Post-quantum commutative group actions from oriented superspecial abelian varieties

Wouter Castryck[1], Bogdan Adrian Dina[2,3], Péter Kutas[2,4], and Gioella Lorenzon[1]

[1] KU Leuven, Belgium
[2] Eötvös Loránd University, Hungary
[3] University of Haifa, Israel
[4] University of Birmingham, United Kingdom

Abstract. We study the cryptographic use of Shimura class group actions on superspecial principally polarized abelian varieties equipped with a CM orientation, within the context of post-quantum cryptography. Besides giving an accessible introduction to the topic, illustrated with many examples, we provide tools for representing CM orientations through higher-dimensional isogenies and for computing the Shimura class group action in the restricted effective group action model. Here, key auxiliary results are a version of Kani's lemma "with CM" and a trick that often allows to halve the ambient dimension. We also give a method for finding superspecial principally polarized abelian surfaces that are oriented by a given CM order, which is extracted from an algebraic framework for studying CM orientations using the Ibukiyama–Katsura–Oort correspondence. We also analyze obstructions to transitivity of the Shimura class group action, related to the characteristic-zero notion of CM type.

1 Introduction

Commutative cryptographic group actions generalize exponentiation in finite cyclic groups and inherit some of its versatility in cryptographic design, such as non-interactive key exchange protocols à la Diffie–Hellman and threshold signatures, see [1] for an overview. At the same time, the concept is general enough to allow for (conjecturally) quantum-resistant instantiations, unlike exponentiation which falls prey to Shor's algorithm. This generality is, however, relative: the only known fruitful source of post-quantum commutative group actions comes from the theory of complex multiplication (CM). Moreover, nearly all known instantiations work in dimension $g = 1$, i.e., using elliptic curves, where the groups are ideal-class groups of imaginary quadratic orders, acting on sets of oriented elliptic curves via isogenies. Their cryptographic use was proposed independently by Couveignes [29] and Rostovtsev–Stolbunov [87], using ordinary elliptic curves over finite fields, but the focus has now shifted to supersingular elliptic curves for the sake of flexibility and efficiency. This is implemented, e.g., in [17,35,5].

CM theory in arbitrary dimension $g \geq 1$ was developed by Shimura and Taniyama mid-20th century [89]. Unlike the elliptic curve case, the acting group

is not simply the ideal-class group of an imaginary quadratic order, but a slightly more complicated class group attached to a CM order of rank $2g$. The complication comes from the fact that it acts on principally polarized abelian varieties (PPAVs); the action needs to be compatible with the polarization data.

Given the growing role of higher-dimensional abelian varieties in isogeny-based cryptography, and given the central place of CM theory in the theory of PPAVs, it is a natural goal to study how Shimura class group actions can be put to cryptographic use. Our purpose is to initiate this study for superspecial PPAVs, which are the appropriate generalizations to higher dimension of supersingular elliptic curves. Sanon's PhD thesis [88] already gave a high-level description of a key exchange protocol from Shimura class groups acting on absolutely simple, ordinary abelian surfaces over a finite field $\mathbb{F}_q$ oriented by their full endomorphism ring. We are unaware of other literature exploring such constructive uses, let alone in the superspecial setting, although [83] mentions the possibility of extending the module-action framework to modules over CM orders, which would generalize our discussion (which corresponds to rank-1 modules). This being said, the Shimura class group action has been on cryptographers' radars for a long time, in tangential studies such as [68,6,7,63,10,51].

When compared to elliptic curves, the cases $g > 1$ offer two potential advantages: *(i)* one can reach larger class groups for a given field size, especially when working over a prime field $\mathbb{F}_p$ as in CSIDH [17], which is not possible in large-discriminant frameworks such as [16,46], and *(ii)* there potentially exist more prime ideals of a given small norm ℓ, allowing for a key space that is more concentrated on small primes. But we mainly view this work as a didactic exploration of CM orientations on superspecial PPAVs and the corresponding Shimura class group actions, with a focus on their cryptographic use and spiced up with many examples. We also note that, already in dimension $g = 1$, orientations and CM actions have proven to be important tools beyond the constructive use of cryptographic group actions [34,45,66].

Contributions and paper organization. Our contributions are organized as follows. First, after reviewing CM orientations and Shimura class groups in Section 2, with further background to be found in the appendices, we introduce some recurring examples of superspecial CM orientations in Section 3. Next, in Section 4 we discuss how to set up and evaluate Shimura class group actions within the **g**-REGA model from [1]. This includes finding the generating set **g** and representing the orientation; two auxiliary tools here are an enhanced version of Kani's lemma incorporating the CM orientation, and a dimension-halving trick often allowing the user to keep working in dimension g (which seems a new observation even for $g = 1$). Along the way we describe the structure of Shimura class groups of CM orders in biquadratic CM fields (i.e., $C_2 \times C_2$ fields), and we break a security measure proposed by Sanon in Remark 4.8. Then in Section 5, we give a purely algebraic treatment of CM orientations on superspecial principally polarized abelian surfaces (PPASs), within the framework of the Ibukiyama–Katsura–Oort (IKO) correspondence from [49]. We show how to evaluate the

Shimura class group action on the IKO-side, and how to produce superspecial PPASs that are oriented by a prescribed CM order; this has applications to the 2-dimensional version of the "one endomorphism" problem (often called OneEnd in the literature [79]), on which we elaborate in Appendix I. Next, in Section 6, we delve into the orbits of the Shimura class group action, which is a rich and much more intricate discussion than in the case $g = 1$, led by the characteristic-zero notion of CM type. An interesting by-product of our discussion is a very efficient method for deciding whether two given oriented supersingular elliptic curves lie in the same orbit under the usual ideal-class group action, see Example 6.3. We conclude in Section 7 with some security considerations.

Appendices. To keep the main body contained, a considerable amount of material was deferred to the appendices. Besides additional background in Appendices A, J and proofs in Appendices D, E, H, they contain several results worth highlighting. Most notably, building on Section 4, we present a key exchange protocol built from superspecial CM-oriented PPASs, all defined over $\mathbb{F}_p$, where the real multiplication part is generated by the p-th power Frobenius endomorphism. The starting orientation is called R2-D4 and described in Appendix C, while concrete parameters tailored for efficient key exchange are discussed in Appendix F. Some compression techniques for CM-oriented superspecial PPASs over $\mathbb{F}_p$ can be found in Appendix G. Along the way, for use in R2-D4 and similar orientations, we give explicit descriptions of the endomorphism rings of certain well-known supersingular elliptic curves in Appendix B, thereby fixing and completing corresponding statements in the existing literature. Away from key exchange, we recall that Appendix I contains results on the 2-dimensional counterpart of OneEnd. In Appendix K we slightly extend a result due to Goren [42], which describes modulo which unramified primes p a given CM-oriented PPAS in characteristic zero has a superspecial reduction; concretely, unlike Goren, we also cover the $C_2 \times C_2$ case, i.e., the case of CM by a biquadratic field.

Acknowledgements. The authors would like to thank Eyal Goren, the reviewers of a prior version of this paper, and the authors of [28,50] for helpful input and feedback. For the statement and proof of Lemma 5.14 we strongly benefited from interactions with Claude Opus 5.5 (Anthropic PBC). All authors are supported by the European Union under the Horizon Europe programme with grant agreement No. 101309092 (PQC-SCAN), as well as by the CELSA alliance through the MaCro project. Castryck and Lorenzon are also supported by the European Research Council (ERC) under the Horizon 2020 programme, grant agreement No. 101020788 (ISOCRYPT), by the Research Council KU Leuven grant C14/24/099, as well as by Cybersecurity Research Flanders with reference number VR20192203. Kutas is partly supported by EPSRC through grant number EP/V011324/1, and together with Dina, he is also supported by the Hungarian Ministry of Innovation and Technology NRDI Office within the framework of the Quantum Information National Laboratory Program. Kutas is also supported by NRDI grant "EXCELLENCE-151343". Lorenzon is funded by

Research Foundation – Flanders (FWO) under a PhD Fellowship fundamental research (project number 1139225N).

2 Background

We review the notions of principally polarized abelian varieties (PPAVs) equipped with a CM orientation and the Shimura class group action. More elementary background can be found in Appendix A, to which we have deferred all proofs.

2.1 CM orientations

Henceforth K denotes a CM field of degree $2g$, i.e., a totally imaginary quadratic extension of a totally real degree-g number field $K_0 \supset \mathbb{Q}$. Observe that K comes equipped with an automorphism $x \mapsto \overline{x}$ with fixed field K_0 that we can think of as complex conjugation. We write K_0^+ for the set of totally positive elements in K_0, i.e., whose minimal polynomial over $\mathbb{Q}$ has positive real roots only.

Definition 2.1. *A K-orientation on a g-dimensional PPAV (A, λ) is a homomorphism $\iota : K \hookrightarrow \mathrm{End}^0(A)$ such that $\iota(\overline{\alpha}) = \iota(\alpha)^\dagger$ for all $\alpha \in K$, where † denotes the Rosati involution. We say that ι is moreover a primitive R-orientation, for some order $R \subset K$, if $\iota^{-1}\,\mathrm{End}(A) = R$.*

Since $\mathrm{End}(A)$ is closed under the Rosati involution, primitive R-orientations can only occur for orders that are self-conjugate, i.e., $\overline{R} = R$. [5] We call such an order a CM order. A sufficient condition is that R has maximal real multiplication, i.e., $R \cap K = \mathbb{Z}_{K_0}$. We remark that we use the terminology of "orientations" for the sake of analogy with cryptographic customs in the case $g = 1$, but we note that many mathematical texts just talk about PPAVs "with CM".

Consider two K-oriented PPAVs $(A_1, \lambda_1, \iota_1), (A_2, \lambda_2, \iota_2)$. Let $\mu \in K_0^+ \cap \iota_1^{-1}\,\mathrm{End}(A_1)$, i.e., $\mu \in K$ is such that $\iota_1(\mu)$ is a totally positive endomorphism on A_1 (e.g., we can choose $\mu = N \in \mathbb{Z}_{>0}$). Consider an $\iota_1(\mu)$-isogeny $\varphi : (A_1, \lambda_1) \to (A_2, \lambda_2)$. Then we say that this polarized isogeny is an oriented isogeny if for all $\alpha \in \iota_1^{-1}\,\mathrm{End}(A_1) \cap \iota_2^{-1}\,\mathrm{End}(A_2)$ we have $\varphi \circ \iota_1(\alpha) = \iota_2(\alpha) \circ \varphi$. We mark this by writing $\varphi : (A_1, \lambda_1, \iota_1) \to (A_2, \lambda_2, \iota_2)$. If both orientations are primitive for the same order $R \subset K$ then we say that it concerns a horizontal oriented isogeny. An oriented isomorphism is an oriented 1-isogeny.

Definition 2.2. *Let K be a CM field of degree $2g$. We write $\mathcal{M}_k(K)$ for the set of K-oriented PPAVs (A, λ, ι) over k, up to oriented isomorphism. For a CM order $R \subset K$, we write $\mathcal{M}_k(R)$ for the subset of primitively R-oriented PPAVs.*

Remark 2.3. If $g = 1$ then we suppress the canonical principal polarization from the notation and denote the elements of $\mathcal{M}_k(K)$ as (E, ι). The set $\mathcal{M}_k(R)$ may be empty. E.g., this is the case if $g = 1$, $p > 0$ and $p \mid [\mathbb{Z}_K : R]$, see Section 6.

[5] E.g., the order $\mathbb{Z}[\sqrt{2} + \sqrt{-1}] \subset \mathbb{Q}(\sqrt{2}, \sqrt{-1})$ does not satisfy $\overline{R} = R$.

2.2 The Shimura class group

To any CM order R in a CM field K we attach its Shimura class group, denoted $\mathfrak{C}(R)$. The nomenclature in the existing literature is heterogeneous: the Shimura class group also goes under the name polarized class group, while our main reference [65] even just talks about "a Picard group", denoting it as $\mathrm{Pic}_* R$. Explicitly, consider pairs $(\mathfrak{a}, \mu)$ where $\mathfrak{a}$ is an invertible fractional R-ideal such that $\mathfrak{a}\bar{\mathfrak{a}} = \mu R$, and $\mu \in K_0^+$ (i.e., μ is a totally positive element of K_0). Endowed with component-wise multiplication, such pairs form a group we denote by I_R. Pairs $(\mu R, \mu\bar{\mu})$ with $\mu \in K^\times$ form a subgroup we denote by P_R. The Shimura class group of R is the quotient

$$\mathfrak{C}(R) := I_R / P_R. \tag{1}$$

We denote by $\sim$ the equivalence induced by P_R on pairs $(\mathfrak{a}, \mu) \in I_R$. If $g = 1$ then the existence of $\mu \in K_0^+ = \mathbb{Q}_{>0}$ such that $\mathfrak{a}\bar{\mathfrak{a}} = \mu R$ is an automatic condition and it determines μ uniquely (namely, it must be the norm of $\mathfrak{a}$). Thus, in this case, it is immediate that $\mathfrak{C}(R) \cong \mathrm{Cl}(R)$.

Lemma 2.4. *For a given invertible fractional ideal $\mathfrak{a} \subset R$ there are at most 2^{g-1} Shimura classes $[(\mathfrak{a}, \mu)] \in \mathfrak{C}(R)$.*

Example 2.5. If $K_0 = \mathbb{Q}(\sqrt{2})$ then $\varepsilon_1 = 1 + \sqrt{2}$ is not totally positive, so there is one class only. If $K_0 = \mathbb{Q}(\sqrt{3})$ then $\varepsilon_1 = 2 + \sqrt{3}$ is totally positive, so here there will be two classes, typically, but there are exceptions: e.g., if $K = K_0(\alpha)$ with $\alpha = \sqrt{-\varepsilon_1}$ then clearly $\varepsilon_1 = \alpha\bar{\alpha}$ so, here too, there is one class only.

As explained in the proof of Theorem 2.7 below, if $\mathfrak{a} \subset R$ then we can act with the Shimura ideal $(\mathfrak{a}, \mu)$ on any primitively R-oriented PPAV (A, λ, ι) by means of an $\iota(\mu)$-isogeny. Since all existing formulae for isogeny evaluation in higher dimension restrict to N-isogenies for some positive integer N, in practice we will restrict to Shimura ideals for which $\mu = N \in \mathbb{Z}_{>0}$, even though in Corollary 5.6 we will explain how the KLPT^2 algorithm from [13] can be used, in principle, to by-pass this restriction in the case of an oriented superspecial PPAS with "known" endomorphism ring. The Shimura classes of the form $[(\mathfrak{a}, N)]$ with $\mathfrak{a} \subset R$ and $N \in \mathbb{Z}_{>0}$ form a subgroup $\mathfrak{C}(R)_{\mathbb{Z}} \subset \mathfrak{C}(R)$ whose index is bounded by the following statement, for which we provide a simple proof; this also follows from [7, Lem. 6].

Lemma 2.6. *We have $\mathfrak{C}(R)^2 \subset \mathfrak{C}(R)_{\mathbb{Z}}$, therefore $[\mathfrak{C}(R) : \mathfrak{C}(R)_{\mathbb{Z}}] \mid \#\mathfrak{C}(R)[2]$.*

The Shimura class group fits into the following exact sequence:

$$1 \to \frac{(R_0^\times)^+}{N_{K/K_0}(R^\times)} \xrightarrow{u \,\mapsto\, [(R,u)]} \mathfrak{C}(R) \xrightarrow{[(\mathfrak{a},\mu)] \,\mapsto\, [\mathfrak{a}]} \mathrm{Cl}(R) \xrightarrow{N_{K/K_0}} \mathrm{Cl}^+(R_0), \tag{2}$$

where $R_0 = R \cap K_0$, $(R_0^\times)^+$ is the subgroup of totally positive units, and $\mathrm{Cl}^+(R_0)$ is the narrow class group (which, as we recall, is the group of invertible fractional

ideals modulo the subgroup of principal fractional ideals admitting a totally positive generator). For instance, this implies that $\#\mathfrak{C}(R)[2] \leq \#\operatorname{Cl}(R)[2] \cdot U$ with $U = [(R_0^\times)^+ : N_{K/K_0}(R^\times)] \leq 2^{g-1}$, where the last bound arises as in the proof of Lemma 2.4. Under mild conditions, the last map is surjective. For example, this is true if $R = \mathbb{Z}_K$ and at least one prime ideal of R_0 ramifies in K [9, Proof of Thm. 3.1] (that theorem also states that if $g = 2$ then this can only fail if $\operatorname{Gal}(K, \mathbb{Q}) \cong C_2 \times C_2$). So the Shimura class number $\mathfrak{h}(R) = \#\mathfrak{C}(R)$ can often be computed as $h(R)U/h^+(R_0)$. An algorithm for computing Shimura class groups was described in [40, §4] (implemented in PARI/GP).

We are now ready to state our main theorem of interest:

Theorem 2.7. *Assume that the set $\mathcal{M}_k(R)$ is non-empty. There is a free action*

$$\mathfrak{C}(R) \times \mathcal{M}_k(R) \to \mathcal{M}_k(R) : (\,[(\mathfrak{a}, \mu)], (A, \lambda, \iota)\,) \mapsto [(\mathfrak{a}, \mu)] \star (A, \lambda, \iota).$$

It can be assumed w.l.o.g. that $\mathfrak{a} \subset R$ and has norm coprime to p, in which case the image arises as the codomain of an $\iota(\mu)$-isogeny with kernel $\cap_{x \in \mathfrak{a}} \ker \iota(x)$.

At this point we make no claims on the orbits of the action. This is discussed in Section 6.

3 First examples of superspecial CM orientations

The focus of our paper lies on the case of superspecial abelian varieties (and surfaces in particular). The easiest way to generate a g-dimensional PPAV that is oriented by a CM field K of degree $2g$ is by considering $A = E^g$, for some supersingular elliptic curve E, equipped with the product polarization λ. The curve E admits an isomorphism $\vartheta : \mathcal{O} \to \operatorname{End}(E)$ from a maximal order $\mathcal{O}$ in the unique definite quaternion algebra $B_{p,\infty}$ over $\mathbb{Q}$ that is ramified at p. By component-wise application we obtain an isomorphism $\vartheta : M_g(\mathcal{O}) \to \operatorname{End}(A)$ that we denote by the same letter. Therefore giving a K-orientation on (A, λ) is the same as giving an embedding $\iota : K \hookrightarrow M_g(B_{p,\infty})$ such that $\iota(\bar{\alpha}) = \iota(\alpha)^*$ for all $\alpha \in K$; here * denotes conjugate-transpose, which under ϑ indeed corresponds to the Rosati involution on A according to Example A.3. For a CM order $R \subset K$, a primitive R-orientation is then a K-orientation mapping R into $M_g(\mathcal{O})$, in such a way that $\iota(\alpha) \notin M_g(\mathcal{O})$ for any $\alpha \in K \setminus R$.

Example 3.1. Let $p \equiv 3 \bmod 4$ and consider the elliptic curve $E : y^2 = x^3 - x$ over $\mathbb{F}_{p^2}$, which is supersingular. Fix a square root $\mathbf{i} \in \mathbb{F}_{p^2} \setminus \mathbb{F}_p$ of -1. Consider $A = E^2$ equipped with the product polarization λ. It is well-known that $\vartheta : i \mapsto (\,(x, y) \mapsto (-x, \mathbf{i}y)\,)$, $j \mapsto \pi_p$ establishes an isomorphism between the maximal order $\mathcal{O} = \mathbb{Z}\langle 1, i, (1+j)/2, (i+k)/2 \rangle \subset B_{p,\infty}$ and $\operatorname{End}(E)$. Assume that $m = p+4$ and $d = (p+1)/4$ are coprime, square-free integers. We claim that

$$\iota \;:\; \frac{1+\sqrt{-m}}{2} \mapsto \begin{pmatrix} \frac{1+j}{2} & i \\ i & \frac{1+j}{2} \end{pmatrix},\ \sqrt{d} \mapsto \begin{pmatrix} 0 & \frac{1+j}{2} \\ \frac{1-j}{2} & 0 \end{pmatrix}$$

is a primitive R-orientation on (A, λ), where $R = \mathbb{Z}[(1+\sqrt{-m})/2, \sqrt{d}]$; as above we identify elements of $M_2(\mathcal{O})$ with their images in $\mathrm{End}(A)$ under ϑ. If we denote the two matrices by M and D, then in order to verify that this indeed concerns a CM orientation, the key properties to check are that

- M and D commute,
- $M^2 - M + (m+1)/4\mathrm{Id}_2 = 0$ and $D^2 - d\mathrm{Id}_2 = 0$,
- $M^* = 1 - M$ and $D^* = D$,

which is easily done. To verify primitivity, note that if $d \not\equiv 1 \bmod 4$ then $R = \mathbb{Z}_K$ (this follows, e.g., from [62, p. 68]), in which case primitivity is automatic. To cover the case $d \equiv 1 \bmod 4$ it is enough to observe that $\iota((1+\sqrt{d})/2) \notin M_2(\mathcal{O})$.

