{10n}{g}$, where the last inequality uses $g \le d \le 2n$.

For each such $t$, division by $g$ gives $\frac{4n}{g} m \equiv \frac{t^2}{g} \pmod{d/g}$, and the coefficient $4n/g$ is invertible modulo $d/g$. Thus $m$ lies in one residue class modulo $d/g$, so there are at most $\frac{ng}{d} + 1$ choices with $1 \le m \le n$. Therefore

$$N_d \le \frac{10n}{g}\left(\frac{ng}{d} + 1\right) \le 30\frac{n^2}{d},$$

where we used $d \le 2n$ in the last step. Substitution into (A.9) gives

$$Q(n) \le 60n^2 \sum_{d \le 2n} \frac{1}{d} \le 60n^2\big(1 + \log(2n)\big).$$

□

## Appendix B. Mamah's heuristic

Mamah [Mam26] stated another heuristic in his analysis of the effectiveness of Wesolowski's algorithm, this heuristic originally commes from [BBD+24]:

**Heuristic M** (Heuristic 1, [Mam26])**.** *For a fixed degree $n$, the probability that a supersingular curve $E$ admits an $n$-isogeny to its Frobenius conjugate $E^{(p)}$ is*

$$\Pr(A_{E,n}) \approx \sqrt{\frac{12n}{p}}.$$

For $1 \le n < p/4$, define

$$\mathsf{A}_{E,n} := \{r_E > 0\},$$

where $E$ is sampled uniformly from $\mathcal{S}_p$. Then Heuristic M is

$$\Pr(\mathsf{A}_{E,n}) \approx \sqrt{\frac{12n}{p}}, \tag{B.1}$$

This heuristic also consider $m, n$ with square divisors, so we derive a non-squarefree bound for $C_p(m,n)$ as below:

**Lemma B.1.** *Let $p \ge 5$ be prime and let $1 \le m, n < p/4$. Put*

$$s(k) := \prod_{\substack{q \mid k \\ q \text{ prime}}} q^{\lfloor v_q(k)/2 \rfloor},$$

*so that $s(k)^2$ is the largest square divisor of $k$. Then*

$$C_p(m,n) = O\left(\sqrt{mn}\, s(\gcd(m,n))\big(1 + \log(2mn)\big)^2\right), \tag{B.2}$$

*where the implied constant is absolute. In particular,*

$$C_p(n,n) = O\left(n^{3/2}\big(1 + \log(2n)\big)^2\right).$$

*Proof.* Fix an integer $t$ arising from an ordered pair of distinct primitive connection-isogeny orbits of degrees $m$ and $n$, respectively, and let $R$ be the associated quaternion order. Put

$$D := 4mn - t^2 > 0, \qquad e_q := v_q(D), \qquad g_q := \min\{v_q(m), v_q(n)\}.$$

By Lemma 5.3, we bound $C_{p,t}(m,n)$ by counting $N_{\max}(R)$.

At $q = p$, the division algebra $(B_{p,\infty})_p$ has a unique maximal order, so $\#\mathcal{B}_p(R) = 1$ [Voi21, Proposition 13.3.4 and Theorem 13.3.11(b)].

Fix a finite prime $q \neq p$. Since $B_{p,\infty}$ is ramified only at $p$ and $\infty$, one has $B_q \simeq M_2(\mathbb{Q}_q)$. The maximal $\mathbb{Z}_q$-orders in $B_q$ are naturally identified with the vertices of the $(q+1)$-regular Bruhat–Tits tree of $\mathbf{PGL}_2(\mathbb{Q}_q)$; see [Voi21, Definition 23.5.5 and Proposition 23.5.8]. Under this identification, the set $\mathcal{B}_q(R)$ determines a convex subtree.

If two vertices in $\mathcal{B}_q(R)$ are at distance $d$, their intersection is an Eichler order of level $q^d$. Since $R_q$ is contained in this intersection, we have $d \leq v_q(\mathrm{discrd}(R)) = e_q$. Thus

$$\operatorname{diam} \mathcal{B}_q(R) \leq e_q.$$

In particular, if $e_q = 0$, then $\mathcal{B}_q(R)$ contains at most one vertex.

We next estimate the depth of $\mathcal{B}_q(R)$. Let $A_i = \mathbb{Z}_q[i]$, $K_i = \mathbb{Q}_q[i]$. Then $K_i$ is either a quadratic field extension of $\mathbb{Q}_q$ or the split algebra $\mathbb{Q}_q \times \mathbb{Q}_q$. Let $\mathcal{O}_{K_i}$ denote its maximal $\mathbb{Z}_q$-order. There exists an integer $f_i \geq 0$ such that $A_i = \mathbb{Z}_q + q^{f_i}\mathcal{O}_{K_i}$.