As a toy example, letting $p = 103$ we find $m = 107$ and $d = 26$. Using that the fundamental unit $\sqrt{26}+5$ of $\mathbb{Q}(\sqrt{26})$ is not totally positive and that $107R_0$ ramifies in K,[6] from (2) we see that the sequence

$$1 \to \mathfrak{C}(R) \xrightarrow{[(\mathfrak{a},\mu)] \,\mapsto\, [\mathfrak{a}]} \mathrm{Cl}(R) \xrightarrow{N_{K/K_0}} \mathrm{Cl}^+(R_0) \to 1$$

is exact. With the use of Magma or SageMath one then verifies that $\mathfrak{C}(R)$ has order 42 and is generated by the classes of

$$\mathfrak{t} = \left((3, \sqrt{-107}-1), 3\right), \qquad \mathfrak{s} = \left((7, \sqrt{-107}\sqrt{26}-2), 7\right)$$

(and cyclically generated by the class of their product $\mathfrak{st}$). Then, for instance, using the $(3,3)$-isogeny formulae from [11] one calculates that $[\mathfrak{t}] \star (A, \lambda, \iota) = (\mathrm{Jac}(H), \lambda_1, \iota_1)$ where $H : y^2 = 102x^6 + 25x^5 + 57x^4 + 4x^3 + 41x^2 + 46x + 14$ and λ_1 is the canonical principal polarization on $\mathrm{Jac}(H)$, and where ι_1 denotes the push-forward orientation (which we do not spell out here). See Section 4.3 for more details on how to evaluate the Shimura class group action in practice.

Example 3.2. Now assume that moreover $d = a^2 + b^2$ for certain integers $a, b \in \mathbb{Z}$ (e.g., $26 = 5^2 + 1^2$ in the toy example). Since d is square-free we can assume that b is odd. Then a similar reasoning shows that

$$\iota' \;:\; \frac{1+\sqrt{-m}}{2} \mapsto \begin{pmatrix} i + \frac{1+j}{2} & 0 \\ 0 & i + \frac{1+j}{2} \end{pmatrix}, \; \sqrt{d} \mapsto \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

is another R-orientation on (A, λ). It is always primitive because using that b is odd one sees that $\iota'((1+\sqrt{d})/2) \notin M_2(\mathcal{O})$. In Example 6.4 we will argue that (A, λ, ι) and (A, λ, ι') cannot lie in the same orbit under the action of $\mathfrak{C}(R)$.

Example 3.3. Away from products, let $p \neq 2, 5$ and consider the genus-2 curve $C : y^2 = x^5 - 1$ together with its Jacobian $A = \mathrm{Jac}(C)$, equipped with the canonical polarization. This PPAS is superspecial if and only if $p \equiv 4 \bmod 5$. For the CM order $R = \mathbb{Z}[\zeta_5]$ we have a primitive orientation $\iota : R \hookrightarrow \mathrm{End}(A)$ sending ζ_5 to the automorphism of A induced by $C \to C : (x, y) \mapsto (\xi_5 x, y)$; here ζ_5, ξ_5 denote primitive 5-th roots of unity in $\mathbb{C}$ and k, respectively.

[6] It can in fact be argued that the existence of a ramifying prime follows from the fact that the fundamental unit is not totally positive.

Example 3.4 (R2-D2). We fix a supersingular elliptic curve $E/\mathbb{F}_{p^2}$ for which $\pi_{p^2} = -p$ and let (A, λ) be its Weil restriction, i.e.,

$$A = W_{\mathbb{F}_{p^2}/\mathbb{F}_p}(E) \cong_{\bar{\mathbb{F}}_p} E \times E^{(p)}$$

equipped with the product polarization. We choose E such that it comes equipped with an explicit isogeny $\tau : E \to E^{(p)}$, satisfying $\hat{\tau} \circ \pi_p = \pi_p \circ \tau$, of square-free degree d not a multiple of p. Such curves were studied in [22] and a method for finding large examples can be found in [16, §5]. Then, an orientation is given by

$$\iota : \mathbb{Q}(\sqrt{d}, \sqrt{-p}) \to \mathrm{End}^0(A) : \sqrt{d} \mapsto \begin{pmatrix} 0 & \hat{\tau} \\ \tau & 0 \end{pmatrix}, \sqrt{-p} \mapsto \Pi_p = \begin{pmatrix} 0 & \pi_p \\ \pi_p & 0 \end{pmatrix},$$

where we note that Π_p is the p-th power Frobenius endomorphism on A. These matrices are fixed, resp. negated, under the Rosati involution (= conjugate-transpose), they square to d, resp. $-p$, and they commute, so this is a well-defined field embedding. It typically concerns a primitive R-orientation with $R = \mathbb{Z}[\sqrt{d}, \sqrt{-p}]$. Indeed, clearly $(1 + \sqrt{d})/2, (1 + \sqrt{-p})/2$ do not map to endomorphisms, so the only concern is whether $(1 + \sqrt{-dp})/2$ maps to an endomorphism. This concern is void if $-dp \not\equiv 1 \bmod 4$ (true for the parameters from [16]) because then $(1+\sqrt{-dp})/2$ is not an algebraic integer. In general, one can verify that the orientation is primitive if and only if $(1 + \pi_p \circ \tau)/2 \notin \mathrm{End}(E)$.

In the above examples, the primitivity of the orientation was easily established. In general, however, finding the primitive order $R = \iota^{-1}\,\mathrm{End}(A)$ from a description of ι on a suborder $R' \subsetneq R$ may be a very hard problem. Indeed, this is similar to computing the endomorphism ring of an ordinary abelian surface A over a finite field $\mathbb{F}_q$ from its Frobenius order $\mathbb{Z}[\pi_q, \hat{\pi}_q]$, for which no efficient algorithms are available [6].

Remark 3.5. R2-D2 is part of a larger family of orientations called R2, discussed in Appendix C. The suffix D2 refers to the Galois group $D_2 \cong C_2 \times C_2$ of $K = \mathrm{Frac}(R)$ over $\mathbb{Q}$.

Remark 3.6. In Section 5 we will develop the theory of CM orientations on the algebraic side of the IKO correspondence. Within that framework, we will discuss an easy technique for producing superspecial PPASs together with a prescribed CM orientation, see Section 5.3. In Appendix C we will put Example 3.4 inside the larger family of explicit $\mathbb{F}_p$-rational orientations that we call R2.

4 Evaluating the Shimura class group action

In this section we describe the main auxiliary tools needed for cryptographic applications of the Shimura ideal class group action on oriented superspecial PPAVs in the **g**-REGA model (e.g., non-interactive key exchange).

4.1 On the generating set g

As mentioned in Section 2.2, we will act with $\mathfrak{C}(R)_{\mathbb{Z}}$ in practice. Recall that this may be a strict subgroup of the full Shimura class group (but when choosing parameters carefully, equality may be enforced). We refer the reader to Section 5.2 for how to act with elements of $\mathfrak{C}(R) \setminus \mathfrak{C}(R)_{\mathbb{Z}}$ in dimension $g = 2$, but this method requires knowledge that is not available to Bob when extracting the shared secret from Alice's public key, and vice versa.

Similar to class group actions like CSIDH [17] and SCALLOP [35], we choose the elements of our generating set $\mathbf{g}$ to be representable by Shimura ideals $(\mathfrak{a}, \ell)$ for small odd primes $\ell \nmid \operatorname{disc} R$,[7] and we collect enough (yet polylogarithmically many) such ideals for these to generate $\mathfrak{C}(R)_{\mathbb{Z}}$ with overwhelming probability (heuristically). Note: the inverse of $[(\mathfrak{a}, \ell)]$ is represented by $(\overline{\mathfrak{a}}, \ell)$. In Appendix D we prove that, in dimension $g = 2$, the options for $\mathfrak{a}$ are the following.

Lemma 4.1. *For a given CM order R of degree 4 and a prime number $\ell \nmid \operatorname{disc} R$ there exists an ideal $\mathfrak{a} \subset R$ such that $(\mathfrak{a}, \ell) \in I_R$ if and only if ℓR splits as either*

$$\ell R = \mathfrak{p}_1\overline{\mathfrak{p}}_1\mathfrak{p}_2\overline{\mathfrak{p}}_2 \qquad \text{or} \qquad \ell R = \mathfrak{p}\overline{\mathfrak{p}} \tag{3}$$

for degree-1 prime ideals $\mathfrak{p}_1, \mathfrak{p}_2$, resp., a degree-2 prime ideal $\mathfrak{p}$. The corresponding options for $\mathfrak{a}$ are $\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_1\overline{\mathfrak{p}}_2, \overline{\mathfrak{p}}_1\mathfrak{p}_2, \overline{\mathfrak{p}}_1\overline{\mathfrak{p}}_2$, resp., $\mathfrak{p}, \overline{\mathfrak{p}}$.

One sees that if ℓ splits completely then we obtain two pairs $\mathfrak{l}, \mathfrak{l}^{-1}, \mathfrak{l}', \mathfrak{l}'^{-1} \in \mathfrak{C}(R)_{\mathbb{Z}}$, while if ℓ splits into two complex-conjugate primes, we end up with one pair $\mathfrak{l}, \mathfrak{l}^{-1} \in \mathfrak{C}(R)_{\mathbb{Z}}$. E.g., this means that if we consider a set of primes $\ell_1, \ldots, \ell_r$ that split completely along with the corresponding set $\mathbf{g} = \{\mathfrak{l}_1, \mathfrak{l}'_1, \ldots, \mathfrak{l}_r, \mathfrak{l}'_r\}$, then the number of expressions of the form

$$\mathfrak{l}_1^{e_1}\mathfrak{l}_1'^{e_1'} \cdots \mathfrak{l}_r^{e_r}\mathfrak{l}_r'^{e_r'}, \qquad |e_i|, |e_i'| \leq B \tag{4}$$

for some bound $B \in \mathbb{Z}_{\geq 0}$ equals $(2B+1)^{2r}$, potentially yielding the same amount of elements of $\mathfrak{C}(R)_{\mathbb{Z}}$. More generally, for arbitrary g, we end up with 2^{g-1} pairs in the completely split case, hence

$$(2B+1)^{r2^{g-1}}$$

many expressions. This growing "key space" for fixed r and B is one of the features of working in higher dimension.

Remark 4.2. The asymptotic proportions of primes factoring as in (3) depend on $\operatorname{Gal}(L, \mathbb{Q})$ where L denotes the Galois closure of K. Using Chebotarev's density theorem, one checks that these proportions are given by $\frac{1}{4}$, resp. $\frac{1}{2}$, in the case $C_2 \times C_2$, by $\frac{1}{4}$, resp. 0, in the case C_4, and by $\frac{1}{8}$, resp. $\frac{1}{4}$, in the case D_4. In practice, see e.g. Appendix F, we can bias this in our favor and choose parameters such that as many small primes ℓ as possible split completely.

[7] This is a convenient assumption but it can be relaxed, to allow for horizontal isogenies coming from ramified primes; since this adds technicality while not adding much to $\mathfrak{C}(R)$ (the corresponding ideal classes have very small order), we omit this.

In order to find the factorization of ℓR, one can look for an order $\mathbb{Z}[\gamma] \subset R$ (i.e., a monogenic suborder) such that every element of R can be written as a $\mathbb{Q}$-linear combination of $1, \gamma, \ldots, \gamma^{2g-1}$ without denominators that are 0 mod ℓ (e.g., this is true as soon as $\ell \nmid [R : \mathbb{Z}[\gamma]]$). Writing $f(x) \in \mathbb{Z}[x]$ for the minimal polynomial of γ, the factorization of ℓR can be obtained from that of $f(x)$ mod ℓ. For $g = 2$ the details are as follows; it is convenient to precompute[8] $r := \gamma_1\bar{\gamma}_1 + \gamma_2\bar{\gamma}_2 \in \mathbb{Z}$, where $\gamma_1 = \gamma, \bar{\gamma}_1, \gamma_2, \bar{\gamma}_2 \in \mathbb{C}$ are the roots of $f(x)$.

- If $f(x) \equiv f_1(x)f_2(x) \bmod \ell$ for irreducible quadratic factors $f_i(x)$ such that $f_1(0) + f_2(0) \equiv r \bmod \ell$ then $\ell R = \mathfrak{p}\overline{\mathfrak{p}}$ with $\mathfrak{p} = (\ell, f_1(\gamma))$.
- If $f(x)$ factors completely mod ℓ then we label the roots as a_1, b_1, a_2, b_2 of $f(x)$ mod ℓ in such a way that $r \equiv a_1b_1 + a_2b_2 \bmod \ell$. Then we can conclude that $\ell R = \mathfrak{p}_1\overline{\mathfrak{p}}_1\mathfrak{p}_2\overline{\mathfrak{p}}_2$ with $\mathfrak{p}_i = (\ell, \gamma - a_i)$ for $i = 1, 2$.

The $C_2 \times C_2$ case. There is an alternative way of describing an "easy" generating set $\mathbf{g}$ in the case of a CM field K with $\mathrm{Gal}(K, \mathbb{Q}) \cong C_2 \times C_2$, even though this time the group generated by $\mathbf{g}$ may be a strict subgroup of $\mathfrak{C}(R)_{\mathbb{Z}}$ (but of very small index). The starting point is that K has three quadratic subfields: the real field K_0 and two imaginary fields K_1, K_2. We let $R_i = R \cap K_i$ for $i = 1, 2$. It is then natural to consider the following map:

$$\kappa_R : \mathrm{Cl}(R_1) \times \mathrm{Cl}(R_2) \to \mathfrak{C}(R) : ([\mathfrak{a}_1], [\mathfrak{a}_2]) \mapsto [((\mathfrak{a}_1 R)(\mathfrak{a}_2 R), N_{K_1/\mathbb{Q}}(\mathfrak{a}_1)N_{K_2/\mathbb{Q}}(\mathfrak{a}_2))],$$

seen as taking values in $\mathfrak{C}(R)_{\mathbb{Z}}$. In Lemma 4.3 below we argue that this group homomorphism is close to being an isomorphism (at least in the case $R = \mathbb{Z}_K$). We can then build $\mathbf{g}$ from a set of "easy" ideal classes

$$[\mathfrak{m}_1], [\mathfrak{m}_2], \ldots, [\mathfrak{m}_s] \in \mathrm{Cl}(R_1), \qquad [\mathfrak{l}_1], [\mathfrak{l}_2], \ldots, [\mathfrak{l}_t] \in \mathrm{Cl}(R_2),$$

e.g., as chosen in CSIDH or SCALLOP, by letting it be the collection of images of $([\mathfrak{m}_i], 1)$, $i = 1, \ldots, s$ and $(1, [\mathfrak{l}_i])$, $i = 1, \ldots, t$.

Lemma 4.3. *Assume that K is a $C_2 \times C_2$ CM field, and let K_0 be the real quadratic subfield and K_1, K_2 the imaginary quadratic subfields. If $\mathbb{Z}_K^\times$ contains no roots of unity other than ± 1, then*

$$|\ker \kappa_{\mathbb{Z}_K}| \leq \frac{1}{2q(K)} \prod_{\ell \text{ prime}} e(\ell),$$

$$|\operatorname{coker} \kappa_{\mathbb{Z}_K}| \leq \frac{1}{4q(K)} |\operatorname{coker} N_{K/K_0}| \prod_{\ell \text{ prime}} e(\ell),$$

where $q(K) = [\mathbb{Z}_K^\times : \mathbb{Z}_{K_0}^\times] \in \{1, 2\}$, $e(\ell)$ is the ramification index of a prime ℓ in K, and $N_{K/K_0} : \mathrm{Cl}(\mathbb{Z}_K) \to \mathrm{Cl}^+(\mathbb{Z}_{K_0})$ denotes the ideal class norm map.

[8] An elegant way of computing r directly from $f(x)$ is through its cubic resolvent: in cases C_4, D_4 it concerns the unique integer root of the cubic resolvent, while in the $C_2 \times C_2$ case it can be characterized as the largest root; see [26].

Proof. The proof follows from results by Lemmermeyer [64] and the exact sequence (2). We refer to Appendix E for all the details. □

As mentioned in Section 2.2, a sufficient condition for $|\operatorname{coker} N_{K/K_0}| = 1$ is that the extension K/K_0 ramifies at a prime of K_0. A theorem characterizing such $C_2 \times C_2$ fields can be found in [92, Thm. 2]. We note that when R is not a maximal order, a similar bound on $\ker \kappa_R$ can also be proven (see Corollary E.1 in Appendix E). We expect that a similar bound on the cokernel also holds for κ_R, but the proof seems less straightforward.

4.2 Kani's lemma with CM

In Section 4.3 we will use the machinery of higher-dimensional isogeny representations for the purpose of representing orientations. Let us quickly revisit the case of a polarized isogeny $\varphi : A_1 \to A_2$ between PPAVs of dimension g, of reduced degree $d \in \mathbb{Z}_{>0}$. The higher-dimensional representation of φ exploits Kani's lemma [53] to encode φ into a higher-dimensional isogeny Φ of smooth reduced degree N, so that Φ, and hence φ, can be evaluated using a chain of isogenies of small degree. The classical method makes use of an auxiliary matrix α with entries in $\mathbb{Z}$, coming from an expression of $n := N - d$ as a sum of $u = 1, 2, 4$ squares (see [84, Prop. 5.15]); by Lagrange's theorem, $u = 4$ always succeeds.

For the purpose of representing orientations, we wish to generalize this to the setting where A_1 and A_2 have an explicit R-orientation for some CM order R in a CM field K of degree $2g$. Two things change. First, the auxiliary matrix may be taken with entries in R rather than in $\mathbb{Z}$, and so the condition on n becomes about relative norms of elements of R rather than about sums of integer squares. Second, the d-isogeny φ does not necessarily have $d \in \mathbb{Z}_{>0}$, as d can be any element of $R_0^+ = R_0 \cap K_0^+$, i.e., any totally positive element of R_0.

Theorem 4.4. *Let $(A_1, \lambda_1, \iota_1)$ and $(A_2, \lambda_2, \iota_2)$ be R-oriented PPAVs of dimension g, and let $\varphi\colon (A_1, \lambda_1, \iota_1) \to (A_2, \lambda_2, \iota_2)$ be an R-oriented d-isogeny with $d \in R_0^+$. Let $n \in R_0^+$ be such that $N := d+n \in \mathbb{Z}$, and let $\alpha = (\alpha_{j,k})_{1\le j,k\le u} \in M_u(R)$ be such that $\alpha^*\alpha = n \operatorname{Id}_u$, where α^* denotes the conjugate-transpose of α. Write $\alpha_{A_i} = \iota_{A_i}(\alpha) := (\iota_i(\alpha_{j,k}))_{1\le j,k\le u} \in \operatorname{End}(A_i^u)$. Then*

$$\Phi = \begin{pmatrix} \varphi \operatorname{Id}_u & \alpha_{A_2}^\dagger \\ -\alpha_{A_1} & \varphi^\dagger \operatorname{Id}_u \end{pmatrix} \;:\; (A_1^u \times A_2^u,\ \lambda_1^{\times u} \times \lambda_2^{\times u}) \to (A_2^u \times A_1^u,\ \lambda_2^{\times u} \times \lambda_1^{\times u})$$

is an N-isogeny of PPAVs of dimension $2ug$; in particular $\Phi^\dagger\Phi = [N]$ and $\deg \Phi = N^{2ug}$. Moreover:

(i) $\varphi = \varrho \circ \Phi \circ \jmath$, where $\jmath\colon A_1 \to A_1^u \times A_2^u$ is the inclusion of the first factor and $\varrho\colon A_2^u \times A_1^u \to A_2$ is the projection onto the first factor;

(ii) if the ideals dR_0 and nR_0 are coprime in R_0 and N is coprime to the base field characteristic, then

$$\begin{aligned} \ker \Phi &= \{(-\alpha_{A_1}^\dagger(P),\, \varphi \operatorname{Id}_u(P)) \mid P \in A_1^u[N]\} \\ &= \{(\iota_1(d)\operatorname{Id}_u(P),\, \alpha_{A_2}\varphi \operatorname{Id}_u(P)) \mid P \in A_1^u[N]\}. \end{aligned}$$

Remark 4.5. Point *(ii)* of Theorem 4.4 describes $\ker \Phi$ in terms of the action of φ on $A_1[N]$, so the images of the $2g$ points in a basis of $A_1[N]$ are enough to determine Φ. Assuming that N is a square, less than that actually suffices: the action of φ on $A_1[\sqrt{N}]$ already determines Φ, by the meet-in-the-middle technique of [86, §6.4], which we detail further in Remark D.1.

To apply Theorem 4.4 one has to produce, for a prescribed $n \in R_0^+$, a matrix $\alpha \in M_u(R)$ such that $\alpha^*\alpha = n\mathrm{Id}_u$, and one wants u as small as possible since Φ is an isogeny between abelian varieties of dimension $2ug$.

As mentioned, if $n \in \mathbb{Z}_{>0}$ then one can construct $\alpha \in M_u(\mathbb{Z})$ with $u = 1, 2, 4$ in the cases where n is a sum of $1, 2, 4$ integer squares. A similar technique applies if $n \in R_0^+$ is a sum of squares of elements in R_0: in [85, Lem. 4.2], Robert gives solutions for $\alpha \in M_{2^k}(R_0)$ when n is a sum of m squares, for some $k \geq \lceil \log_2 m \rceil$ where equality can be assumed if $m \leq 8$. E.g., for $n = \sum_{i=0}^{m-1} \vartheta_i^2$ with $\vartheta_i \in R_0$ the solutions for $m = 1, 2, 4$ are

$$\alpha = (\vartheta_0), \qquad \alpha = \begin{pmatrix} \vartheta_0 & \vartheta_1 \\ -\vartheta_1 & \vartheta_0 \end{pmatrix}, \qquad \alpha = \begin{pmatrix} \vartheta_0 & \vartheta_1 & \vartheta_2 & -\vartheta_3 \\ -\vartheta_1 & \vartheta_0 & \vartheta_3 & \vartheta_2 \\ -\vartheta_2 & -\vartheta_3 & \vartheta_0 & -\vartheta_1 \\ \vartheta_3 & -\vartheta_2 & \vartheta_1 & \vartheta_0 \end{pmatrix}. \tag{5}$$

There is no guarantee that $n \in R_0^+$ can always be written as a sum of squares in R_0, and if it can, more than 4 squares might be needed. A theorem by Siegel [90, Thm. 1] states that all totally positive integral elements of a totally real number field K_0 are a sum of integral squares only when $K_0 = \mathbb{Q}$ or $K_0 = \mathbb{Q}(\sqrt{5})$. Peters showed that if a totally positive integral element n of a real quadratic field is a sum of integral squares, then it is a sum of at most 5 integral squares, and gave an explicit criterion to detect such n [81, Sz. 2].

In the following, we give solutions for $\alpha \in M_u(R)$ in the cases where n is a sum of $k+1$ relative norms of elements of R, in which case one can always take $u = 2^k$, and in the cases where n is a sum of t relative norms of elements of R and $m = 2^d$ squares in R_0, in which case one ends up with $u = 2^{k+t}$ with $k \geq d$ as above. These solutions are more general than what is required for our applications in Section 4.3, but they are of independent interest, and do specialize to meet our needs.

Proposition 4.6. *Let $X, Y \in M_m(R)$ and set*

$$\alpha = \begin{pmatrix} X & Y^* \\ -Y & X^* \end{pmatrix} \in M_{2m}(R).$$

(a) If $X^*X = XX^* = x\,\mathrm{Id}_m$ and $Y^*Y = YY^* = y\,\mathrm{Id}_m$ for some $x, y \in R_0$, and if $XY = YX$, then

$$\alpha^*\alpha = \alpha\alpha^* = (x+y)\,\mathrm{Id}_{2m}.$$

(b) Conversely, for arbitrary $X, Y \in M_m(R)$ and $n \in R_0$ one has $\alpha^*\alpha = n\,\mathrm{Id}_{2m}$ if and only if

$$XY = YX, \qquad X^*X + Y^*Y = XX^* + YY^* = n\,\mathrm{Id}_m.$$

(c) *Let $k \in \mathbb{Z}_{\geq 0}$ and $\vartheta_0, \ldots, \vartheta_k \in R$. Define $\alpha_0 := (\vartheta_0) \in M_1(R)$ and, for $1 \leq i \leq k$,*

$$\alpha_i := \begin{pmatrix} \alpha_{i-1} & \bar{\vartheta}_i \operatorname{Id}_{2^{i-1}} \\ -\vartheta_i \operatorname{Id}_{2^{i-1}} & \alpha_{i-1}^* \end{pmatrix} \in M_{2^i}(R).$$

Then $\alpha_k^* \alpha_k = n \operatorname{Id}_{2^k}$ with $n = \sum_{i=0}^{k} N_{K/K_0}(\vartheta_i)$. In particular, if n is a sum of $k+1$ relative norms of elements of R there is a matrix $\alpha \in M_{2^k}(R)$ such that $\alpha^* \alpha = n \operatorname{Id}_{2^k}$.

(d) *Let $d, t \in \mathbb{Z}_{\geq 0}$ and let $m = 2^d$. Consider $\vartheta_0, \ldots, \vartheta_{m-1} \in R_0$. Let $\alpha_0 \in M_{2^k}(R_0)$ such that $\alpha_0^* \alpha_0 = \alpha_0 \alpha_0^* = n_1 \operatorname{Id}_{2^k}$ with $n_1 = \sum_{i=0}^{m-1} \vartheta_i^2$, for $k \geq d$ as above. Consider $\vartheta_1', \ldots, \vartheta_t' \in R$ and let $n_2 = \sum_{i=1}^{t} N_{K/K_0}(\vartheta_i')$. Define, for $1 \leq i \leq t$,*

$$\alpha_i := \begin{pmatrix} \alpha_{i-1} & \bar{\vartheta}_i' \operatorname{Id}_{2^{k+i-1}} \\ -\vartheta_i' \operatorname{Id}_{2^{k+i-1}} & \alpha_{i-1}^* \end{pmatrix} \in M_{2^{k+i}}(R).$$

Then $\alpha_t^* \alpha_t = (n_1 + n_2) \operatorname{Id}_{2^{k+t}}$. In particular, if n is a sum of t relative norms of elements of R and $m = 2^d$ squares of elements in R_0, there exists a matrix $\alpha \in M_{2^{k+t}}(R)$ such that $\alpha^* \alpha = n \operatorname{Id}_{2^{k+t}}$.

Proposition 4.6(d) reduces to [85, Lem. 4.2] when $n_2 = 0$, which a priori only offers a solution $\alpha \in M_{2^k}(R_0)$ but it is clear that if n_1 is a sum of squares in $\mathbb{Z}$ then also $\alpha \in M_{2^k}(\mathbb{Z})$ (as is apparent from (5) for $k = d = 0, 1, 2$).

4.3 Representing the orientation

Usage. Let (A, λ) be a g-dimensional PPAV together with a primitive orientation $\iota : R \to \operatorname{End}(A)$ by a CM order R of rank $2g$. Say we want to act on (A, λ, ι) with a Shimura ideal of the form $((\ell, f_1(\gamma)), \ell)$, for a suitable $\gamma \in R$ generating $K/\mathbb{Q}$ and a suitable degree-g polynomial $f_1(x) \in \mathbb{Z}[x]$; that is, a Shimura ideal of the kind described in Section 4.1. Following the proof of Theorem 2.7, this amounts to computing the subgroup

$$H := A[\ell] \cap A[f_1(\iota(\gamma))] \cong \left(\frac{\mathbb{Z}}{\ell\mathbb{Z}}\right)^g,$$

which is maximal isotropic with respect to the ℓ-Weil pairing. Generators for H can be obtained from the kernel of a matrix of $f_1(\iota(\gamma))$ acting on $A[\ell]$.

Remark 4.7. This can be sped up by using that H equals the image of $f_2(\iota(\gamma))$ on $A[\ell]$, where $f_2(x)$ denotes the cofactor of $f_1(x)$ in the factorization of the minimal polynomial $f(x)$ of γ. This generalizes an optimization that is well-known in the elliptic curve setting, see e.g. [16, §3.2].

The result of the action is then the codomain of an ℓ-isogeny with kernel H, equipped with the induced principal polarization and the push-forward orientation. For this approach, it is crucial to be able to evaluate $\iota(\gamma)$ on arbitrary ℓ-torsion points. Thus, we need to represent the orientation in a way that is both compact and allows the user to evaluate it efficiently.

Remark 4.8. As a side remark, observe that we can act with $((\ell, f_1(\gamma)), \ell)$ without knowing what R is, as long as $\gamma \in R$. One could hope to thwart Kuperberg's attack [23,58] on the vectorization problem on $\mathfrak{C}(R)$ by keeping R secret and just providing γ: as mentioned at the end of Section 3, extracting the primitive order R from a non-primitive orientation (by $\mathbb{Z}[\gamma]$ in this case) might indeed be an intractable problem. Interestingly, this countermeasure was proposed by Sanon [88, pp. 107–108], albeit in the context of ordinary PPAVs oriented by Frobenius. But it does not provide the hoped-for security: determining the stabilizer S of the (non-free) action by $\mathbb{Z}[\gamma]$ is an abelian hidden subgroup problem, which is solved in polynomial time using Kitaev's algorithm [54]. One then has a free action by $\mathfrak{C}(\mathbb{Z}[\gamma])/S$ which can be targeted using Kuperberg's algorithm.

Higher-dimensional representation. Let us first assume that we can efficiently evaluate endomorphisms coming from R_0. Then, in order to represent $\iota(\gamma)$, the routine approach, whose elliptic curve counterpart was first described in SCALLOP-HD [21], would be as follows: write $d = \overline{\gamma}\gamma \in R_0^+$ and suppose that we can find a smooth integer N, ideally a power of 2, such that $N - d$ can be effectively written as a sum of s squares of elements of R_0.[9] Then using the auxiliary matrix

$$\alpha \in M_u(R_0), \qquad u = 2^{\lceil \log_2 s \rceil}$$

from Proposition 4.6(d), we can embed $\iota(\gamma)$ in an N-isogeny between PPAVs of dimension $2ug$ as outlined in Theorem 4.4. This is an endomorphism of A^{2u}, in fact. This endomorphism can be computed from interpolation data for $\iota(\gamma)$, i.e., from the point pairs $P, \iota(\gamma)(P)$ with P running over a basis of $A[N]$. Recall from Remark 4.5 that if N is a square then using Robert's meet-in-the-middle technique it suffices to provide interpolation data on $A[\sqrt{N}]$. When evaluating the action, this interpolation data can be pushed through: if $\varphi : A \to A/H$ is a horizontal isogeny of degree coprime to N, then the point pairs $\varphi(P), \varphi(\iota(\gamma)(P))$ form a valid set of interpolation points for $(\varphi_* \iota)(\gamma)$.[10]

Halving the ambient dimension. In the common scenario where $\gamma = \sqrt{-d}$ and $N - d$ is a sum of at most 3 squares of elements of R_0, we can reduce the ambient dimension to ug. This observation seems new even for $g = 1$. To see this, let α be as in (5), where in the case of 3 squares we let $\vartheta_2 = 0$, i.e., $N - d = \vartheta_0^2 + \vartheta_1^2 + \vartheta_3^2$. Define

- $\xi := \sqrt{-d} + \alpha$ if $u = 1$ (i.e., $N - d$ is the square of an element of R_0),
- $\xi := \sqrt{-d} \cdot \mathrm{diag}(1, -1) + \alpha$ if $u = 2$ (sum of two squares),

[9] If $N - d$ is not a sum of squares then one can retry with a different N or γ; see our paragraph on representing R_0 at the end of this section for a workaround.

[10] In applications, this interpolation data must be masked using a change of basis for security reasons. In practice, one changes to a deterministic basis of $A[N]$ (or of $A[\sqrt{N}]$ in the case of meet-in-the-middle) for reasons of data compression, and this handles the security issue at once.

– $\xi := \sqrt{-d} \cdot \mathrm{diag}(1, -1, 1, -1) + \alpha$ if $u = 4$ (sum of 3 squares and $\vartheta_2 = 0$).

One then checks that $\iota_A(\xi)$ is a polarized endomorphism of A^u of reduced degree N. If we can evaluate this endomorphism at any point of A^u, then we can also evaluate $\iota(\gamma)$ at any point of A, assuming we know how to evaluate $\iota_A(\alpha)$. It is easy to see that the interpolation points $\varphi(P), \varphi(\iota(\sqrt{-d})(P))$ can be used to determine the kernel of $\iota(\xi)$. The method remains compatible with meet-in-the-middle: if the interpolation data is restricted to $A[\sqrt{N}]$ then this remains good enough to compute "half the kernel", i.e., $\ker(\iota(\xi)) \cap A^u[\sqrt{N}]$. But the data can also be used to compute half the kernel of the Rosati dual, i.e., $\ker(\iota(\xi^*)) \cap A^u[\sqrt{N}]$, allowing one to complete the path as wanted. Observe that the case $u = 1$ avoids the need for working in higher dimension altogether; for $g = 1$ this essentially specializes to the non-HD versions of SCALLOP [35,2].

Representing R_0. We now return to the problem of representing endomorphisms coming from R_0. In dimension $g = 2$ where $K_0 = \mathbb{Q}(\sqrt{c})$ for some square-free $c \in \mathbb{Z}_{>0}$, one can proceed by finding a smooth-degree representation of the endomorphism $\iota(\sqrt{c})$, which is a c-isogeny, using the methods from the previous paragraph but where now $\alpha \in M_u(\mathbb{Z})$. In the special case where $c = p$, this is actually not needed: here the representation comes for free, see the next Section 4.4. In arbitrary dimension, the most robust solution seems to be circumventing the problem and looking for some $\gamma \in R$ generating $K/\mathbb{Q}$ such that $\overline{\gamma}\gamma \in \mathbb{Z}$, so that the methods from the previous paragraph can be run over $\mathbb{Z}$ right away. The existence of such an element γ is not a strong condition on R.[11] E.g., if $R = \mathbb{Z}_K$ then any element of the form $(\beta/\bar{\beta})N_{K/\mathbb{Q}}(\beta)$ does the job, as long as $\beta/\bar{\beta}$ generates $K/\mathbb{Q}$. Moreover, by ensuring that $N_{K/\mathbb{Q}}(\beta)$ is smooth, one gets for free that $\overline{\gamma}\gamma = N_{K/\mathbb{Q}}(\beta)^2$ is smooth as well, again avoiding all need to go to higher dimension. However, this typically leads to generators γ having a large norm (corresponding to $\iota(\gamma)$ having a large degree), and it gives little control over $[R : \mathbb{Z}[\gamma]]$, meaning we potentially lose many small primes ℓ that could have been used to compute the Shimura class group action.

4.4 Instantiating key exchange over $\mathbb{F}_p$

In Appendix F we describe concrete key exchange protocols built from Shimura class group actions on superspecial PPASs. These are built from the R2 family of orientations: R2-D2 was described in Example 3.4, while R2-C4 and R2-D4 will be discussed in Appendix C. A common feature of these orientations is that all PPASs and isogenies involved when computing the action are defined over $\mathbb{F}_p$ and the p-th power Frobenius endomorphism Π_p is in the image of ι. This choice is motivated by CSIDH [17], where working with the Frobenius orientation on supersingular elliptic curves comes with three distinguishing features: *(i)* the

[11] Note that on PPAVs over a finite field $\mathbb{F}_q$ one always has $\pi_q \hat{\pi}_q = q$, so in the simple, ordinary case (where π_q generates the endomorphism algebra) one can simply work with Frobenius.

fact that every curve E in the orbit is defined over $\mathbb{F}_p$ makes for a more compact object; *(ii)* the orientation does not need representation, because the Frobenius endomorphism is always known explicitly; *(iii)* it is easy to choose parameters ensuring the splitting of many small primes ℓ (as well as control the field of definition of the points in a corresponding kernel).

In higher dimension, these features remain partially present:

- One can no longer avoid the need for representing the orientation altogether, but at least the Frobenius part of the orientation still comes for free. E.g., in the R2-D4 and R2-C4 orientations described in Appendix C, where $K_0 = \mathbb{Q}(\sqrt{p})$ and $\iota(\sqrt{p}) = \Pi_p$, this removes the need for representing an endomorphism of the form $\iota(\sqrt{c})$, $c \in \mathbb{Z}_{>0}$, overall making for a more compact orientation representation. Moreover, the $\mathbb{F}_p$-rationality can be exploited to compress this data further: this is explained in Appendix G.
- In all R2 orientations we have $\mathbb{Q}(\sqrt{\pm p}) \subset K$, so by choosing $\pm p$ to be a quadratic residue mod ℓ_i for small primes $\ell_1, \ldots, \ell_r$ we roughly double the chance that ℓ_i splits completely (from $\frac{1}{4}$ to $\frac{1}{2}$ in cases $C_2 \times C_2, C_4$ and $\frac{1}{8}$ to $\frac{1}{4}$ in case D_4). Moreover, parameter generation is flexible enough to boost those chances further; e.g., see our approach to R2-D4 taken in Appendix F.1.

Remark 4.9. There may be reasons to stay away from Frobenius orientations, e.g., for more control over $\#\mathfrak{C}(R)$ or more flexibility in choosing parameters, as in the SCALLOP family [35,21,2] (a general technique for constructing such orientations in dimension $g = 2$ will be outlined in Section 5.3). In the $C_2 \times C_2$ case, there is another very good reason for avoiding Frobenius orientations: based on a preliminary version of the current paper, the authors of [50] observed that the vectorization problem for R2-D2, which we recall is an orientation by a CM order of the form $\mathbb{Z}[\sqrt{d}, \sqrt{-p}]$, reduces to vectorization problems in $\mathrm{Cl}(\mathbb{Z}[\sqrt{-p}])$ and $\mathrm{Cl}(\mathbb{Z}[\sqrt{-dp}])$. This relies on a twisting trick akin to [18,96] and, as far as we are aware, does not obviously generalize to arbitrary $C_2 \times C_2$ orientations, because it depends on a property that is specific to cases where p ramifies in K, namely that twists coming from "inequivalent CM types" can nevertheless end up in the same orbit; see Section 6, and Example 6.12 in particular. For the sake of reference, we have still included (a shortened version of) our R2-D2-based key exchange protocol in Appendix F.2.

5 CM orientations under the IKO correspondence

The IKO framework, introduced in [49], gives a convenient Deuring-like dictionary between the geometry of superspecial PPASs over $\overline{\mathbb{F}}_p$ and their connecting isogenies on one hand, and the algebra of positive definite invertible matrices in $M_2(B_{p,\infty})$ and their "connecting matrices" on the other hand. The goal of this section is to incorporate CM orientations and Shimura class group actions in this framework and provide applications. Omitted proofs can be found in Appendix H. Note: in this section the letter g is used to denote an IKO matrix, rather than the dimension of an abelian variety, which is assumed 2 throughout.

5.1 The IKO correspondence: a slight extension

Fix a supersingular elliptic curve E in characteristic $p > 0$ along with an isomorphism $\vartheta : \mathcal{O} \to \mathrm{End}(E)$ from some maximal order $\mathcal{O} \subset B_{p,\infty}$, as at the beginning of Section 3. By Deligne's theorem [41] any superspecial PPAS is isomorphic to (A, λ) for some principal polarization $\lambda : A \to A^\vee$ on $A = E^2$. Let λ_0 denote the product polarization. Then the automorphism $\lambda_0^{-1} \circ \lambda$ corresponds under ϑ (when applied component-wise) to some totally positive Hermitian norm-1 matrix, i.e.,

$$g = \begin{pmatrix} s & r \\ \bar{r} & t \end{pmatrix}, \qquad s, t \in \mathbb{Z}_{>0},\ r \in \mathcal{O},\ st - r\bar{r} = 1;$$

see also Example A.5. We call g the g-matrix or IKO-matrix associated with (A, λ). To identify equivalent principal polarizations λ, λ', i.e., for which $(A, \lambda) \cong (A, \lambda')$, one should work with g-matrices up to equivalence, where we say $g' \sim g$ if and only if $g' = u^* g u$ for some $u \in \mathrm{GL}_2(\mathcal{O})$.

Since we are just considering A with different principal polarizations, we always have $\mathrm{End}(A) \cong M_2(\mathcal{O})$ and the piece of data that varies with λ is the Rosati involution. We know that the Rosati involution for the product polarization λ_0 is just conjugate-transpose $b \mapsto b^*$. In general, for a superspecial PPAS corresponding to a matrix g, the Rosati involution is $b \mapsto b^\dagger = g^{-1} b^* g$ which follows from the definition of the g-matrix and the definition of the Rosati involution. In that respect, superspecial surfaces can also be characterized by positive involutions on $M_2(\mathcal{O})$.

Isogenies are incorporated in the IKO correspondence as follows. Let λ_1, λ_2 be principal polarizations on A with g-matrices g_1, g_2 and let $\mu \in M_2(\mathcal{O})$ be a matrix fixed by the Rosati involution with respect to g_1 that is totally positive, i.e., the (necessarily real) roots of its minimal polynomial over $\mathbb{Q}$ are positive. Then $\gamma \in M_2(\mathcal{O})$ corresponds (under ϑ) to a $\vartheta(\mu)$-isogeny $(A, \lambda_1) \to (A, \lambda_2)$ if and only if

$$\begin{aligned} \lambda_1 \circ \vartheta(\mu) = \widehat{\vartheta(\gamma)} \circ \lambda_2 \circ \vartheta(\gamma) &\Leftrightarrow \lambda_0^{-1} \circ \lambda_1 \circ \vartheta(\mu) = \lambda_0^{-1} \circ \widehat{\vartheta(\gamma)} \circ \lambda_0 \circ \lambda_0^{-1} \circ \lambda_2 \circ \vartheta(\gamma) \\ &\Leftrightarrow g_1 \mu = \gamma^* g_2 \gamma. \end{aligned}$$

This (slightly) extends the corresponding calculation from [13, §2.3], which was restricted to $\mu = N\mathrm{Id}_2$ for positive integers N.

Remark 5.1. For $\gamma = \mathrm{Id}_2$ we have the following identity as a special case: $g_1 \mathrm{Id}_2 = \mathrm{Id}_2^* g_1 \mathrm{Id}_2$. An interesting aspect of this trivial equation is that it says both that (A, λ_1) and (A, λ_1) are 1-isogenous (trivially), and that (A, λ_0) and (A, λ_1) are $\vartheta(g_1)$-isogenous. Since this reasoning applies to arbitrary g_1, this shows the great impact of the choice of μ (e.g., when specified as the second component of a Shimura ideal) on the resulting codomain as a PPAV.

So this means that if λ_1 is a principal polarization on A with g-matrix g_1 and if $H \subset A$ is a finite subgroup that is maximal with the property that

$$e_{\vartheta(\mu)}(P, \lambda_1(Q)) = 1, \quad \forall P, Q \in H,$$

then a g-matrix for the induced principal polarization on $A \cong A/H$ is

$$g_2 := \gamma^{-*} g_1 \mu \gamma^{-1},$$

with $\gamma \in M_2(\mathcal{O})$ any matrix such that $\ker \vartheta(\gamma) = H$. This follows, e.g., from Proposition A.4. In particular, g_2 is necessarily a valid g-matrix, i.e., a positive-definite Hermitian matrix of norm 1 contained in $M_2(\mathcal{O})$.

Remark 5.2. Most of these properties are easily confirmed:

- Using $\mu^\dagger = g_1^{-1}\mu^* g_1$ one checks that $g_2^* = g_2$.
- By definition of a $\vartheta(\mu)$-isogeny, the degree of $\vartheta(\mu)$ equals the squared degree of $\vartheta(\gamma)$, hence in view of [14, Lem. 2.10] we see that g_2 has norm 1.
- It concerns a positive-definite matrix: for this it suffices to see that $\gamma^* g_2 \gamma = g_1\mu$ is positive-definite, and this in turn follows from the total positivity of μ and the fact that after base-changing to $B_{p,\infty} \otimes_{\mathbb{Q}} \mathbb{C} \cong M_2(\mathbb{C})$ if needed one can rewrite $g_1\mu = h(h^*\mu h)h^{-1}$ with $hh^* = g_1$. The existence of such an h follows from Cholesky decomposition, which we have included as Lemma 5.3 below.