Since $i^2 = -pm$, the discriminant-conductor formula gives $(-4pm)\mathbb{Z}_q = \operatorname{disc}(A_i) = q^{2f_i}\operatorname{disc}(\mathcal{O}_{K_i})$. Hence

$$2f_i \leq v_q(4pm).$$

Similarly, for $A_j = \mathbb{Z}_q[j] = \mathbb{Z}_q + q^{f_j}\mathcal{O}_{K_j}$, where $j^2 = -pn$, we obtain $2f_j \leq v_q(4pn)$.

Set $f_q := \min\{f_i, f_j\}$. Since $p$ is prime integer, we obtain

$$f_q \leq \begin{cases} \lfloor g_q/2 \rfloor, & q \neq 2, p, \\ 1 + \lfloor g_2/2 \rfloor, & q = 2. \end{cases} \tag{B.3}$$

Since $A_i, A_j \subseteq R_q$, we have $\mathcal{B}_q(R) \subseteq \mathcal{B}_q(A_i) \cap \mathcal{B}_q(A_j)$. Let $S_i = \mathcal{B}_q(\mathcal{O}_{K_i})$ and $S_j = \mathcal{B}_q(\mathcal{O}_{K_j})$. The standard branch description of quadratic orders [ACS16, §1] gives

$$\mathcal{B}_q(A_i) \subseteq N_{f_i}(S_i), \qquad \mathcal{B}_q(A_j) \subseteq N_{f_j}(S_j),$$

where $N_f(S)$ denotes the closed $f$-neighborhood of $S$. Consequently, choosing the stem corresponding to the smaller conductor exponent $f_q$ to be $S$, we obtain

$$\mathcal{B}_q(R) \subseteq N_{f_q}(S).$$

The stem $S$ has one of three possible shapes: an apartment if the corresponding quadratic algebra is split, a single vertex if it is an unramified quadratic field extension, or a single edge if it is a ramified quadratic field extension; see [ACS16, §1].

We now translate these geometric descriptions into a uniform vertex bound. If $S$ is an apartment, put $\Lambda_q = S$. Otherwise, extend the vertex or edge $S$ to an apartment $\Lambda_q$. In either case, $S \subseteq \Lambda_q$, and therefore $\mathcal{B}_q(R) \subseteq N_{f_q}(S) \subseteq N_{f_q}(\Lambda_q)$.

Since nearest-point projection onto $\Lambda_q$ does not increase distances, the projection of $\mathcal{B}_q(R)$ contains at most $e_q + 1$ vertices. For each vertex of $\Lambda_q$, the corresponding projection fiber inside $N_{f_q}(\Lambda_q)$ contains

$$1 + (q-1)\sum_{r=0}^{f_q-1} q^r = q^{f_q}$$

vertices, with the sum interpreted as empty when $f_q = 0$. Hence

$$\#\mathcal{B}_q(R) \leq (e_q + 1)q^{f_q}.$$

Combining this with (B.3), we obtain

$$\#\mathcal{B}_q(R) \le (e_q+1)q^{\lfloor g_q/2\rfloor} \qquad (q \ne 2,p),$$

and

$$\#\mathcal{B}_2(R) \le 2(e_2+1)2^{\lfloor g_2/2\rfloor}.$$

Since $m,n<p$, we have $p \nmid \gcd(m,n)$. The local-global product correspondence yields

$$\begin{aligned} N_{\max}(R) &= \prod_q \#\mathcal{B}_q(R) \\ &\le 2\left(\prod_{q\ne p} q^{\lfloor g_q/2\rfloor}\right)\left(\prod_{q\ne p}(e_q+1)\right) \\ &\le 2s(\gcd(m,n))\tau(D). \end{aligned}$$

Combining this with Lemma 5.3, we obtain

$$\begin{aligned} C_{p,t}(m,n) &\le 48N_{\max}(R) \\ &\le 96s(\gcd(m,n))\tau(4mn-t^2). \end{aligned}$$

Summing over all admissible $t$, we have

$$C_p(m,n) \le 96s(\gcd(m,n)) \sum_{\substack{t\in\mathbb{Z}\\ t^2<4mn}} \tau(4mn-t^2).$$