Verifying that $g_2 \in M_2(\mathcal{O})$ in an elementary way is more intricate. For instances of H that arise as the kernel of a Shimura ideal, such a verification can be found in Appendix H.1; see also Remark 5.5.

Lemma 5.3 (Cholesky decomposition in $M_2(B_{p,\infty})$). *Let $g \in M_2(B_{p,\infty})$ be a positive-definite Hermitian matrix, i.e., $g = g^*$ and its (necessarily real) eigenvalues are positive. Then there exists $h \in M_2(B_{p,\infty})$ such that $hh^* = g$.*

5.2 Orientations of g-matrices and the Shimura class group action

Let K be a quartic CM field and let $R \subset K$ be a CM order. We now describe K-oriented superspecial PPASs (A, λ, ι) within the framework of the IKO correspondence, where as above we assume $A = E^2$. Let g be the g-matrix associated with λ and let $\iota : R \hookrightarrow \mathrm{End}(A) \cong M_2(\mathcal{O})$ be a primitive orientation. Write $R_0 = \mathbb{Z}[a]$ for some real generator a and $R = R_0[b_1, b_2]$ for generators $b_i \in R$.[12] Upon identifying a, b_1, b_2 with their images under the orientation, we can think of them as commuting elements of $M_2(\mathcal{O})$ satisfying $g^{-1}a^* g = a$ and $g^{-1}b_i^* g \in R$ for $i = 1, 2$. We refer to the tuple $g, a, b_1, b_2 \in M_2(\mathcal{O})$ as algebraic data associated with (A, λ, ι).

Theorem 5.4. *Assuming the generalized Riemann hypothesis (GRH), there is a polynomial-time algorithm which upon input of*

- *algebraic data g, a, b_1, b_2 of an oriented superspecial PPAS (A, λ, ι),*

[12] In many cases one can do with a single generator b_i: this is true if and only if R is a free R_0-module. E.g., for this it suffices that R_0 is maximal and $h(K_0) = 1$.

– *an integral Shimura ideal* $(\mathfrak{a}, \mu)$ *of* $\mathbb{Z}[a, b_1, b_2]$ *such that* $p \nmid N(\mathfrak{a})$,[13]

returns algebraic data of the codomain of the corresponding $\iota(\mu)$*-isogeny, equipped with the induced principal polarization and orientation.*

Proof. The endomorphisms that vanish on the kernel of $\mathfrak{a}$ form a left ideal of $M_2(\mathcal{O})$ generated by the elements of $\mathfrak{a}$. A generator γ of this (principal) ideal can be found in polynomial time, using either [14] or [78, §4]; the former algorithm is more efficient but the latter method is heuristic-free (modulo a dependency on GRH). As explained in the previous section, the g-matrix of the codomain is then $g' = \gamma^{-*} g \mu \gamma^{-1}$. From the requirement that γ should represent an oriented isogeny, it is then clear that the push-forward orientation is given by $a' = \gamma a \gamma^{-1}$, $b_i' = \gamma b_i \gamma^{-1}$. All this data is straightforward to compute. □

Remark 5.5. It is instructive to try and *define* the Shimura class group action (on superspecial PPASs) via $g \mapsto \gamma^{-*} g \mu \gamma^{-1}$, $a \mapsto \gamma a \gamma^{-1}$, $b_i \mapsto \gamma b_i \gamma^{-1}$ and then give a direct matrix-arithmetic proof that this action is well-defined. The main tasks here are showing that $g', a', b_i' \in M_2(\mathcal{O})$, that a' is fixed under the Rosati involution induced by g', and that $\mathbb{Z}[a', b_1', b_2']$ is stable under this Rosati involution. The details of such a direct approach can be found in Appendix H.1.

All explicit formulae for isogeny evaluation (that we are aware of) restrict to N-isogenies for some $N \in \mathbb{Z}_{>0}$. In view of this, we have the following important consequence of Theorem 5.4: if a g-matrix of (A, λ) is known, then we can act with arbitrary Shimura ideal classes, even if they do not belong to $\mathfrak{C}(R)_{\mathbb{Z}}$, in polynomial time. (Note: here we assume that (A, λ) is given to us as a Jacobian with the canonical polarization, or as a product of elliptic curves with the product polarization.) We point out that knowing a g-matrix of (A, λ) is polynomial-time equivalent to knowing the endomorphism ring of A, in the strong sense of [73, Def. 2.14]; see [73, §3] for details.

Corollary 5.6. *Let R be a CM order in a quartic CM field and let p be a prime number. There exists a heuristic polynomial-time algorithm for computing $[(\mathfrak{a}, \mu)] \star (A, \lambda, \iota)$ for all $[(\mathfrak{a}, \mu)] \in \mathfrak{C}(R)$ and $(A, \lambda, \iota) \in \mathcal{M}_{\overline{\mathbb{F}}_p}(R)$, assuming that a g-matrix of (A, λ) is known.*

Proof. We assume that the orientation ι is given in higher-dimensional representation as in Section 4.3, but the argument should carry over to any other valid way of representing ι. From a g-matrix of (A, λ) we can compute an explicit isomorphism $\varphi : (A, \lambda) \to (E^2, \lambda_0 \circ g)$ using [73, Thm. 5.3]. This is then also an oriented isomorphism between (A, λ, ι) and $(E^2, \lambda_0 \circ g, \varphi_* \iota)$, and the given orientation data are easily pushed through and interpolated to obtain algebraic data g, a, b_1, b_2. Algebraic data g', a', b_1', b_2' for $[(\mathfrak{a}, \mu)] \star (A, \lambda, \iota)$ is then obtained using Theorem 5.4. Using the IKO-to-surface conversion algorithm from [13, Thm. 5.1] we can realize this codomain as a Jacobian (or a product of elliptic

[13] We included the assumption $p \nmid N(\mathfrak{a})$ for consistency with Theorem 2.7. But, in fact, our algebraic treatment works equally well in the inseparable case.

curves with the product polarization), where an explicit isomorphism φ' can be found through another use of [73, Thm. 5.3]. The algebraic data a', b'_1, b'_2 can then be converted into interpolation data and pushed through φ' to obtain the desired output. □

Remark 5.7. Corollary 5.6 is heuristic due to several calls to the powersmooth KLPT^2 algorithm from [13], which relies on some heuristics. Note that the improved KLPT^2 algorithm from [78] can likely be made heuristic-free (perhaps under GRH; this would work with very similar ideas as in the case of 1-dimensional KLPT [97]).

5.3 Constructing g-matrices with a given orientation

An important application of the algebraic viewpoint is the construction of superspecial PPASs with a prescribed orientation. Indeed, on the geometry side there is no systematic way to find orientations other than ad-hoc constructions, as in Section 3 or Appendix C, or CM reductions, which only provide orientations by orders with a relatively small discriminant.

As an entry point, we consider the next lemma:

Lemma 5.8. *Let g be a g-matrix and let $a \in M_2(\mathcal{O}) \setminus \mathbb{Q}\mathrm{Id}_2$ be a symmetric element with respect to Rosati. Then its centralizer $\mathrm{Cent}(a) \subset M_2(B_{p,\infty})$ is isomorphic to the quaternion algebra $C = \mathbb{Q}(a) \otimes B_{p,\infty}$ over $\mathbb{Q}(a)$, and Rosati acts as quaternion conjugation on C.*

This shows that one can obtain orientations as follows. Take a symmetric element $a \in M_2(\mathcal{O})$ and consider the intersection $M_2(\mathcal{O}) \cap \mathrm{Cent}(a)$. This intersection is invariant under Rosati as both $M_2(\mathcal{O})$ and $\mathrm{Cent}(a)$ are. Then taking quadratic subrings of $M_2(\mathcal{O}) \cap \mathrm{Cent}(a)$ will yield orientations. This is reminiscent of the 1-dimensional case where orientations are simply quadratic subrings of maximal orders. We note that intersections of the form $M_2(\mathcal{O}) \cap \mathrm{Cent}(a)$ were intensively studied in [75] where it was shown that they are either maximal orders or Eichler orders of level p in the quaternion algebra $\mathbb{Q}(a) \otimes B_{p,\infty}$ as soon as p is unramified in $\mathbb{Q}(a)$ [75, Thm. 2.5.27].

Our goal is an algorithm for creating superspecial PPASs that are primitively oriented by a prescribed CM order R. For convenience we assume $R = \mathbb{Z}_K$. We expect that the method below can be tweaked so that it covers a much more general class of CM orders, but this would add a layer of technicality that lies beyond our current scope. Note that primitivity comes for free if $R = \mathbb{Z}_K$, so we are merely concerned with finding a K-oriented superspecial PPAS (A, λ, ι) such that $\mathbb{Z}_K \subset \iota^{-1}\,\mathrm{End}(A)$.

Our approach is to generalize the 1-dimensional blueprint, where one takes an element in $B_{p,\infty}$ of prescribed minimal polynomial and then puts that element into a maximal order (after which one finds a corresponding elliptic curve using the effective Deuring correspondence). Here the situation is somewhat more complex as one has to embed an element of $M_2(B_{p,\infty})$ into a maximal order of $M_2(B_{p,\infty})$ in a way that is compatible with the Rosati involution; this

compatibility is automatic in the 1-dimensional case. First we prove two useful lemmas. The first lemma can be thought of as a base-point-free orientation framework. Indeed, it is sometimes inconvenient (especially when trying to generalize 1-dimensional proofs) to work with g-matrices, and rather use the following equivalent framework.

Lemma 5.9. *Let Λ be a maximal order in $M_2(B_{p,\infty})$ that is stable under conjugate-transpose. Then there exists an isomorphism of rings with involutions between Λ with conjugate-transpose and $M_2(\mathcal{O})$ with some Rosati involution induced by a g-matrix. This isomorphism can be computed in polynomial time.*

Proof. First, we find γ such that $M_2(\mathcal{O}) = \gamma^{-1}\Lambda\gamma$. Such a γ can be found in polynomial time using [78] or [14] as it is equivalent to a principal ideal problem. One can assume that the reduced norm of γ is coprime to p. As $M_2(\mathcal{O})$ is closed under conjugate-transpose one has that $(\gamma^{-1}\Lambda\gamma)^* = \gamma^{-1}\Lambda\gamma$. This can be rearranged as $\gamma\gamma^*\Lambda = \Lambda\gamma\gamma^*$ which means that $\gamma\gamma^*$ generates a two-sided ideal of Λ, hence $(\gamma\gamma^*) = (N)$ for some integer N. This implies that $\gamma\gamma^*/N$ has norm 1 and hence the same holds for $\gamma^*\gamma/N$. Call this latter matrix g and note that it indeed concerns a g-matrix. Then it is easy to see that $(\Lambda, *)$ is isomorphic as a ring with involution to $(M_2(\mathcal{O}), \dagger_g)$ where $\dagger_g$ denotes the Rosati involution induced by g. □

The second lemma is a very generic fact from the algorithmic theory of quaternion algebras (and might already be in the literature but we could not find a reference). Let C be a quaternion algebra over a field K, not of characteristic 2. It is a well-known result that C is split by a quadratic extension $K(\sqrt{d})$, i.e., $C \otimes_K K(\sqrt{d}) \cong M_2(K(\sqrt{d}))$, if and only if C contains $K(\sqrt{d})$ as a subfield. The next lemma shows that if one can find an explicit isomorphism between $C \otimes_K K(\sqrt{d})$ and $M_2(K(\sqrt{d}))$, then one can find a subfield of C that is isomorphic to $K(\sqrt{d})$:

Lemma 5.10. *Let K be a number field. Let C be a quaternion algebra over K given by a quaternion basis over K and let $d \in K \setminus K^2$. Then one can decide whether $K(\sqrt{d})$ embeds into C in quantum polynomial time. Furthermore, if it embeds, one can find an explicit embedding in quantum polynomial time.*

Remark 5.11. Let us emphasize that the quantum part is not necessary in practice, but it makes the lemma heuristic-free. Indeed, our proof of Lemma 5.10 in Appendix H.2 invokes a quantum algorithm from [60] but this can be replaced with a heuristic algorithm for solving quadratic forms.

We now proceed to finding $\mathbb{Z}_K$-oriented superspecial PPASs. The following is a first step:

Lemma 5.12. *Given a quartic CM field K, there exists a quantum polynomial-time algorithm that finds an embedding $\iota : K \hookrightarrow M_2(B_{p,\infty})$ such that $\iota(\bar{\alpha}) = \iota(\alpha)^*$ for all $\alpha \in K$, or decides that no such embedding exists.*

Proof. We first find an element $a \in M_2(B_{p,\infty})$ such that a is symmetric with respect to conjugate-transpose and $\mathbb{Q}(a) \cong K_0$. We will essentially show how one can generate an element of given trace and norm for symmetric matrices. A symmetric matrix has the following form:

$$\begin{pmatrix} s & r \\ \bar{r} & t \end{pmatrix}$$

where $s, t \in \mathbb{Q}$ and $r \in B_{p,\infty}$. Let the trace of a be M and its norm be N. This gives us two equations: $s + t = M,\ st - \mathrm{n}(r) = N$. Such s, t, r can be found by reducing to a quadratic form in 6 variables that always admits a solution and can be solved with Castel's algorithm [12] in polynomial time (this is a variation of an algorithm due to Simon that does not require factoring). Now we find the CM field K inside $\mathrm{Cent}(a)$, or decide that it cannot be found, using Lemma 5.10 with $C = \mathrm{Cent}(a)$ and $K = \mathbb{Q}(a)$. □

Remark 5.13 (continuation of Remark 5.11). Here too the word "quantum" can be omitted at the cost of introducing some heuristics.

Our main technical tool for converting this into an algorithm for finding $\mathbb{Z}_K$-orientations is the following:

Lemma 5.14. *Let $K \subset M_2(B_{p,\infty})$ be a quartic CM field such that complex conjugation in K amounts to conjugate-transpose in $M_2(B_{p,\infty})$ and assume that the conductor ideal in $\mathbb{Z}_K$ of the order $\mathbb{Z}_K \cap M_2(\mathcal{O})$ is principal and generated by a totally positive $\mu \in \mathbb{Z}_{K_0}$ that is coprime to p. Let $\xi \in M_2(B_{p,\infty})$ be such that $\xi\xi^* = \mu$. Let $\gamma \in M_2(B_{p,\infty})$ be such that $M_2(\mathcal{O})\gamma$ equals the left fractional ideal $M_2(\mathcal{O})\mathbb{Z}_K$ generated by $\mathbb{Z}_K$. Let $\Lambda = (\gamma\xi)^{-1}M_2(\mathcal{O})\gamma\xi$. Then $\Lambda^* = \Lambda$ and*

$$\iota : \mathbb{Z}_K \to \Lambda : \alpha \mapsto \xi^{-1}\alpha\xi$$

is an injective ring homomorphism such that $\iota(\bar{\alpha}) = \iota(\alpha)^*$ for all $\alpha \in \mathbb{Z}_K$.

Corollary 5.15. *If we assume that $h^+(K_0)$ is polynomially bounded,*[14] *there exists a heuristic polynomial-time algorithm which upon input of a quartic CM field K returns a primitively $\mathbb{Z}_K$-oriented superspecial PPAS in characteristic p, or returns that no such surface exists.*

Proof. Repeatedly run Lemma 5.12 until the conductor ideal $\mathfrak{c}$ of $\mathbb{Z}_K \cap M_2(\mathcal{O})$ is generated (as an ideal in $\mathbb{Z}_K$) by some totally positive $\mu \in \mathbb{Z}_{K_0}$ whose norm is not a multiple of p. We expect this to happen with good probability. Indeed,

[14] This is typically a mild assumption. By Lemma E.4 we have $h^+(K_0)/h(K_0) \in \{1, 2, 4\}$. The Cohen–Lenstra heuristics [24] predict that the odd part of $h(K_0)$ is extremely small, e.g., it is just 1 for about 76% of all real quadratic number fields. By an extension of the Cohen–Lenstra heuristics (see [57, §1] for a discussion) the even part of $h_0(K_0)$ comes almost entirely from $\mathrm{Cl}(K_0)[2]$, which by genus theory is roughly of size 2^t with t the number of prime factors of $\operatorname{disc} K_0$. In view of the Erdös–Kac theorem this number is typically of size $\log \operatorname{disc} K_0$.

since $\bar{\mathfrak{c}} = \mathfrak{c}$ we expect that $\mathfrak{c} = \mathfrak{c}_0\mathbb{Z}_K$ for some ideal $\mathfrak{c}_0 \subset \mathbb{Z}_{K_0}$; the only exception here is where $\mathfrak{c}$ has a factor $\mathfrak{p}^e$ with e odd and $\mathfrak{p}$ a prime ideal that is ramified over K_0. We then expect $\mathfrak{c}_0$ to be generated by some totally positive μ with probability about $1/h^+(K_0)$. Using the previous technical lemma we can find an embedding of $\mathbb{Z}_K$ inside a self-conjugate maximal order $\Lambda \subset M_2(B_{p,\infty})$ in a way that is compatible with Rosati. Note that all ingredients can be computed efficiently, including the Cholesky decomposition from Lemma 5.3 which involves solving quadratic equations over $\mathbb{Q}$ in 5 variables, which can be done efficiently using Castel's algorithm [12]. Now invoking Lemma 5.9 we get an orientation for some g-matrix in $M_2(\mathcal{O})$. This can be converted into a K-oriented superspecial PPAS as in the proof of Corollary 5.6. □

5.4 Applications

In Appendix I we discuss applications of the previously described algebraic machinery to the "one endomorphism problem". In dimension 1 it is known that knowing one non-scalar endomorphism provides significant extra information:

- Given one non-scalar endomorphism one can compute the entire endomorphism ring in sub-exponential time [45].
- Given an endomorphism of small degree (polylogarithmic in p), one can compute the full endomorphism ring in polynomial time [66].
- Given an oracle for computing non-scalar endomorphisms, one can compute the endomorphism ring in polynomial time [79].

We investigate the first two problems in dimension 2, where we observe that adjustments have to be made, as there are non-scalar (even polarized) endomorphisms that will likely not provide enough information for computing the full endomorphism ring: an example is

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

on E^2 for any supersingular elliptic curve E. The proper analogue of a non-scalar endomorphism in dimension 2 is one whose minimal polynomial defines a quartic CM field over $\mathbb{Q}$ (a "generic" endomorphism satisfies this).

6 Orbits of the Shimura class group action

It is classical that if $g = 1$ and $p = 0$ then the action $\star : \mathrm{Cl}(R) \times \mathcal{M}_k(R) \to \mathcal{M}_k(R)$ is "essentially" transitive, although our algebraic treatment of orientations comes with the subtlety, emphasized by Onuki [76], that there are in fact two orbits. One can jump back and forth between these orbits via $(E, \iota) \mapsto (E, \iota \circ \sigma)$, with $\sigma : \alpha \mapsto \overline{\alpha}$ denoting complex conjugation on R. In positive characteristic, on the other hand, we know that if $p > 0$ and $p \mid [\mathbb{Z}_K : R]$ then $\mathcal{M}_k(R)$ is empty, while if $p \nmid [\mathbb{Z}_K : R]$ then:

(i) If p splits in K then all curves in $\mathcal{M}_k(R)$ are ordinary and one can jump between the two orbits by quadratic twisting.

(ii) If p ramifies in K then all curves in $\mathcal{M}_k(R)$ are supersingular; here the action becomes transitive.

(iii) If p is inert in K then all curves in $\mathcal{M}_k(R)$ are supersingular, but in this case there are again two orbits. Instead of $(E,\iota) \mapsto (E, \iota\circ\sigma)$, there is a neat alternative way to jump between the orbits due to Onuki: it can be assumed that all curves are defined over $\mathbb{F}_{p^2}$, and then one can jump back and forth using the p-power Frobenius map $(E,\iota) \mapsto (E^{(p)}, \iota^{(p)})$.

This can be assembled from [61, §13, Thm. 5 & Thm. 12] and [76, Prop. 3.3].

In higher dimension this picture is much more subtle.[15] The literature does not seem to contain the ingredients needed for a complete answer, which in any case lies beyond the scope of our paper. But we can sketch the overall picture, highlighting the main ingredient: the CM type. Behind the scenes, this notion is already responsible for the (typically) two orbits in the elliptic curve case. Much of what follows is standard to specialists: our main goal is to illustrate the theory with examples that are relevant to the paper at hand.

6.1 An obstruction to transitivity

Our starting point is the ring homomorphism $\varrho_{(A,\lambda,\iota)} : R \to \mathrm{End}_k(D(A)) : \alpha \mapsto \iota(\alpha)^*$, where $\iota(\alpha)^*$ denotes the pull-back map under $\iota(\alpha)$ on the g-dimensional k-vector space $D(A) = H^0(A, \Omega^1_A)$ of regular 1-forms on A. This homomorphism gives $D(A)$ the structure of an R-module, in the sense of representation theory.