Applying (2.13) with $M=mn$, we obtain

$$\sum_{\substack{t\in\mathbb{Z}\\ t^2<4mn}} \tau(4mn-t^2) = O\left(\sqrt{mn}\big(1+\log(2mn)\big)^2\right).$$

Consequently,

$$C_p(m,n) = O\left(\sqrt{mn}\, s(\gcd(m,n))\big(1+\log(2mn)\big)^2\right).$$

Finally, taking $m=n$ and using $s(n)\le\sqrt{n}$, we conclude that

$$C_p(n,n) = O\left(n^{3/2}\big(1+\log(2n)\big)^2\right). \qquad \square$$

Recall that the number of all connection isogenies bewteen $E$ and $E^{(p)}$ with degree $n$ $r_E(n) = \sum_{d^2\mid n} r_E^\circ(\frac{n}{d^2})$. By Proposition 3.6, we have

$$A_p^{BH}(n) = \sum_{E\in\mathcal{S}_p} r_E^\circ(n) = \frac{\sqrt{pn}}{\pi}L(1,\chi_{-4pn}) + \frac{\sqrt{pn}}{2\pi}\mathbf{1}_{\{pn\equiv 3 \pmod 4\}}L(1,\chi_{-pn})$$

Set $\mathcal{L}_{p,n} := L(1,\chi_{-4pn}) + \frac{1}{2}\mathbf{1}_{\{pn\equiv 3 \pmod 4\}}L(1,\chi_{-pn})$ and $\mathcal{H}_{p,n} := \sum_{d^2\mid n}\frac{\mathcal{L}_{p,n/d^2}}{d}$, we have that

$$\sum_{E\in\mathcal{S}_p} r_E^\circ = A_p^{BH} = \frac{\sqrt{np}}{\pi}\mathcal{L}_{n,p}, \quad \sum_{E\in\mathcal{S}_p} r_E(n) = \sum_{E\in\mathcal{S}_p}\sum_{d^2\mid n} r_E^\circ\left(\frac{n}{d^2}\right) = \frac{\sqrt{np}}{\pi}\mathcal{H}_{n,p}.$$

**Theorem B.2.** *Let $p\ge 5$ be prime and put $N := \lfloor (p/2)^{1/3}\rfloor$. For every $1\le n\le N$, we have Heuristic M holds in the logarithmic sense, i.e.*

$$\log \Pr(\mathsf{A}_{E,n}) = \frac{1}{2}\log\frac{n}{p} + O(\log\log p). \tag{B.4}$$

*Proof.* Averaging $r_E(n)$ over $\mathcal{S}_p$, we obtain

$$\begin{aligned}\mathbb{E}[r_E] &= \frac{1}{h_p}\sum_{d^2|n} A_p^{\mathrm{BH}}(n/d^2)\\ &= \frac{\sqrt{pn}}{\pi h_p}\sum_{d^2|n}\frac{\mathcal{L}_{p,n/d^2}}{d}\\ &= \frac{\sqrt{pn}}{\pi h_p}\mathcal{H}_{p,n}.\end{aligned} \tag{B.5}$$

We next bound the contribution from multiple connecting isogenies. For all $r \in \mathbb{Z}_{\geq 0}$,

$$0 \leq r - \mathbf{1}_{\{r>0\}} \leq \binom{r}{2}.$$

It gives

$$0 \leq \mathbb{E}[r_E] - \Pr(\mathsf{A}_{E,n}) \leq \mathbb{E}\left[\binom{r_E}{2}\right] = \frac{1}{2h_p}\sum_{\substack{d^2|n\\ e^2|n}} C_p\left(\frac{n}{d^2}, \frac{n}{e^2}\right). \tag{B.6}$$

By Lemma B.1,

$$C_p\left(\frac{n}{d^2}, \frac{n}{e^2}\right) = O\left(\frac{n}{de}\, s\left(\gcd\left(\frac{n}{d^2}, \frac{n}{e^2}\right)\right)\left(1+\log(2n)\right)^2\right).$$

Then we bound the sum on the right hand side: for $d^2, e^2 \mid n$, we have $s\left(\gcd\left(\frac{n}{d^2}, \frac{n}{e^2}\right)\right) = \frac{s(n)}{\operatorname{lcm}(d,e)}$. The sum $\sum_{\substack{e^2|n, d^2|n\\ e\geq 1, d\geq 1}} \frac{s(n)}{\operatorname{lcm}(d,e)}$ is uniformly bounded, since

$$\begin{aligned}\sum_{\substack{e^2|n, d^2|n\\ d,e\geq 1}} \frac{1}{de\operatorname{lcm}(d,e)} &\leq \sum_{d,e\geq 1}\frac{1}{de\operatorname{lcm}(d,e)}\\ &= \sum_{d,e\geq 1}\frac{\gcd(d,e)}{d^2e^2}\\ &= \zeta(2)^2\sum_{r\geq 1}\frac{\varphi(r)}{r^4}\\ &= \frac{\zeta(2)^2\zeta(3)}{\zeta(4)} < \infty.\end{aligned}$$

Consequently,

$$\sum_{\substack{d^2|n\\ e^2|n}} C_p\left(\frac{n}{d^2}, \frac{n}{e^2}\right) = O\left(ns(n)\big(1+\log(2n)\big)^2\right).$$

By $h_p = p/12 + O(1)$,

$$\mathbb{E}\left[\binom{r_E}{2}\right] = O\left(\frac{ns(n)}{p}\big(1+\log(2n)\big)^2\right).$$

Combining this with (B.5) and (B.6), as well as $h_p = p/12 + O(1)$ proves

$$\Pr(\mathsf{A}_{E,n}) = \frac{12}{\pi}\sqrt{\frac{n}{p}}\mathcal{H}_{p,n} + O\left(\frac{ns(n)}{p}\big(1+\log(2n)\big)^2\right), \tag{B.7}$$

where the implied constant is absolute.

Consider

$$\frac{\Pr(\mathsf{A}_{E,n})}{\frac{12}{\pi}\sqrt{\frac{n}{p}}\mathcal{H}_{p,n}} = 1 + O\left(\sqrt{\frac{n}{p}}\frac{s(n)(1+\log(2n))^2}{\mathcal{H}_{p,n}}\right).$$

For every fixed $\eta > 0$, Siegel's theorem [Sie35], together with the induced-character formula and Mertens' product estimate, gives

$$\mathcal{L}_{p,n} \geq c_\eta p^{-\eta}$$

uniformly for $1 \leq n \leq N$ for a constant $c_\eta > 0$. Since all terms in $\mathcal{H}_{p,n}$ are positive,

$$\mathcal{H}_{p,n} \geq \mathcal{L}_{p,n} \geq c_\eta p^{-\eta}.$$

Since $s(n) \leq \sqrt{n}$ and $n \leq N \leq p^{\frac{1}{3}}$,

$$\frac{\Pr(\mathsf{A}_{E,n})}{(12/\pi)\sqrt{n/p}\,\mathcal{H}_{p,n}} = 1 + O_\eta\left(p^{-1/6+\eta}(\log p)^2\right).$$

Choosing any fixed $0 < \eta < 1/6$ proves uniformly for $1 \leq n \leq N$:

$$\frac{\Pr(\mathsf{A}_{E,n})}{\frac{12}{\pi}\sqrt{\frac{n}{p}}H_{p,n}} = 1 + o(1).$$

Finally, for every fixed $\eta > 0$,

$$c_\eta p^{-\eta} \leq \mathcal{H}_{p,n} \leq C(\log p)^2.$$

Since $\eta > 0$ is arbitrary, this implies

$$\log \mathcal{H}_{p,n} = O(\log\log p)$$

uniformly for $1 \leq n \leq N$.

Taking logarithms, we obtain

$$\begin{aligned}\log \Pr(\mathsf{A}_{E,n}) &= \frac{1}{2}\log\frac{n}{p} + \log \mathcal{H}_{p,n} + O(1)\\ &= \frac{1}{2}\log\frac{n}{p} + O(\log\log p).\end{aligned}$$

This proves (B.4). □

Remark. From the proof, we have:

$$\frac{\Pr(\mathsf{A}_{E,n})}{\sqrt{12n/p}} = \frac{\sqrt{12}}{\pi}\mathcal{H}_{p,n} + O\left(p^{-1/6}(\log p)^2\right). \tag{B.8}$$

Thus, to disprove the proposed uniform fixed-constant asymptotic in this family, it suffices to construct a sequence $(p_j, n_j)$ for which

$$\mathcal{H}_{p_j,n_j} \not\longrightarrow \frac{\pi}{\sqrt{12}}.$$

However, there are lots of sequence without this limit. So the constant term in Heuristic M does not always hold.

[1]School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, Anhui, China

[2]Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China

*Email address*: yiouyang@ustc.edu.cn

*Email address*: veivei@mail.ustc.edu.cn

*Email address*: xuzheng1@mail.ustc.edu.cn