Lemma 6.1. *If $(A,\lambda,\iota), (A',\lambda',\iota') \in \mathcal{M}_k(R)$ lie in the same orbit under the action of $\mathfrak{C}(R)$ then $D(A) \cong D(A')$ as R-modules.*

Proof. Every Shimura ideal class has a representative $(\mathfrak{a},\mu)$ such that $N(\mathfrak{a})$ is not a multiple of p. Then the isogeny $\varphi_\mathfrak{a} : A \to A'$ is separable, so the map $\varphi_\mathfrak{a}^* : D(A') \to D(A)$ is an isomorphism of k-vector spaces. The diagram

$$\begin{array}{ccc} D(A') & \xrightarrow{\iota'(\alpha)^*} & D(A') \\ {\scriptstyle\varphi_\mathfrak{a}^*}\downarrow & & \downarrow{\scriptstyle\varphi_\mathfrak{a}^*} \\ D(A) & \xrightarrow[\iota(\alpha)^*]{} & D(A) \end{array}$$

commutes for every $\alpha \in R$, because $\varphi_\mathfrak{a}$ is oriented. This indeed amounts to saying that $\varphi_\mathfrak{a}^*$ is an isomorphism of R-modules. □

Remark 6.2. The lemma generalizes to non-horizontal isogenies: for any two CM orders $R, R' \subset \mathbb{Z}_K$ such that $p \nmid [\mathbb{Z}_K : R][\mathbb{Z}_K : R']$ it holds that if $(A,\lambda,\iota) \in \mathcal{M}_k(R)$, $(A',\lambda',\iota') \in \mathcal{M}_k(R')$ are connected by a separable oriented isogeny, then $D(A), D(A')$ are isomorphic as $\mathbb{Z}_K$-modules (the assumption $p \nmid [\mathbb{Z}_K : R]$ naturally allows to extend the domain of $\varrho_{(A,\lambda,\iota)}$ to $\mathbb{Z}_K$, and likewise for R').

[15] E.g., from Remark C.6 we see that $\mathcal{M}_k(R)$ may be non-empty even when $p \mid [\mathbb{Z}_K : R]$.

Example 6.3. We start by shedding a different light on Onuki's motivating example from [76, p4]. Let $R = \mathbb{Z}[i]$ be the ring of Gaussian integers, assume $p \neq 2$ and let $\mathbf{i} \in k$ be a square root of -1. Consider $E : y^2 = x^3 + x$ over k along with the primitive R-orientations

$$\iota : i \mapsto ((x, y) \mapsto (-x, \mathbf{i}y)), \qquad \iota' : i \mapsto ((x, y) \mapsto (-x, -\mathbf{i}y))$$

on E. The space $D(E)$ is generated by dx/y, which under $\iota(i)$ pulls back to $d(-x)/\mathbf{i}y = (-1/\mathbf{i})dx/y = \mathbf{i}dx/y$ while under $\iota'(i)$ it pulls back to $-\mathbf{i}dx/y$. Thus the representations map i to the multiplication maps by $\mathbf{i}$ and $-\mathbf{i}$, respectively. Therefore (E, ι), (E, ι') must lie in different orbits. In general, this gives a very efficient method for testing whether, in case *(iii)* above, two given oriented supersingular elliptic curves lie in the same orbit or not. In other words, the decisional version of the "effective R-vectorization problem" from [96, Prob. 2] is very easy to solve.

Example 6.4. Let E_1, E_2 be elliptic curves over k and let $A = E_1 \times E_2$. Then $D(A) \cong D(E_1) \oplus D(E_2)$. So, upon picking a generator ω_i of $D(E_i)$ for $i = 1, 2$ we can view $(\omega_1, 0), (0, \omega_2)$ as a basis for $D(A)$. The elements of $\mathrm{End}(A)$ can be viewed as matrices

$$\begin{pmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{pmatrix}, \qquad \alpha_{ij} \in \mathrm{Hom}(E_i, E_j).$$

With respect to our basis of $D(A)$, the matrix of the pull-back operator along such a matrix $(\alpha_{ji})_{i,j}$ is given by $(c_{ij})_{i,j}$, where $c_{ij} \in k$ is the constant such that $\alpha_{ij}^* \omega_j = c_{ij}\omega_i$. Let us have a look at what this implies for the two orientations ι, ι' from Examples 3.1 and 3.2.

Using that differentials pull back to zero under inseparable isogenies, i.e., $\pi_p^* \omega_i = 0$ for $i = 1, 2$, we find that $\varrho_{(A,\lambda,\iota)}$ maps

$$\frac{1 + \sqrt{-m}}{2} \mapsto \begin{pmatrix} \frac{1}{2} & \mathbf{i} \\ \mathbf{i} & \frac{1}{2} \end{pmatrix}, \qquad \sqrt{d} \mapsto \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}.$$

When rewritten with respect to the basis $(\omega_1, \omega_2), (\omega_1, -\omega_2)$ this becomes

$$\frac{1 + \sqrt{-m}}{2} \mapsto \begin{pmatrix} \frac{1}{2} + \mathbf{i} & 0 \\ 0 & \frac{1}{2} - \mathbf{i} \end{pmatrix}, \qquad \sqrt{d} \mapsto \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix},$$

showing that $\varrho_{(A,\lambda,\iota)}$ arises as a direct sum of two 1-dimensional representations, mapping

$$\begin{aligned} x_1 + x_2\frac{1 + \sqrt{-m}}{2} + x_3\sqrt{d} + x_4\sqrt{d}\frac{1 + \sqrt{-m}}{2} \quad &\mapsto \\ x_1 + (\frac{1}{2} \pm \mathbf{i})x_2 \pm \frac{1}{2}x_3 + \frac{1}{2}(\mathbf{i} \pm \frac{1}{2})x_4 \end{aligned} \tag{6}$$

for all $x_i \in \mathbb{Z}$. A similar reasoning for $\varrho_{(A,\lambda,\iota')}$ gives

$$\frac{1 + \sqrt{-m}}{2} \mapsto \begin{pmatrix} \frac{1}{2} + \mathbf{i} & 0 \\ 0 & \frac{1}{2} + \mathbf{i} \end{pmatrix}, \qquad \sqrt{d} \mapsto \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

when expressing the images with respect to $(b\omega_1, (1/2 - a)\omega_2), (b\omega_1, (-1/2 - a)\omega_2)$. Thus we again have a direct sum of two 1-dimensional representations, namely

$$x_1 + (\frac{1}{2} + \mathbf{i})x_2 \pm \frac{1}{2}x_3 \pm \frac{1}{2}(\frac{1}{2} + \mathbf{i})x_4 \tag{7}$$

From (6) and (7) we see that ι, ι' give rise to non-isomorphic R-module structures on $D(A)$. So by Lemma 6.1 we must be in different orbits under the $\mathfrak{C}(R)$-action.

We emphasize that there is an important difference between the obstructions from Example 6.3 and Example 6.4. In the former case, precomposition of ι' with complex conjugation removes the obstruction. For the second example this is not true, nor is it possible to achieve this using another automorphism of R: the obstruction to transitivity is really more fundamental here. This is related to the notion of "inequivalent" CM types discussed in the next section.

Example 6.5. Revisiting the $\mathbb{Z}[\zeta_5]$-orientation from Example 3.3, we have $D(A) \cong H^0(C, \Omega^1_C) = \langle dx/y, xdx/y \rangle$. Through

$$\iota(\zeta_5)^* \frac{dx}{y} = \frac{d(\xi_5 x)}{y} = \xi_5 \frac{dx}{y}, \qquad \iota(\zeta_5)^* x\frac{dx}{y} = (\xi_5 x)\frac{d(\xi_5 x)}{y} = \xi_5^2 x\frac{dx}{y}$$

we see that, here too, $\varrho_{(A,\lambda,\iota)}$ decomposes into two 1-dimensional representations, mapping ζ_5 to multiplication-by-ξ_5 and multiplication-by-ξ_5^2, respectively.

6.2 CM types in characteristic $p = 0$

In the above examples, the decomposability of $\varrho_{(A,\lambda,\iota)}$ into 1-dimensional sub-representations $\varrho_i : R \to k$, $i = 1, \ldots, g$, is not a coincidence. Moreover, these 1-dimensional representations are not arbitrary. In characteristic $p = 0$ this is classical and captured by the notion of CM type. In this case the ϱ_i naturally extend their domains to all of K and take values in some subfield $L \subset k$ that is isomorphic to a Galois closure of K. It can be argued [93] that the resulting size-g subset of $\mathrm{Hom}(K, L)$ is a CM type, in the following sense:

Definition 6.6. *Let K be a CM field and let $L \subset \mathbb{C}$ be isomorphic to a Galois closure of K. A* CM type *of K is a subset of* $\mathrm{Hom}(K, L)$ *containing exactly one element from each pair of complex conjugate embeddings $\varphi, \bar{\varphi} : K \to L$.*

E.g., this pattern is clearly visible in Example 6.5.

Famously, it was proven by Shimura and Taniyama that for $p = 0$ and under mild assumptions on R, the CM type is the only obstruction to transitivity. One sufficient condition is that $R = \mathbb{Z}_K$,[16] see [93, Thm. 5.2]:

[16] The "right" sufficient condition reads that R is Gorenstein, which is implied, e.g., by the existence of an index-2 subfield $K' \subset K$ such that $R \cap K'$ is a maximal order [4] (e.g., maximal real multiplication is a sufficient condition).

Theorem 6.7 (Shimura–Taniyama, 1961). *Consider a CM field K and let (A,λ,ι), $(A',\lambda',\iota') \in \mathcal{M}_k(\mathbb{Z}_K)$ have the same CM types; here k denotes an algebraically closed field of characteristic 0. There exists a unique $[(\mathfrak{a},\mu)] \in \mathfrak{C}(\mathbb{Z}_K)$ such that $(A',\lambda',\iota') = [(\mathfrak{a},\mu)] \star (A,\lambda,\iota)$.*

The number of CM types of K is easily seen to be 2^g, so under the aforementioned mild assumptions on R this is our best guess for the number of orbits in characteristic 0. However we note that, very exceptionally, the number of orbits may be smaller because certain CM types may not occur, even when $R = \mathbb{Z}_K$.[17] Another important difference with the elliptic curve case is that it may be impossible to jump between orbits by precomposing the orientation with an automorphism of K. This comes from the existence of inequivalent CM types: any automorphism τ of K commutes with complex conjugation (this is because K_0 can be characterized as the maximal totally real subfield of K) and therefore, if Φ is a CM type, then so is $\Phi' = \Phi \circ \tau$, in which case Φ, Φ' are said to be equivalent. In our main case of interest $g = 2$ we have from [93, Lem. 3.4]:

Lemma 6.8 (Inequivalent CM types for $g = 2$). *Let K be a CM field of degree 4 with Galois closure L and let $G = \mathrm{Gal}(L,\mathbb{Q})$. Let $\mathrm{Hom}(K,L) = \{\mathrm{id}, \bar{\mathrm{id}}, \varphi, \bar{\varphi}\}$. Write $\Phi = \{\mathrm{id}, \varphi\}$ and $\Phi' = \{\mathrm{id}, \bar{\varphi}\}$. Then:*

- *if $G \cong C_2 \times C_2$ or $G \cong D_4$ then Φ and Φ' are inequivalent,*
- *if $G \cong C_4$ then Φ, Φ' are equivalent.*

So the number of inequivalent CM types in these cases is 2, 2, 1, respectively.

6.3 Back to positive characteristic

We now move to our main case of interest $k = \overline{\mathbb{F}}_p$. We expect that under mild assumptions on R, the obstruction to transitivity from Lemma 6.1 should be the only one. Again we formulate this for maximal orders only:[18]

Theorem 6.9. *Let $k = \overline{\mathbb{F}}_p$ and let K be a CM field. Let $(A,\lambda,\iota), (A',\lambda',\iota') \in \mathcal{M}_k(\mathbb{Z}_K)$ be such that $\varrho_{(A,\lambda,\iota)} \cong \varrho_{(A',\lambda',\iota')}$. Then there exists a unique $[(\mathfrak{a},\mu)] \in \mathfrak{C}(\mathbb{Z}_K)$ such that $(A',\lambda',\iota') = [(\mathfrak{a},\mu)] \star (A,\lambda,\iota)$.*

Proof. It was shown in [43, Lem. 4.4.1] that every $(A,\lambda,\iota) \in \mathcal{M}_k(\mathbb{Z}_K)$ admits a lift to characteristic 0. As explained in their proof, the possible CM types of such a lift are controlled by Kottwitz' Φ-determinant condition: explicitly, fixing an integral basis $\alpha_1, \alpha_2, \ldots, \alpha_{2g}$ of $\mathbb{Z}_K$, for a CM type Φ one can consider the so-called Kottwitz polynomial

$$f_\Phi \;=\; \prod_{\varphi \in \Phi} \left(x_1\varphi(\alpha_1) + x_2\varphi(\alpha_2) + \ldots + x_{2g}\varphi(\alpha_{2g})\right) \;\in \mathbb{Z}_{K^*}[x_1, x_2, \ldots, x_{2g}].$$

[17] This can only happen if the last map in (2) is not surjective, see the proof of [51, Lem. 19]. In dimension $g = 2$ all such examples come from the $C_2 \times C_2$ case.

[18] Here too, it is likely that this assumption can be relaxed significantly, perhaps to R Gorenstein and $p \nmid [\mathbb{Z}_K : R]$ but we did not find references supporting this.

Here $K^* \subset L$ is the fixed field of the group of automorphisms $\tau \in \mathrm{Gal}(L, \mathbb{Q})$ for which $\tau \circ \Phi = \Phi$ (this is known as the "reflex field" of K with respect to Φ). The Φ-determinant condition then states that a lift with CM type Φ can be found if and only if there exists a prime $\mathfrak{p}^*$ of K^* above p such that $\det(\iota(x_1\alpha_1 + x_2\alpha_2 + \ldots + x_{2g}\alpha_{2g})^*) \equiv f_\Phi \bmod \mathfrak{p}^*$. In particular, if $\varrho_{(A,\lambda,\iota)} \cong \varrho_{(A',\lambda',\iota')}$ then it is possible to find lifts of (A, λ, ι) and (A', λ', ι') having the same CM type, and the connecting Shimura ideal class, whose existence is predicted by Theorem 6.7, will also connect (A, λ, ι) and (A', λ', ι'). □

It is not hard to check that if p is unramified then the Φ-Kottwitz condition cannot be satisfied with respect to more than one CM type Φ. In this case, at least under the assumption $R = \mathbb{Z}_K$, the number of orbits in characteristic p equals the number of orbits in characteristic 0, which is 2^g in most cases. But if p ramifies then there may be multiple options for Φ, meaning that the number of orbits decreases (i.e., several CM types "coincide" mod p).

Example 6.10 (CSIDH). Assume $g = 1$ and let $K = \mathbb{Q}(\sqrt{-p})$ for some prime number p. For simplicity we assume $p \equiv 1 \bmod 4$ so that $\mathbb{Z}_K = \mathbb{Z}[\sqrt{-p}]$ but this does not matter for our conclusions. The two possible Kottwitz polynomials are $x_1 \pm \sqrt{-p}x_2 \in \mathbb{Z}_K[x_1, x_2]$ (in both cases $K^* = K$). Modulo $\mathfrak{p}^* = (\sqrt{-p})$ both polynomials collide to x_1, which equals $\det(\iota(x_1 + \sqrt{-p}x_2)^*) \in \mathbb{F}_p[x_1, x_2]$ for the Frobenius orientation $\iota : \sqrt{-p} \mapsto \pi_p$ on any supersingular elliptic curve over $\mathbb{F}_p$. So in this case the CM types coincide, i.e., there is only one orbit, in accordance with case *(ii)* at the beginning of this section.

Example 6.11. Revisiting Example 6.4, one easily checks that both (A, λ, ι) and (A, λ, ι') satisfy the Kottwitz condition for a unique CM type, namely $\{\mathrm{id}, \varphi\}$ and $\{\mathrm{id}, \bar{\varphi}\}$, respectively, where

$$\varphi : \frac{1 + \sqrt{-m}}{2} \mapsto \frac{1 - \sqrt{-m}}{2}, \ \sqrt{d} \mapsto -\sqrt{d}$$

(and these CM types are inequivalent). So here we expect $2^2 = 4$ orbits.

Example 6.12 (R2-D2). For the R2-D2 orientations described in Example 3.4, one checks that $\varrho_{(A,\lambda,\iota)}$ maps

$$\sqrt{d} \mapsto \begin{pmatrix} 0 & d/c_\tau \\ c_\tau & 0 \end{pmatrix}, \qquad \sqrt{-p} \mapsto 0$$

when expressed with respect to the basis $(\omega, 0), (0, \omega')$, where ω, ω' denote regular 1-forms on $E, E^{(p)}$, respectively. One checks that

$$\det\left(\iota\left(x_1 + x_2\sqrt{d} + x_3\sqrt{-p} + x_4\sqrt{d}\sqrt{-p}\right)^*\right) = x_1^2 - dx_2^2.$$

This satisfies the Kottwitz condition with respect to all CM types of $\mathbb{Q}(\sqrt{d}, \sqrt{-p})$. So here we expect a unique orbit, coming from a 4-to-1 collision (this is modulo the disclaimer that we have naively applied Theorem 6.9 to the non-maximal order $\mathbb{Z}[\sqrt{d}, \sqrt{-p}]$). The resulting transitivity of the Shimura class group action in this setting plays an important role in the attack from [50].

Remark 6.13. In general, for $C_2 \times C_2$ fields, one checks that all 4 Kottwitz polynomials have the same reduction mod p if and only if p ramifies in both imaginary quadratic subfields. If p ramifies in exactly one imaginary quadratic subfield then we are left with 3 distinct reductions. As mentioned, in the unramified case the 4 Kottwitz polynomials have distinct reductions.

In summary, for a given CM order R in a CM field K, a rough roadmap to understanding the orbits of $\mathfrak{C}(R)$ on the set of primitively R-oriented superspecial PPAVs is as follows. Let us assume for simplicity that $R = \mathbb{Z}_K$. By Theorem 6.7 the number of orbits is at most the number 2^g of CM types of K. But the actual number of orbits is usually smaller:

- We first recall that it may (exceptionally) happen that there are no primitively $\mathbb{Z}_K$-oriented PPAVs in characteristic 0 with a certain given CM type Φ. In view of the proof of Theorem 6.9 this then also reduces the number of orbits over $\overline{\mathbb{F}}_p$. Note that this property is necessarily shared by the entire equivalence class of Φ.
- The most important decrease comes from the fact that the abelian varieties in certain orbits may not be superspecial. E.g., recall from *(i)* that if $g = 1$ then we end up with ordinary elliptic curves as soon as p splits in K. In higher dimension this is more subtle and a full characterization is lacking, but Appendix K contains the details for $g = 2$ under the assumption that p does not ramify in K; this (slightly) extends results by Goren [42].
- As we have just discussed, if p ramifies in K then several CM types may "coincide" mod p. This is controlled by Kottwitz' determinant condition.

A detailed understanding of all this, as well as of the precise assumptions on R under which Theorems 6.7 and 6.9 remain true, is what prevents us from giving a complete orbit description as we did for $g = 1$ at the beginning of the section.

Remark 6.14. Any oriented superspecial PPAV (A, λ, ι) can be defined over $\mathbb{F}_{p^2}$, so we can consider its Frobenius conjugate

$$(A, \lambda, \iota)^{(p)} = (A^{(p)}, \lambda^{(p)}, \iota^{(p)}).$$

We believe that $(A, \lambda, \iota)^{(p)}$ always lies in the orbit of $(A, \lambda, \iota \circ \sigma)$, perhaps under mild assumptions on R. This would extend Onuki's observation for $g = 1$ and we believe that it should follow from Ekedahl's product criterion [39, Prop. 2.4]; see [42, §2] for an accessible statement. But we have not analyzed this in detail.

7 Security considerations

We conclude with some security aspects of cryptosystems built from Shimura class group actions. Most threats we consider are analogs of existing attack scenarios in the case $g = 1$, that we expect to carry over to higher dimension. But the necessary tools are often insufficiently developed, due to limited attention and greater technical difficulty. The first 4 aspects have guided us in the parameter selection for the R2-D4 key exchange protocol from Appendix F.1:

Endomorphism ring computation. It is to be expected that the vectorization problem for Shimura class group actions on oriented superspecial PPAVs reduces to the problem of computing their endomorphism rings; this would be in complete analogy with the elliptic curve case [18,96]. We note that the endomorphism ring computation problem has several interpretations (e.g., computing an explicit basis of $4g^2$ self-maps, determining its structure as an abstract order with involution, computing an IKO-style matrix, ...) but they are most likely all equivalent: see [73] for the surface case, while in arbitrary dimension some aspects are treated in [78]. Up to our knowledge, the fastest known algorithms run in time

$$p^{1/3+o(1)} \text{ if } g = 1, \qquad p^{4/5+o(1)} \text{ if } g = 2, \qquad (\log p)^{O(g^2)} p^{g-1} \text{ if } g > 2,$$

see [95], [47] and [27, Thm. 6.15], respectively. In practice this attack scenario is largely overruled by Kuperberg's quantum attack, which we discuss next.

Generic attacks on the abelian hidden shift problem. The vectorization problem for commutative group actions can be modeled as an instance of the "abelian hidden shift problem" [23], for which several quantum sub-exponential time algorithms were devised by Kuperberg [58,59]. These run in time $L_{\#G}(1/2)$ with G the acting group; so in our case $\#G = \mathfrak{h}(R)$. The concrete cost of Kuperberg's algorithm is not well-understood, even in the elliptic curve case; one of the fuzzy factors is the cost of an oracle query [8,20], and this becomes even more intricate as the dimension grows. In Appendix F.1 we work with $\log_2 \mathfrak{h}(R) \approx 2048$, which is on par with CSIDH-4096, commonly believed to reach NIST-I security. In the **g**-REGA setting, it is customary to sample the secret keys from a strict subset of the acting group, but classical meet-in-the-middle approaches allow to solve the vectorization problem in time $O(\sqrt{\#\mathcal{K}})$ where $\mathcal{K}$ denotes the key space, i.e., the set of expressions of the form (4). We fix this at 256 bits for NIST-I.

Using non-maximal orders. In the elliptic curve case, the vectorization problem associated with a non-maximal order R reduces to any overorder in which R has a smooth index, by "walking up the volcano". This lowers the size of the class group and therefore weakens the problem, see e.g. [16, §6.3]. In higher dimension this picture is more complicated, but it is certain that in many scenarios (e.g., if the order is Bass), such ascending walks are again possible; see [3] and Appendix J for more context. So it seems wise to work with maximal orders, or at least with orders R for which $[\mathbb{Z}_K : R]$ is very small.[19] This is also a protective measure against the active "overorder attack" described below.

What about "imprimitive" CM types? This discussion is specific to higher dimension. We mentioned that key exchange from the R2-D2 orientation discussed in Example 3.4 got broken in [50], in the sense that the vectorization

[19] It may also be secure to work with CM orders R whose index in $\mathbb{Z}_K$ is large, but then it should be free of small factors so that ascension is infeasible, as in SCALLOP [35].

problem for $\mathfrak{C}(R)$ reduces to that for $\mathrm{Cl}(R_1), \mathrm{Cl}(R_2)$ with R_1, R_2 the two (largest possible) imaginary quadratic subrings of R. The attack relies on the colliding inequivalent CM types discussed in Example 6.12 and does not obviously generalize to arbitrary $C_2 \times C_2$ orientations: this is a compelling open question. More generally, one might wonder whether there is a general security issue with mod-p reductions of oriented PPAVs (A, λ, ι) whose CM type Φ is imprimitive, meaning there is a CM subfield $K' \subset K$ and a CM type Φ' of K' such that Φ consists of those embeddings whose restriction to K' is an element of Φ'. It is known that this is equivalent (in characteristic 0) to A being isogenous to a product of lower-dimensional abelian varieties [93, Thm. 5.2(3)]. All CM types of $C_2 \times C_2$ fields are imprimitive, while they are always primitive in the cases C_4, D_4.

Genus theory attacks. In the elliptic curve case, the decisional Diffie–Hellman assumption for class group actions can often be broken using genus theory [19,15]. This attack strategy generalizes to higher dimension, as was illustrated in [19, §6] and worked out in full generality in [50]. It can be countered by making sure one acts with elements from $\mathfrak{C}(R)^2$ (which is automatic if $\#\mathfrak{C}(R)$ is odd).

Backdoors and active attacks. One realistic threat is that a malicious authority or another adversary could lure Alice or Bob into acting with their secret key on a dishonestly generated PPAV (A, λ, ι), primitively oriented only by an overorder $R' \supsetneq R$ rather than R; then in this case the key could be recovered along the lines of Remark 4.8. This attack can be prevented by public key validation, i.e., by checking that A is superspecial, which can be done efficiently thanks to recent work [28], and that ι is a primitive R-orientation. If R has a large index in $\mathbb{Z}_K$, however, the non-primitivity of ι might be hard to detect; see our corresponding discussion at the end of Section 3. Thus, for this reason as well, it is safer to consider orientations by orders that are near-maximal.

References

1. N. Alamati, L. De Feo, H. Montgomery, and S. Patranabis. Cryptographic group actions and applications. In *Asiacrypt 2020*, volume 12492 of *Lecture Notes in Computer Science*, pages 411–439. Springer, 2020.
2. B. Allombert, J.-F. Biasse, J. K. Eriksen, P. Kutas, C. Leonardi, A. Page, R. Scheidler, and M. T. Bagi. Faster SCALLOP from non-prime conductor suborders in medium sized quadratic fields. In *PKC 2025*, volume 15676 of *Lecture Notes in Computer Science*, pages 333–363. Springer, 2025.
3. S. Arpin, S. Marseglia, and C. Springer. Isogeny graphs of abelian varieties and singular ideals in orders. Preprint available at `https://arxiv.org/abs/2508.03570`.
4. H. Bass. Torsion free and projective modules. *Transactions of the American Mathematical Society*, 102(2):319–327, 1962.
5. A. Basso, G. Borin, R. Rueger, and S. Schaeffler. CORAL: Faster isogeny group action for post-quantum NIKE. In *Crypto 2026*, volume 16803 of *Lecture Notes in Computer Science*, pages 506–536. Springer, 2026.

6. G. Bisson. Computing endomorphism rings of abelian varieties of dimension two. *Mathematics of Computation*, 84(294):1977–1989, 2015.
7. G. Bisson and M. Streng. On polarised class groups of orders in quartic CM-fields. *Mathematical Research Letters*, 24(2):247–270, 2017.
8. X. Bonnetain, M. Duparc, D. Lazzarini, L. Maino, C. Martindale, C. Petit, S. Schaeffler, A. Schrottenloher, and A. Sferlazza. Quantum security analysis of unrestricted isogeny-based group actions. Preprint available at `https://eprint.iacr.org/2026/2215`.
9. R. Bröker, D. Gruenewald, and K. Lauter. Explicit CM theory for level 2-structures on abelian surfaces. *Algebra & Number Theory*, 5(4):495–528, 2011.
10. E. H. Brooks, D. Jetchev, and B. Wesolowski. Isogeny graphs of ordinary abelian varieties. *Research in Number Theory*, 3:art. no. 28, 2017.
11. N. Bruin, E. V. Flynn, and D. Testa. Descent via $(3, 3)$-isogeny on Jacobians of genus 2 curves. *Acta Arith.*, 165(3):201–223, 2014.
12. P. Castel. Solving quadratic equations in dimension 5 or more without factoring. In *ANTS-X*, volume 1 of *The Open Book Series*, pages 213–233. MSP, 2013.
13. W. Castryck, T. Decru, P. Kutas, A. Laval, C. Petit, and Y. B. Ti. KLPT^2: Algebraic pathfinding in dimension two and applications. In *Crypto 2025*, volume 16000 of *Lecture Notes in Computer Science*, pages 167–200. Springer, 2025.
14. W. Castryck, J. K. Eriksen, R. Invernizzi, and F. Vercauteren. The principal ideal problem for endomorphism rings of superspecial abelian varieties. *Research in Number Theory*. To appear (proceedings of ANTS-XVI).
15. W. Castryck, M. Houben, F. Vercauteren, and B. Wesolowski. On the decisional Diffie–Hellman problem for class group actions on oriented elliptic curves. *Research in Number Theory*, 8(4):art. no. 99, 2022. Proceedings of ANTS-XV.
16. W. Castryck, R. Invernizzi, G. Lorenzon, J. Meers, and F. Vercauteren. Orient Express: Using Frobenius to express oriented isogenies. In *Latincrypt 2023*, volume 14168 of *Lecture Notes in Computer Science*, pages 145–173. Springer, 2023.
17. W. Castryck, T. Lange, C. Martindale, L. Panny, and J. Renes. CSIDH: An efficient post-quantum commutative group action. In *Asiacrypt 2018*, volume 11274 of *Lecture Notes in Computer Science*, pages 395–427. Springer, 2018.
18. W. Castryck, L. Panny, and F. Vercauteren. Rational isogenies from irrational endomorphisms. In *Eurocrypt 2020*, volume 12106 of *Lecture Notes in Computer Science*, pages 523–548. Springer, 2020.
19. W. Castryck, J. Sotáková, and F. Vercauteren. Breaking the decisional Diffie–Hellman problem for class group actions using genus theory: Extended version. *Journal of Cryptology*, 35(4):art. no. 24, 2022.
20. J. Chávez-Saab, J.-J. Chi-Dominguez, S. Jaques, and F. Rodriguez-Henriquez. The SQALE of CSIDH: sublinear Vélu quantum-resistant isogeny action with low exponents. *Journal of Cryptographic Engineering*, 12(3):349–368, 2022.
21. M. Chen, A. Leroux, and L. Panny. SCALLOP-HD: Group action from 2-dimensional isogenies. In *PKC 2024*, volume 14603 of *Lecture Notes in Computer Science*, pages 190–216. Springer, 2024.
22. M. Chenu and B. Smith. Higher-degree supersingular group actions. *Mathematical Cryptology*, 1(2):85–101, 2022.
23. A. Childs, D. Jao, and V. Soukharev. Constructing elliptic curve isogenies in quantum subexponential time. *Journal of Mathematical Cryptology*, 8(1):1–29, 2014.
24. H. Cohen and H. W. Lenstra. Heuristics on class groups of number fields. In *Number Theory Noordwijkerhout 1983*, pages 33–62, Berlin, Heidelberg, 1984. Springer.

25. L. Colò and D. Kohel. Orienting supersingular isogeny graphs. *Journal of Mathematical Cryptology*, 14:414–437, 2020.
26. K. Conrad. Galois groups of cubics and quartics (not in characteristic 2). Expository paper, available at `https://kconrad.math.uconn.edu/blurbs/galoistheory/cubicquartic.pdf`.
27. M. Corte-Real Santos, H. Kir, O. Taibi, and B. Wesolowski. Rapid mixing in superspecial isogeny graphs. Preprint available at `https://eprint.iacr.org/2026/2429`.
28. M. Corte-Real Santos, G. Lorenzon, and K. Reijnders. Supersingularity and superspeciality verification of abelian surfaces. Preprint available at `https://eprint.iacr.org/2026/2259`.
29. J.-M. Couveignes. Hard homogeneous spaces. Unpublished note available at `https://eprint.iacr.org/2006/291`.
30. D. A. Cox. *Primes of the Form x^2+ny^2: Fermat, Class Field Theory, and Complex Multiplication*. Pure and Applied Mathematics. Wiley, 2nd edition, 2013.
31. P. Dartois, J. K. Eriksen, T. B. Fouotsa, A. Herlédan Le Merdy, R. Invernizzi, D. Robert, R. Rueger, F. Vercauteren, and B. Wesolowski. PEGASIS: Practical effective class group action using 4-dimensional isogenies. In *Crypto 2025*, volume 16000 of *Lecture Notes in Computer Science*, pages 67–99. Springer, 2025.
32. P. Dartois, J. K. Eriksen, R. Invernizzi, and F. Vercauteren. qt-Pegasis: Simpler and faster effective class group actions. In *Eurocrypt 2026*, volume 16544 of *Lecture Notes in Computer Science*, pages 636–665. Springer, 2026.
33. P. Dartois, A. Leroux, D. Robert, and B. Wesolowski. SQIsignHD: New dimensions in cryptography. In *Eurocrypt 2024*, volume 14651 of *Lecture Notes in Computer Science*, pages 3–32. Springer, 2024.
34. L. De Feo, C. D. de Saint Guilhem, T. B. Fouotsa, P. Kutas, A. Leroux, C. Petit, J. Silva, and B. Wesolowski. Séta: Supersingular encryption from torsion attacks. In *Asiacrypt 2021*, volume 13093 of *Lecture Notes in Computer Science*, pages 249–278. Springer, 2021.
35. L. De Feo, T. B. Fouotsa, P. Kutas, A. Leroux, S.-P. Merz, L. Panny, and B. Wesolowski. SCALLOP: Scaling the CSI-FiSh. In *PKC 2023*, volume 13940 of *Lecture Notes in Computer Science*, pages 345–375. Springer, 2023.
36. B. A. Dina, S. Ionica, and J. Sijsling. Isogenous hyperelliptic and non-hyperelliptic Jacobians with maximal complex multiplication. *Mathematics of Computation*, 92:349–383, 2023.
37. Y. A. Drozd, V. V. Kirichenko, and A. V. Roiter. Hereditary and Bass orders. *Math. USSR-Izvestiya*, 1(6):1357–1375, 1967. Translated from Izv. Akad. Nauk SSSR Ser. Mat. **31** (1967), 1415–1436.
38. D. Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.
39. T. Ekedahl. On supersingular curves and abelian varieties. *Mathematica Scandinavica*, 60:151–178, 1987.
40. A. Enge and E. Thomé. Computing class polynomials for abelian surfaces. *Experimental Mathematics*, 23(2):129–145, 2014.
41. D. Goldstein and M. Schacher. Norms in central simple algebras. *Pacific Journal of Mathematics*, 292(2):373–388, 2018.
42. E. Z. Goren. On certain reduction problems concerning abelian surfaces. *Manuscripta mathematica*, 94(1):33–44, 1997.
43. E. Z. Goren and K. E. Lauter. Class invariants for quartic CM fields. *Annales de L'Institut Fourier*, 57(2):457–480, 2007.

44. C. Greither. On the two generator problem for the ideals of a one-dimensional ring. *Journal of Pure and Applied Algebra*, 24(3):265–276, 1982.
45. A. Herlédan Le Merdy and B. Wesolowski. The supersingular endomorphism ring problem given one endomorphism. *IACR Communications in Cryptology*, 2(1):art. no. 6, 2025.
46. M. Houben. Efficient post-quantum commutative group actions from orientations of large discriminant. In *Asiacrypt 2025*, volume 16248 of *Lecture Notes in Computer Science*, pages 141–173. Springer, 2025.
47. M. Houben and L. L. Pham. Computing endomorphisms of superspecial principally polarized abelian surfaces in time and memory $p^{4/5+o(1)}$. Preprint available at `https://eprint.iacr.org/2026/2530`.
48. J. G. Huard, B. K. Spearman, and K. S. Williams. Integral bases for quartic fields with quadratic subfields. *Journal of Number Theory*, 51(1):87–102, 1995.
49. T. Ibukiyama, T. Katsura, and F. Oort. Supersingular curves of genus two and class numbers. *Compositio Mathematica*, 57(2):127–152, 1986.
50. R. Invernizzi, J. K. Eriksen, A. Sferlazza, P. Tasbihgou, and F. Vercauteren. Cryptanalysis of higher-dimensional module actions via twisting and genus theory. Private communication.
51. S. Ionica and E. Thomé. Isogeny graphs with maximal real multiplication. *Journal of Number Theory*, 207:385–422, 2020.
52. G. Ivanyos. *Algorithms for algebras over global fields*. PhD thesis, Hungarian Academy of Sciences, 1996.
53. E. Kani. The number of curves of genus two with elliptic differentials. *Journal für die reine und angewandte Mathematik*, 1997(485):93–122, 1997.
54. A. Y. Kitaev. Quantum measurements and the Abelian stabilizer problem. Unpublished note available at `https://arxiv.org/abs/quant-ph/9511026`.
55. M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol. *The Book of Involutions*, volume 44 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, 1998.
56. D. Kohel, K. Lauter, C. Petit, and J.-P. Tignol. On the quaternion ℓ-isogeny path problem. *LMS Journal of Computation and Mathematics*, 17(A):418–432, 2014. Proceedings of ANTS-XI.
57. P. Koymans and C. Pagano. On the distribution of $\mathrm{Cl}(K)[l^\infty]$ for degree l cyclic fields. *Journal of the European Mathematical Society*, 24(4):1189–1283, 2022.
58. G. Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. *SIAM Journal on Computing*, 35(1):170–188, 2005.
59. G. Kuperberg. Another subexponential-time quantum algorithm for the dihedral hidden subgroup problem. In *TQC 2013*, volume 22 of *Leibniz International Proceedings in Informatics*, pages 20–34, 2013.
60. P. Kutas and M. Montessinos. Efficient computations in central simple algebras using Amitsur cohomology. *Journal of Algebra*, 665:255–281, 2025.
61. S. Lang. *Complex multiplication*, volume 255 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 1983.
62. S. Lang. *Algebraic Number Theory*, volume 110 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 1994.
63. K. Lauter and D. Robert. Improved CRT algorithm for class polynomials in genus 2. In *ANTS-X*, volume 1 of *The Open Book Series*, pages 437–461. MSP, 2013.
64. F. Lemmermeyer. Kuroda's class number formula. *Acta Arithmetica*, 66(3):245–260, 1994.
65. H. Lenstra, J. Pila, and C. Pomerance. A hyperelliptic smoothness test. II. *Proceedings of the London Mathematical Society*, 84(1):105–146, 2002.

66. J. Love and D. Boneh. Supersingular curves with small noninteger endomorphisms. In *ANTS-XIV*, volume 4 of *The Open Book Series*, pages 7–22. MSP, 2020.
67. S. Marseglia. Cohen–Macaulay type of orders, generators and ideal classes. *Journal of Algebra*, 658:247–276, 2024.
68. C. Martindale. *Isogeny graphs, modular polynomials, and applications*. PhD thesis, Université de Bordeaux and Universiteit Leiden, 2018.
69. K. McMurdy. Explicit representation of the endomorphism rings of supersingular elliptic curves. Unpublished note available at `https://pages.ramapo.edu/~kmcmurdy/research/McMurdy-ssEndoRings.pdf`.
70. J. S. Milne. Abelian varieties. Course notes available at `https://www.jmilne.org/math/CourseNotes/av.html`.
71. J. S. Milne. Algebraic number theory. Course notes available at `https://www.jmilne.org/math/CourseNotes/ant.html`.
72. J. S. Milne. Complex multiplication. Course notes available at `https://www.jmilne.org/math/CourseNotes/cm.html`.
73. M. Montessinos. Equivalent computational problems for superspecial abelian surfaces. Preprint available at `https://eprint.iacr.org/2026/120`.
74. J. Neukirch. *Algebraic Number Theory*, volume 322 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, Heidelberg, 1999. Translated from the 1992 German original by Norbert Schappacher, with a foreword by G. Harder.
75. M.-H. Nicole. *Superspecial abelian varieties, theta series and the Jacquet-Langlands correspondence*. PhD thesis, McGill University, 2005.
76. H. Onuki. On oriented supersingular elliptic curves. *Finite Fields and Their Applications*, 69:art. no. 101777, 2021.
77. A. Page and D. Robert. Introducing Clapoti(s): Evaluating the isogeny class group action in polynomial time. Unpublished note available at `https://eprint.iacr.org/2023/1766`.
78. A. Page, D. Robert, and J. Soumier. On the conversion of module representations for higher dimensional supersingular isogenies. In *Crypto 2026*, volume 16803 of *Lecture Notes in Computer Science*, pages 415–446. Springer, 2026.
79. A. Page and B. Wesolowski. The supersingular endomorphism ring and one endomorphism problems are equivalent. In *Eurocrypt 2024*, volume 14656 of *Lecture Notes in Computer Science*, pages 388–417. Springer, 2024.
80. L. Panny, C. Petit, and M. Stopar. KLaPoTi: An asymptotically efficient isogeny group action from 2-dimensional isogenies. *IACR Communications in Cryptology*, 2(3):art. no. 5, 2025.
81. M. Peters. Quadratische Formen über Zahlringen. *Acta Arithmetica*, 24(2):157–164, 1973.
82. D. Robert. The geometric interpretation of the Tate pairing and its applications. Preprint available at `https://eprint.iacr.org/2023/177`.
83. D. Robert. The module action for isogeny based cryptography. Preprint available at `https://eprint.iacr.org/2024/1556`.
84. D. Robert. On the efficient representation of isogenies. Survey written for NuTMiC 2024, available at `https://eprint.iacr.org/2024/1071`.
85. D. Robert. Some notes on algorithms for abelian varieties. Compilation of notes available at `https://eprint.iacr.org/2024/406`.
86. D. Robert. Breaking SIDH in polynomial time. In *Eurocrypt 2023*, volume 14008 of *Lecture Notes in Computer Science*, pages 472–503. Springer, 2023.
87. A. Rostovtsev and A. Stolbunov. Public-key cryptosystem based on isogenies. Unpublished note available at `https://eprint.iacr.org/2006/145`.

88. S. P. Sanon. *Endomorphism rings of ordinary abelian varieties*. PhD thesis, Technische Universität Kaiserslautern, 2022.
89. G. Shimura. *Abelian varieties with complex multiplication and modular functions*, volume 46 of *Princeton Mathematical Series*. Princeton University Press, 1998.
90. C. L. Siegel. Sums of mth powers of algebraic integers. *Annals of Mathematics*, 46(2):313–339, 1945.
91. J. H. Silverman. *The Arithmetic of Elliptic Curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 2009.
92. B. K. Spearman and K. S. Williams. Unramified quadratic extensions of a quadratic field. *The Rocky Mountain Journal of Mathematics*, 25(2):783–788, 1995.
93. M. Streng. *Complex multiplication of abelian surfaces*. PhD thesis, Leiden University, 2010.
94. J. Voight. *Quaternion algebras*, volume 288 of *Graduate Texts in Mathematics*. Springer, 2021.
95. B. Wesolowski. The supersingular isogeny problem in time and memory $p^{1/3+o(1)}$. Preprint available at `https://eprint.iacr.org/2026/1486`.
96. B. Wesolowski. Orientations and the supersingular endomorphism ring problem. In *Eurocrypt 2022*, volume 13277 of *Lecture Notes in Computer Science*, pages 345–371. Springer, 2022.
97. B. Wesolowski. The supersingular isogeny path and endomorphism ring problems are equivalent. In *FOCS 2022*, pages 1100–1111. IEEE Computer Society Press, 2022.
98. K. S. Williams. Integers of biquadratic fields. *Canadian Mathematical Bulletin*, 13(4):519–526, 1970.

A Background on the background

A.1 Cryptographic group actions

Let X be a finite set and G a finite commutative group in which one can compute efficiently. A cryptographic group action $\star : G \times X \to X$ is an action of G on X meeting some hardness properties, most notably one-wayness: given $x_1, x_2 \in X$ it is (computationally) hard to retrieve $g \in G$, if it exists, such that $g \star x_1 = x_2$. This problem is often referred to as vectorization in G. An effective group action (EGA) is one where the group action can be evaluated efficiently at all inputs. Only fairly recently, it was discovered how to build effective cryptographic group actions from isogenies between elliptic curves, see [31,32,77,80].

For certain applications such as non-interactive key exchange it is enough to have a so-called restricted effective group action (REGA). More explicitly, the action is called a $\mathbf{g}$-REGA for some generating subset $\mathbf{g} = \{g_1, \ldots, g_n\}$ of G if n is polylogarithmic in $|G|$ and there exist efficient algorithms to compute $g_i \star x$, $g_i^{-1} \star x$ for all $g_i \in \mathbf{g}, x \in X$. Most isogeny-based group actions in the literature fall into this category, see e.g. [17,25,22,16,46,5]. This will also be the case for the Shimura class group actions studied in this paper, where potential generalizations of [32] to construct an EGA from this are left as future work. We refer to [1] for more details on REGAs and EGAs and their range of applications.

A.2 Polarized isogenies

In the next few paragraphs we recall some basic facts and terminology needed for studying the Shimura class group action. All statements below are well-known to specialists and kept concise; we refer the reader to [93,72,70] and the references included below for more details.

Abelian varieties and isogenies. Let k be an algebraically closed field of characteristic p, where p is either a prime or 0. An abelian variety over k is a smooth projective algebraic variety, equipped with a base point $0 \in A$ and geometric morphisms $A \times A \to A : (P, Q) \mapsto P + Q$, $A \to A : P \mapsto -P$ turning A into an algebraic group; elliptic curves are abelian varieties of dimension 1. A homomorphism $\varphi : A_1 \to A_2$ between abelian varieties A_1, A_2 is a morphism such that $\varphi(0) = 0$; this is automatically a group homomorphism at the level of points; if $\dim A_1 = \dim A_2$ and φ is surjective then it is called an isogeny. The set of endomorphisms (= self-homomorphisms) of an abelian variety A becomes a ring when equipped with point-wise addition and composition; it is denoted $\mathrm{End}(A)$. We identify $\mathbb{Z}$ with the subring of $\mathrm{End}(A)$ consisting of scalar multiplications and write $\mathrm{End}^0(A) = \mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$ for the endomorphism algebra. Assuming $p > 0$, an abelian variety A/k of dimension g is called superspecial if it is isomorphic to E^g for some supersingular elliptic curve E/k. If $g \geq 2$ then the choice of E does not matter: by a theorem of Deligne [41] any two products of g supersingular elliptic curves are isomorphic.

Principal polarizations. A polarization on an abelian variety A is the isogeny

$$\lambda : A \to A^\vee \cong \mathrm{Pic}^0(A) : P \mapsto [t_{-P}D - D]$$

attached to an ample divisor D on A (where t_{-P} denotes point-wise translation by $-P$). The reader not familiar with the language of divisors can just think of a polarization as an isogeny to a certain companion abelian variety, called the dual abelian variety, and satisfying a certain "positivity" condition; this is good enough for almost all that follows. If this isogeny is of degree 1 then it is called a principal polarization and we call (A, λ) a principally polarized abelian variety, or PPAV in short (in the specific case of surfaces we will often write PPAS). While elliptic curves have a unique, canonical, principal polarization, this is not true in general for abelian varieties.

Example A.1 (Jacobians). Consider a smooth projective curve C of genus $g \geq 1$ along with the Abel–Jacobi map $u : C \hookrightarrow \mathrm{Jac}(C) \cong \mathrm{Pic}^0(C) : P \mapsto [P - P_0]$, where $P_0 \in C$ is some fixed base point. The codimension 1 subvariety of $\mathrm{Jac}(C)$ consisting of all sums of $g - 1$ points in the image of u can be seen as a divisor D on $\mathrm{Jac}(C)$. The associated polarization $\lambda : \mathrm{Jac}(C) \to \mathrm{Jac}(C)^\vee$ is a principal polarization, known as the canonical principal polarization on $\mathrm{Jac}(C)$.

Example A.2 (products). If $(A_1, \lambda_1), (A_2, \lambda_2)$ are PPAVs, with principal polarizations coming from divisors D_1, D_2, then the product abelian variety $A_1 \times A_2$

comes with the product polarization $\lambda_1 \times \lambda_2$, attached to the divisor $D_1 \times A_2 + A_1 \times D_2$, which is again principal.

Rosati involution. With every homomorphism $\varphi : A_1 \to A_2$ comes a dual homomorphism $\hat{\varphi} : A_2^\vee \to A_1^\vee$ (which, in the case of a surjective homomorphism, is obtained by pulling back divisors on A_2 under φ). If λ_1, λ_2 are principal polarizations on A_1, A_2 then this gives rise to an adjoint homomorphism

$$\varphi^\dagger = \lambda_1^{-1} \circ \hat{\varphi} \circ \lambda_2 : A_2 \to A_1.$$

When applied to endomorphisms of a PPAV (A, λ), this defines an involution on $\mathrm{End}(A)$ called the Rosati involution, which extends by linearity to the entire endomorphism algebra $\mathrm{End}^0(A)$. The aforementioned positivity condition on λ is reflected in the fact that the quadratic form $\alpha \mapsto \mathrm{Tr}(\alpha\alpha^\dagger)$ on $\mathrm{End}^0(A)$, where $\mathrm{Tr}(x)$ denotes the trace of the multiplication-by-x map, viewed as a linear map on the $\mathbb{Q}$-vector space $\mathrm{End}^0(A)$, is positive-definite.

Example A.3. Any homomorphism $\varphi : A_1 \times A_2 \to B_1 \times B_2$ between two products can be seen as a matrix

$$\begin{pmatrix} \varphi_{11} & \varphi_{21} \\ \varphi_{12} & \varphi_{22} \end{pmatrix}$$

of homomorphisms $\varphi_{ij} : A_i \to B_j$. If A_1, A_2, B_1, B_2 come equipped with principal polarizations, then with respect to the respective product polarizations we have

$$\varphi^\dagger = \begin{pmatrix} \varphi_{11}^\dagger & \varphi_{12}^\dagger \\ \varphi_{21}^\dagger & \varphi_{22}^\dagger \end{pmatrix}, \tag{8}$$

which can be thought of as a "conjugate-transpose". This naturally generalizes to homomorphisms between products of arbitrarily many abelian varieties.

Polarized isogenies. If $\varphi : A_1 \to A_2$ is an isogeny and λ_2 is a principal polarization on A_2 then $\hat{\varphi} \circ \lambda_2 \circ \varphi$ is a polarization on A_1, which cannot be principal in general, if only for degree reasons. According to a theorem of Mumford, if λ_1 is a principal polarization on A_1 then the polarizations on A_1 are exactly those isogenies of the form $\lambda_1 \circ \alpha$ with α a totally positive endomorphism on A_1: by this we mean that $\alpha = \alpha^\dagger$ and the (necessarily real) eigenvalues of the multiplication-by-α map are positive. Applying this to the foregoing, this means that for $\hat{\varphi} \circ \lambda_2 \circ \varphi$ to be a polarization, there must exist a totally positive $\alpha \in \mathrm{End}(A_1)$ such that the diagram

$$\begin{array}{ccccc} A_1 & \xleftarrow{\alpha} & A_1 & \xrightarrow{\varphi} & A_2 \\ & \searrow^{\lambda_1} & & & \downarrow{\scriptstyle \lambda_2} \\ & & A_1^\vee & \xleftarrow[\widehat{\varphi}]{} & A_2^\vee \end{array}$$

commutes. We then say that $\varphi : (A_1, \lambda_1) \to (A_2, \lambda_2)$ is a polarized isogeny with respect to α, or an α-isogeny for short. An important special case is $\alpha = N \in \mathbb{Z}_{>0}$. We note that the degree of an N-isogeny equals N^g; the integer N itself is then called its reduced degree, denoted $\mathrm{degrd}(\varphi)$. An isomorphism between PPAVs is a 1-isogeny. We state the next result for later use [70, Prop. 13.8]:

Proposition A.4. *Let (A_1, λ_1) be a PPAV, let $H \subset A_1$ be a finite subgroup whose order is not a multiple of p, and let $\alpha \in \mathrm{End}(A_1)$ be totally positive. Let $\varphi : A_1 \to A_2$ be a separable isogeny with kernel H. There exists a principal polarization λ_2 on A_2 such that $\varphi : (A_1, \lambda_1) \to (A_2, \lambda_2)$ is an α-isogeny if and only if H is maximal with the property that*

$$e_\alpha(P, \lambda_1(Q)) = 1, \quad \forall P, Q \in H,$$

where e_α denotes the α-Weil pairing on A_1. Moreover λ_2 is unique in this case.

Example A.5 (IKO matrices). Let E be a supersingular elliptic curve over k and assume that $p > 0$; fix an isomorphism $\vartheta : \mathcal{O} \to \mathrm{End}(E)$ from a suitable maximal order $\mathcal{O} \subset B_{p,\infty}$ in the unique quaternion algebra ramified at p and ∞. Consider the superspecial abelian surface $A = E^2$ equipped with the product polarization λ_1. In this case Mumford's theorem says that the principal polarizations on A are precisely the maps of the form $\lambda_1 \circ \vartheta(g)$ where

$$g = \begin{pmatrix} s & r \\ \bar{r} & t \end{pmatrix}, \qquad s, t \in \mathbb{Z}_{>0}, r \in \mathcal{O}, st - r\bar{r} = 1,$$

and ϑ is applied entry-wise; the abbreviation IKO stands for Ibukiyama, Katsura and Oort, the authors of [49] where this is detailed; see Section 5.

A.3 Omitted proofs

Proof (of Lemma 2.4). Let $\varepsilon_1, \ldots, \varepsilon_{g-1}$ be a set of fundamental units of R_0, chosen such that they are all positive under some fixed real embedding. Since $(R, \varepsilon_i^2) \in P_R$ for all i, it is clear that all Shimura classes for which $\mathfrak{a} = R$ are among expressions of the form

$$[(R, \varepsilon_1^{e_1} \cdots \varepsilon_{g-1}^{e_{g-1}})], \ e_i \in \{0, 1\}. \tag{9}$$

The lemma follows because $\mathfrak{a}\bar{\mathfrak{a}} = \mu_1 R = \mu_2 R$ iff $\mu_1^{-1}\mu_2 \in R_0^+$ is a unit. □

Remark A.6. Not all expressions of the form (9) are valid Shimura classes: this is the case only if $\varepsilon_1^{e_1} \cdots \varepsilon_{g-1}^{e_{g-1}}$ is totally positive. It may well happen that this is true only when $e_1 = e_2 = \ldots = e_{g-1} = 0$; in such cases $[(\mathfrak{a}, \mu)]$ is always uniquely determined by $\mathfrak{a}$ (however, the existence of μ still imposes a condition on $\mathfrak{a}$). Moreover, even in the case of a totally positive unit $\varepsilon \neq 1$, the class $[(R, \varepsilon)]$ may still be trivial: if $\varepsilon = \alpha\bar{\alpha}$ for some $\alpha \in R$, then $(R, \varepsilon) = (\alpha R, \alpha\bar{\alpha}) \in P_R$.

Proof (of Lemma 2.6). Consider a Shimura ideal $(\mathfrak{a}, \mu)$ with $\mathfrak{a} \subset R$ and let n be an integer such that $\frac{n}{\mu}\mathfrak{a}$ is integral. Then

$$[(\mathfrak{a}, \mu)]^2 = [(\mathfrak{a}^2, \mu^2)] = [(\frac{n}{\mu}\mathfrak{a}^2, \frac{n}{\mu}\frac{\bar{n}}{\bar{\mu}}\mu^2)] = [(\frac{n}{\mu}\mathfrak{a}^2, n^2)],$$

where we have used that $n = \bar{n}$, $\mu = \bar{\mu}$. □

Remark A.7. Based on computational evidence, we expect the dichotomy

$$[\mathfrak{C}(R) : \mathfrak{C}(R)_{\mathbb{Z}}] \in \{1, \#\mathfrak{C}(R)[2]\}$$

to hold for any CM order $R \subset K$.

Proof (of Theorem 2.7). Although it is hard to pinpoint a reference not putting additional assumptions on R and k, or not working in the language of complex tori [89] or Deligne modules [65], the arguments are standard. We highlight the constructive aspects of the proof, because these indicate how to work with the action, and are more concise about the remaining details. Let $(A_1, \lambda_1, \iota_1) \in \mathcal{M}_k(R)$ and $[(\mathfrak{a}, \mu)] \in \mathfrak{C}(R)$. We show how to construct

$$(A_2, \lambda_2, \iota_2) = [(\mathfrak{a}, \mu)] \star (A_1, \lambda_1, \iota_1),$$

a priori as an element of $\mathcal{M}_k(K)$, but we will then be able to conclude that we always land in $\mathcal{M}_k(R)$. By replacing $\mathfrak{a}$ with an equivalent fractional ideal if needed and scaling μ accordingly, we can assume that $\mathfrak{a}$ is integral of norm not divisible by p. We then consider the finite subgroup

$$H = \bigcap_{x \in \mathfrak{a}} \ker \iota_1(x)$$

of A_1, which has order $N_{K/\mathbb{Q}}(\mathfrak{a})$. Let $\varphi : A_1 \to A_2 := A_1/H$ be the corresponding quotient isogeny (in the language of $\mathfrak{a}$-multiplications, this amounts to acting with $\mathfrak{a}$ on the unpolarized abelian variety A_1, see [72, §7]). Writing $\mu = \sum_i x_i \bar{y}_i$ for $x_i, y_i \in \mathfrak{a}$ and letting $z_i = x_i\bar{y}_i/\mu \in R$ so that $\sum_i z_i = 1$, for any $P, Q \in H$ we have

$$\begin{aligned} e_{\iota(\mu)}(P, \lambda_1(Q)) &= \prod_i e_{\iota(\mu)}(\iota(z_i)(P), \lambda_1(Q)) = \prod_i e_{\iota(x_i)\iota(y_i)^\dagger}(P, \lambda_1(Q)) \\ &= \prod_i e_{\iota(y_i)^\dagger}(\iota(x_i)(P), \lambda_1(Q)) = \prod_i e_{\iota(y_i)^\dagger}(0, \lambda_1(Q)) = 1 \end{aligned}$$

where we have used some standard properties of the Weil pairing which, e.g., can be found in [82, §3.1], and the fact that $\lambda_1(Q)$ lies in the kernel of the dual of $\iota(y_i)^\dagger$. So by Proposition A.4 there is a unique principal polarization λ_2 on A_2 turning φ into an $\iota_1(\mu)$-isogeny. We then let $\iota_2 : K \to \mathrm{End}^0(A_2)$ be the push-forward orientation, i.e.,

$$\iota_2 : \alpha \mapsto (\varphi \circ \iota_1(\alpha) \circ \varphi^\dagger)/(\varphi \circ \varphi^\dagger).$$

From the multiplicativity of the unpolarized action [72, Prop. 7.28] it is not hard to deduce that acting in this way with $(\mathfrak{ab}, \mu\nu)$ amounts to the same thing as first acting with $(\mathfrak{a}, \mu)$, then with $(\mathfrak{b}, \nu)$. Since it is straightforward that Shimura ideals of the form $(\mu R, \mu\bar{\mu})$ with $\mu \in R$ act trivially, this defines an action of $\mathfrak{C}(R)$ as wanted, and freeness is easily established. To see that this lands in $\mathcal{M}_k(R) \subset \mathcal{M}_k(K)$ it is enough to observe that every ideal class admits two representing ideals with coprime norms, see e.g. [63, Proof of Lem. 27] for more details on this step. □

B Endomorphism rings of some well-known supersingular elliptic curves

We include explicit descriptions of the endomorphism rings of certain "extremal" supersingular elliptic curves:

- $y^2 = x^3 + x$ in characteristic $p \equiv 3 \bmod 4$, of j-invariant 1728: being the cryptographers' default example of a supersingular elliptic curve, this is of course very well-known and included for completeness only,
- $y^2 = x^3 + 1$ in characteristic $p \equiv 5 \bmod 6$, of j-invariant 0: an explicit description of its endomorphism ring can be found in [69, Prop. 3.2] but this description is wrongly attributed to $y^2 = x^3 - 1$, see Remark B.3,
- $y^2 = x^3 + 12x^2 + 18x$ in characteristic $p \equiv 5 \bmod 8$, which has j-invariant 8000: we are unaware of prior work describing the endomorphism ring as explicitly as in Lemma B.4 below; in [56, Lem. 3] it was studied as an abstract maximal order in $B_{p,\infty}$ but, unfortunately, that statement contains a typo, see Remark B.5 below.

We do not cover other congruence classes of p, but we note that also the maximal orders from [56, Lem. 4] for the cases $p \equiv 1 \bmod 4$ contain typos, as was pointed out in [14, Ftn. 4]. Since [69,56] are frequently cited references and the typos have been copied into the existing literature, we hope that this appendix helps in cleaning up the landscape.

Lemma B.1. *Let $p \equiv 3 \bmod 4$ be a prime number and consider the quaternion algebra*

$$B_{p,\infty} = \left(\frac{-1, -p}{\mathbb{Q}}\right)$$

ramified at p and ∞, with the usual generators $1, i, j, k$ (i.e., we have $i^2 = -1$, $j^2 = -p$ and $k = ij = -ji$) over $\mathbb{Q}$. Fix a square root $\mathbf{i} \in \mathbb{F}_{p^2}$ of -1. The elliptic curve $E : y^2 = x^3 + x$ over $\overline{\mathbb{F}}_p$ with j-invariant 1728 admits the endomorphism

$$\iota : (x, y) \mapsto (-x, \mathbf{i}y)$$

satisfying $\iota^2 = [-1]$ and

$$1 \mapsto [1], \quad i \mapsto \iota, \quad j \mapsto \pi_p, \quad k \mapsto \iota\pi_p \tag{10}$$

defines a ring isomorphism

$$\mathbb{Z}\left\langle 1, i, \frac{i+j}{2}, \frac{1+k}{2}\right\rangle \to \mathrm{End}(E),$$

where π_p denotes the p-th power Frobenius endomorphism on E.

Proof. As mentioned, this is standard, and easier than the lemmata below, so we will be concise. It is immediate that $\iota^2 = [-1]$ and, using $p \equiv 3 \bmod 4$, it is easy to check that $\iota\pi_p = -\pi_p\iota$. Thus the assignment (10) defines an isomorphism $B_{p,\infty} \to \mathrm{End}^0(E)$ at the level of algebras. To complete the proof, it suffices to observe that $(i+j)/2$, $(1+k)/2$ map to endomorphisms on E, which in view of [91, Cor. III.4.11] is equivalent to

$$E[2] \subset \ker(\iota + \pi_p), \ker([1] + \iota\pi_p),$$

but that is immediate from $E[2] = \{0, (0,0), (\mathbf{i}, 0), (-\mathbf{i}, 0)\}$. This is indeed enough, because an explicit calculation shows that $\mathbb{Z}\langle 1, i, (i+j)/2, (1+k)/2\rangle$ is a subring of $B_{p,\infty}$ and a discriminant computation shows that it concerns a maximal order.

Lemma B.2. *Let $p \equiv 5 \bmod 6$ be a prime number and consider the quaternion algebra*

$$B_{p,\infty} = \left(\frac{-3, -p}{\mathbb{Q}}\right)$$

ramified at p and ∞, with the usual generators $1, i, j, k$ (i.e., we have $i^2 = -3$, $j^2 = -p$ and $k = ij = -ji$) over $\mathbb{Q}$. Fix a primitive 3rd root of unity $z \in \mathbb{F}_{p^2}$. The elliptic curve $E : y^2 = x^3 + 1$ over $\overline{\mathbb{F}}_p$ with j-invariant 0 admits the endomorphism

$$\omega : (x, y) \mapsto (zx, y)$$

satisfying $\omega^2 + \omega + [1] = 0$. Writing $\Omega = [2]\omega + [1]$ and π_p for the p-th power Frobenius endomorphism on E, the assignment

$$1 \mapsto [1], \quad i \mapsto \Omega, \quad j \mapsto \pi_p, \quad k \mapsto \Omega\pi_p$$

defines an isomorphism

$$\mathbb{Z}\left\langle 1, \frac{-1+i}{2}, j, \frac{3+i+3j+k}{6}\right\rangle \to \mathrm{End}(E) \tag{11}$$

of rings.

Proof. It is easy to check that the (horizontal) line connecting two points of the form (zx_0, y_0), (z^2x_0, y_0) has (x_0, y_0) as its third point of intersection with E. This implies that $\omega^2 + \omega + [1] = 0$, and as a special case we find that $(0, 1)$ and $(0, -1)$ are points of order 3.

An easy calculation shows $\Omega^2 = [-3]$. Since $\pi_p^2 = [-p]$ and $\pi_p\omega = \omega^2\pi_p$ so that $\Omega\pi_p = -\pi_p\Omega$, the assignment from the statement of the lemma is seen to

define an isomorphism between $B_{p,\infty}$ and $\mathrm{End}^0(E)$. Clearly $(-1+i)/2 \mapsto \omega$, so it suffices to show that

$$[3] + \Omega + [3]\pi_p + \Omega\pi_p = [4] + [2]\omega + [4]\pi_p + [2]\omega\pi_p \tag{12}$$

annihilates $E[6]$: then the left-hand side of (11) lands inside $\mathrm{End}(E)$, and this left-hand side is a maximal order, as can be checked explicitly.

It obviously annihilates $E[2]$, so it suffices to concentrate on $E[3]$. An analysis of the 3-division polynomial $3x(x^3+4)$ of E shows that, besides $(0, \pm 1)$, the points of order 3 are

$$(t, \pm\sqrt{-3}),\ (zt, \pm\sqrt{-3}),\ (z^2t, \pm\sqrt{-3})$$

with $t = (-4)^{-(p-2)/3} \in \mathbb{F}_p$. We now do a case-by-case analysis of how π_p acts on these points, keeping in mind that they are of order 3:

- on $(0, \pm 1)$ it acts as the identity, so expression (12) acts as $\omega - [1]$ which clearly kills these points,
- on $(t, \pm\sqrt{-3})$ it acts by negation, easily seen to make expression (12) vanish,
- on $(zt, \pm\sqrt{-3})$ it acts as $-\omega$, again making expression (12) vanish after reduction modulo [3],
- similarly, on $(z^2t, \pm\sqrt{-3})$ it acts as $-\omega^2$, which similarly makes expression (12) vanish modulo [3].

This concludes the proof.

Remark B.3. In contrast with what is claimed in [69, Prop. 3.2] the statement fails for $E : y^2 = x^3 - 1$ as soon as $p \equiv 3 \bmod 4$. In this case 3 is a square in $\mathbb{F}_p$ and one finds that the points $(t, \pm\sqrt{3}) \in E[3]$ with $t = 4^{-(p-2)/3} \in \mathbb{F}_p$ are fixed by Frobenius. Expression (12) acts on these points as $\omega - [1]$ which fails to annihilate them.

Lemma B.4. *Let $p \equiv 5 \bmod 8$ be a prime number and consider the quaternion algebra*

$$B_{p,\infty} = \left(\frac{-2, -p}{\mathbb{Q}}\right)$$

ramified at p and ∞, with the usual generators $1, i, j, k$ (i.e., we have $i^2 = -2$, $j^2 = -p$ and $k = ij = -ji$) over $\mathbb{Q}$. Fix a square root $z \in \mathbb{F}_{p^2}$ of -2. The elliptic curve $E : y^2 = x^3 + 12x^2 + 18x$ over $\overline{\mathbb{F}}_p$ with j-invariant 8000 admits the endomorphism

$$\vartheta : (x, y) \mapsto \left(\frac{-y^2}{2x^2}, \frac{z}{4}y(1 - \frac{18}{x^2})\right)$$

satisfying $\vartheta^2 = [-2]$ and

$$1 \mapsto [1], \quad i \mapsto \vartheta, \quad j \mapsto \pi_p, \quad k \mapsto \vartheta\pi_p \tag{13}$$

defines a ring isomorphism

$$\mathbb{Z}\left\langle 1, i, \frac{1+i+j}{2}, \frac{i+2j\pm k}{4}\right\rangle \to \mathrm{End}(E) \tag{14}$$

where the sign $\pm$ *should be read as* $+$ *if* $p \equiv 5, 13 \bmod 48$ *and as* $-$ *if* $p \equiv 29, 37 \bmod 48$*, and* π_p *denotes the* p*-th power Frobenius endomorphism.*

Proof. In the proof all occurrences of $\pm$ should be read as $+$ if $p \equiv 5, 13 \bmod 48$ and as $-$ if $p \equiv 29, 37 \bmod 48$, and conversely for $\mp$. Throughout we fix a square root $\mathbf{i} \in \mathbb{F}_p$ of -1.

It is easy to check that ϑ is a well-defined endomorphism of degree 2 satisfying $\vartheta^2 = [-2]$. It is also easy to see that $\pi_p^2 = [-p]$ and $\vartheta\pi_p = -\pi_p\vartheta$, showing that (13) defines an isomorphism between $B_{p,\infty}$ and $\mathrm{End}^0(E)$. It thus suffices to show that

$$E[2] \subset \ker([1] + \vartheta + \pi_p), \qquad E[4] \subset \ker(\vartheta + [2]\pi_p \pm \vartheta\pi_p),$$

because this implies that the left-hand side of (14) indeed lands inside $\mathrm{End}(E)$, and this left-hand side is a maximal order, regardless of the sign chosen, as can be checked via an explicit discriminant calculation.

We do this in an artisanal way. First, write $O = (0, 0)$, $T_0 = (-6 + 3\mathbf{i}z, 0)$, $T_1 = (-6 - 3\mathbf{i}z, 0)$ for the points of order 2 on E. Using

$$\vartheta(O) = 0, \quad \vartheta(T_0) = \vartheta(T_1) = O$$

and

$$\pi_p(O) = O, \quad \pi_p(T_\ell) = T_{1-\ell} \text{ for } \ell = 0, 1$$

one indeed checks that $[1] + \vartheta + \pi_p$ annihilates $E[2]$ as wanted.

The remainder of the proof is a bit more tedious. By explicit calculation, the points of order 4 are seen to be groupable as follows:

- the halves of O, which have x-coordinates $3\mathbf{i}z$, $-3\mathbf{i}z$,
- the halves of T_0, with x-coordinates

$$-6 + 3\mathbf{i}z + 6\sqrt{1 - \mathbf{i}z}, \ -6 + 3\mathbf{i}z - 6\sqrt{1 - \mathbf{i}z},$$

- the halves of T_1, with x-coordinates

$$-6 - 3\mathbf{i}z + 6\sqrt{1 + \mathbf{i}z}, \ -6 - 3\mathbf{i}z - 6\sqrt{1 + \mathbf{i}z}.$$

These x-coordinates all live in $\mathbb{F}_{p^2}$, i.e., $1 + \mathbf{i}z$ and $1 - \mathbf{i}z$ are squares in $\mathbb{F}_{p^2}$. This follows, e.g., from $\pi_{p^2} = [-p]$, implying that for any point $P \in E[4]$ it holds that $\pi_{p^2}(P) = [-p]P = -P$ where we have used $p \equiv 1 \bmod 4$. For the same reason none of the corresponding y-coordinates is defined over $\mathbb{F}_{p^2}$, i.e., the points of order 4 all have field of definition $\mathbb{F}_{p^4}$.

It is immediate to see how $[2]\pi_p = \pi_p[2]$ acts on these halves: the respective results are O, T_1, T_0, independent of which half was chosen. Thus it suffices to show that

$$\vartheta(1 \pm \pi_p) = (1 \mp \pi_p)\vartheta$$

also maps the halves of O to O, the halves of T_0 to T_1 and the halves of T_1 to T_0, with the sign depending on $p \bmod 48$ as in the statement of the lemma.

Let us first concentrate on the halves of O. Since $O \in \ker \vartheta$ and since $\vartheta^2 = [-2]$, these halves must be mapped inside $\{T_0, T_1\}$ under ϑ. But this means that $(1 \pm \pi_p)\vartheta$ maps every half of O to $T_0 + T_1 = O$ as wanted.

Now consider a half of T_ℓ for $\ell \in \{0, 1\}$. Since $\vartheta(T_\ell) = O$ this half must be mapped to a half of O under ϑ. It follows that $(1 \mp \pi_p)\vartheta$ maps the half of T_ℓ to a sum of two halves of O with different x-coordinates. It is clear that this is indeed a point of order 2 different from O, hence an element of $\{T_0, T_1\}$, but we need to show that it concerns $T_{1-\ell}$.

We concentrate on a half H of T_0 (the other case is analogous). An explicit calculation shows that

$$\vartheta(H) = (-3\mathbf{i}z, y_0)$$

for some y-coordinate $y_0 \in \mathbb{F}_{p^4}$. Then

$$\pi_p\vartheta(H) = (3\mathbf{i}z, y_0^p)$$

and the usual addition formula shows that the x-coordinate of $(1 \mp \pi_p)\vartheta(H)$ is given by

$$\left(\frac{\mp y_0^p - y_0}{6\mathbf{i}z}\right)^2 - 12 = \frac{1}{72}(y_0^p \pm y_0)^2 - 12 = \frac{1}{72}((y_0^2)^p \pm 2y_0^{p+1} + y_0^2) - 12.$$

Since $y_0^2 = (-3\mathbf{i}z)^3 + 12(-3\mathbf{i}z)^2 + 18(-3\mathbf{i}z) = 108(2 - \mathbf{i}z)$ this equals

$$-6 \pm 3 \cdot 108^{(p-1)/2}(2 - \mathbf{i}z)^{(p+1)/2} = -6 \pm 3\left(\frac{3}{p}\right)(2 - \mathbf{i}z)^{(p+1)/2},$$

with $(\frac{\cdot}{p})$ the Legendre symbol. So upon setting this equal to the x-coordinate of T_1, it suffices to show

$$\pm\left(\frac{3}{p}\right)(2 - \mathbf{i}z)^{(p+1)/2} = -\mathbf{i}z.$$

Note, as a sanity check, that both sides square to 2, but we have to argue that we have the same square root of 2 on both sides. For this we rely on the congruence

$$(2 + \sqrt{2})^{(p+1)/2} \equiv \sqrt{2} \bmod p$$

in $\mathbb{Z}[\sqrt{2}]$, which holds if and only if $p \equiv 3, 13 \bmod 16$ assuming p is inert: to prove this, the main trick is to rewrite $2 + \sqrt{2} = \zeta^{-1}(1 + \zeta)^2$ for a primitive 8th root of unity ζ. So using $p \equiv 5 \bmod 8$ we find that the equality holds with $+$ if and only if $p \equiv 5, 13 \bmod 48$ and with $-$ if and only if $p \equiv 29, 37 \bmod 48$, as wanted.

Remark B.5. The two orders in (14) are isomorphic via $j \mapsto -j$, so also if $p \equiv 29, 37 \bmod 48$ then it is possible to give an isomorphism from the maximal order with a $+$ sign, but the isomorphism changes. These orders are essentially taken from [56, Lem. 3] but, there, the third basis element reads $(1 + j + k)/2$ rather than $(1 + i + j)/2$. This is an error: for the sign $-$ the resulting order is not maximal but of index 3 (for the sign $+$ it is).

C R2: three families of $\mathbb{F}_p$-rational orientations

In this section we study families of oriented PPASs defined over $\mathbb{F}_p$, where the orientations are such that the p-th power Frobenius endomorphism is among the orienting endomorphisms; in this context, we will denote this Frobenius endomorphism by Π_p to avoid confusion with p-th power Frobenius maps between elliptic curves, for which we keep writing π_p. We have already encountered R2-D2 in Example 3.4. The other orientations are called R2-C4 and R2-D4. Besides paying tribute to a series of astromech droids, these names make reference to the Galois group (of the Galois closure) of the underlying CM field K, which is either isomorphic to $D_2 \cong C_2 \times C_2$, C_4 or D_4 (indeed, A_4 and S_4 are impossible in the CM case). The R2 stands for "rational and 2-dimensional".[20] Our main motivation for including Π_p in the ring of orienting endomorphisms is that its representation comes "for free", i.e., we do not need to carry around corresponding interpolation data; see Section 4.3. This comes with important gains, similar to why, in the elliptic curve case, CSIDH stands out in terms of efficiency.

For each of the families, the starting surface is the Weil restriction

$$A = W_{\mathbb{F}_{p^2}/\mathbb{F}_p}(E) \cong_{\bar{\mathbb{F}}_p} E \times E^{(p)}$$

of a well-chosen supersingular elliptic curve $E/\mathbb{F}_{p^2}$, equipped with the product polarization λ. Recall: the p-th power Frobenius endomorphism on A is given by

$$\Pi_p = \begin{pmatrix} 0 & \pi_p \\ \pi_p & 0 \end{pmatrix}$$

with $\pi_p : E \to E^{(p)}$ the Frobenius isogeny. In R2-D2 we use a curve E for which $\pi_{p^2} = -p$, from which it follows that $\Pi_p^2 = -p$. In particular this requires that our orienting order R contains $\sqrt{-p}$. For this reason, for R2-C4 and R2-D4 one must proceed differently: indeed, C_4-fields and D_4-fields that are CM cannot contain any imaginary quadratic subfields. Instead we choose $E/\mathbb{F}_{p^2}$ such that $\pi_{p^2} = p$, so that $\Pi_p^2 = p$. To our knowledge, the use of PPAS over $\mathbb{F}_p$ on which the p-th power Frobenius endomorphism does not act as $\sqrt{-p}$ has not been considered in a cryptographic context before, so these constructions may be found interesting for independent reasons.

Remark C.1. Since Π_p commutes with all other orienting endomorphisms, all oriented PPAS in the orbit of (A, λ, ι) under the action of $\mathfrak{C}(R)$ can again be assumed to be defined over $\mathbb{F}_p$. For use in Appendix G, let us also note that the assumptions $\pi_{p^2} = \pm p$ give rise, respectively, to isomorphisms

$$A(\mathbb{F}_p) \cong E(\mathbb{F}_{p^2}) \cong \left(\frac{\mathbb{Z}}{(p \mp 1)\mathbb{Z}}\right)^2, \; A(\mathbb{F}_{p^2}) \cong E(\mathbb{F}_{p^2}) \times E(\mathbb{F}_{p^2}) \cong \left(\frac{\mathbb{Z}}{(p \mp 1)\mathbb{Z}}\right)^4$$

of abelian groups.

[20] We invite the reader to complete our series with a key exchange protocol for which the name C-3PO would be well-suited.

More precisely, we consider the following method for generating Frobenius-oriented PPASs over $\mathbb{F}_p$, which produces all possible Galois groups. As we have mentioned, we must start from a supersingular elliptic curve $E/\mathbb{F}_{p^2}$ on which $\pi_p^2 = p$, with known endomorphism ring, and let (A, λ) be its Weil restriction. Our strategy is to let

$$K = \mathbb{Q}\left(\sqrt{a + b\sqrt{p}}\right)$$

for $a, b \in \mathbb{Z}$ such that $a + |b|\sqrt{p} < 0$ (so that it concerns a quartic CM field) that will be determined later. We look for an orientation of the form

$$\iota : \sqrt{p} \mapsto \Pi_p = \begin{pmatrix} 0 & \pi_p \\ \pi_p & 0 \end{pmatrix}, \sqrt{a + b\sqrt{p}} \mapsto \begin{pmatrix} \alpha & \beta\pi_p \\ \pi_p\beta & \frac{1}{p}\pi_p\alpha\pi_p \end{pmatrix}$$

for suitable

$$\alpha \in \mathrm{End}(E),\ \beta \in \frac{1}{p}\mathrm{Hom}(E^{(p)}, E)\pi_p \supsetneq \mathrm{End}(E)$$

of trace $\mathrm{tr}(\alpha) = \mathrm{tr}(\beta) = 0$. An easy calculation shows that the right-most matrix commutes with Π_p and is negated when taking the conjugate-transpose. Let us compute

$$\begin{aligned}\begin{pmatrix} \alpha & \beta\pi_p \\ \pi_p\beta & \frac{1}{p}\pi_p\alpha\pi_p \end{pmatrix}^2 &= \begin{pmatrix} p\beta^2 + \alpha^2 & 0 \\ 0 & \pi_p\beta^2\pi_p + \frac{1}{p}\pi_p\alpha^2\pi_p \end{pmatrix} + \begin{pmatrix} 0 & (\alpha\beta + \beta\alpha)\pi_p \\ \pi_p(\alpha\beta + \beta\alpha) & 0 \end{pmatrix} \\ &= (-p\,\mathrm{n}(\beta) - \mathrm{n}(\alpha))\iota(1) + \mathrm{tr}(\alpha\beta)\iota(\sqrt{p})\end{aligned}$$

where we have used that $\mathrm{tr}(\beta) = \mathrm{tr}(\alpha) = 0$. Thus for $a = -p\,\mathrm{n}(\beta) - \mathrm{n}(\alpha)$ and $b = \mathrm{tr}(\alpha\beta)$ the map $\iota : K \to \mathrm{End}^0(A)$ is well-defined. Our key tool to control the Galois group of K is the following lemma:

Lemma C.2. *Under the above assumptions on a, b, the Galois group (of the Galois closure) of K is isomorphic to*

1. *$D_2 \cong C_2 \times C_2$ if $a^2 - b^2p$ is a square,*
2. *C_4 if $a^2 - b^2p$ is not a square but $p(a^2 - pb^2)$ is,*
3. *D_4 if neither $a^2 - pb^2$ nor $p(a^2 - pb^2)$ is a square.*

In particular, if $p \equiv 3 \bmod 4$ then K is either a $C_2 \times C_2$-field or a D_4-field.

Proof. The classification can be found in [26, Lem. 4.7]. As for the last claim, the classification shows that $\mathrm{Gal}(K, \mathbb{Q}) \cong C_4$ if and only if $a^2 - b^2p = pk^2$ for some $k \in \mathbb{Z}$. In particular $a = pa'$ for some integer a' such that $pa'^2 = b^2 + k^2$. Since $p \equiv 3 \bmod 4$ this is only possible if a', b, k are all even. But then we can take out a factor 4 and the argument repeats, leading to a contradiction by infinite descent. □

In the concrete orientations below, the curve E will be constructed as a twist of some well-known supersingular elliptic curve $E_0 : y^2 = f(x)$ defined over $\mathbb{F}_p$:

$$E : vy^2 = f(x), \qquad v \in \mathbb{F}_{p^2} \text{ non-square.}$$

Note that $\pi_{p^2} = -p$ on E_0 (assuming $p > 3$) and therefore we indeed have $\pi_{p^2} = p$ on E. The twisting isomorphisms

$$\tau : E \to E_0 : (x, y) \mapsto (x, uy), \qquad \tau^{(p)} : E^{(p)} \to E_0 : (x, y) \mapsto (x, u^p y),$$

where $u \in \mathbb{F}_{p^4} \setminus \mathbb{F}_{p^2}$ satisfies $u^2 = v$, combine (component-wise) into an isomorphism of PPAVs

$$T : W_{\mathbb{F}_{p^2}/\mathbb{F}_p}(E) \to E_0^2$$

(for the product polarizations). Using $\pi_p \tau = \tau^{(p)} \pi_p$ and $\pi_p \tau^{(p)} = -\tau \pi_p$ one checks

$$T \circ \Pi_p \circ T^{-1} = \begin{pmatrix} 0 & -\pi_p \\ \pi_p & 0 \end{pmatrix}, \quad T \circ \begin{pmatrix} \alpha & \beta\pi_p \\ \pi_p\beta & \frac{1}{p}\pi_p\alpha\pi_p \end{pmatrix} \circ T^{-1} = \begin{pmatrix} \widetilde{\alpha} & -\widetilde{\beta}\pi_p \\ \pi_p\widetilde{\beta} & -\frac{1}{p}\pi_p\widetilde{\alpha}\pi_p \end{pmatrix},$$

where the (quasi-)endomorphisms $\widetilde{\alpha} = \tau\alpha\tau^{-1}$ and $\widetilde{\beta} = \tau\beta\tau^{-1}$ denote the "lollipops" of α, β. Thus, one can also think of ι as an orientation of (E_0^2, λ_0), where for any

$$\widetilde{\alpha} \in \mathrm{End}(E_0), \quad \widetilde{\beta} \in \mathrm{End}(E_0)\frac{1}{\pi_p} = \frac{1}{\pi_p}\,\mathrm{End}(E_0)$$

with trace 0, we obtain an orientation by

$$\mathbb{Z}[\sqrt{a + b\sqrt{p}}] \tag{15}$$

with $a = -p\,\mathrm{n}(\widetilde{\beta}) - \mathrm{n}(\widetilde{\alpha})$ and $b = \mathrm{tr}(\widetilde{\alpha}\widetilde{\beta})$; the advantage of this viewpoint is that it allows us to work directly in $\mathrm{End}(E_0)$, but it becomes less apparent that the orienting endomorphisms on the Weil restriction are $\mathbb{F}_p$-rational.

Remark C.3. This viewpoint also makes for a clearer link with the machinery from Section 5.3: applying Lemma 5.8 to $g = \mathrm{Id}_2$ and the symmetric element

$$P = \begin{pmatrix} 0 & -\pi_p \\ \pi_p & 0 \end{pmatrix} \in M_2(\mathrm{End}(E_0)),$$

we are looking for quadratic subrings of $\mathbb{Q}(P) \otimes_{\mathbb{Q}} \mathrm{End}(E_0)$ of the form (15).

Example C.4 (R2-D4). In view of Lemma C.2, let us assume $p \equiv 3 \bmod 4$. We start from $E_0 : y^2 = x^3 + x$, whose endomorphism ring is of course very well known, see Lemma B.1: it is isomorphic to the maximal order

$$\mathcal{O}_0 = \mathbb{Z}\langle 1, i, \frac{i+j}{2}, \frac{1+k}{2}\rangle \subset B_{p,\infty} = \left(\frac{-1, -p}{\mathbb{Q}}\right)$$

where an isomorphism $\mathcal{O}_0 \to \mathrm{End}(E_0)$ maps i to $(x, y) \mapsto (-x, \mathbf{i}y)$ and j to p-th power Frobenius π_p.

We proceed as above and look for suitable trace-0 elements $\alpha \in \mathcal{O}_0$, $\beta \in \mathcal{O}_0\frac{1}{j}$, where we have dropped the $\widetilde{\cdot}$'s to lighten notation. A $\mathbb{Z}$-basis of the trace-0 sublattice of $\mathcal{O}_0$ is given by i, $\frac{i+j}{2}$, k, so a general trace-0 element is

$$\alpha = a_1 i + a_2 \frac{i+j}{2} + a_3 k.$$

For β, a direct computation of the reduced trace on $\frac{1}{j}\mathcal{O}_0$ shows that $\frac{1}{j}$, $\frac{i}{j}$, $\frac{1+k}{2j}$ form a $\mathbb{Z}$-basis of its trace-0 sublattice, so

$$\beta = \frac{1}{j}\big(b_1 + b_2 i + b_3 \frac{1+k}{2}\big)$$

for integers $a_1, a_2, a_3, b_1, b_2, b_3$. This yields

$$\begin{aligned} a &= -a_1^2 - a_1 a_2 - \frac{p+1}{4} a_2^2 - p a_3^2 - b_1^2 - b_1 b_3 - \frac{p+1}{4} b_3^2 - b_2^2, \\ b &= a_2 b_1 + (a_1 + a_2) b_3 - 2 a_3 b_2. \end{aligned}$$

For each choice of a_i, b_i such that $a^2 - b^2 p$ is a non-square (generically the case) we obtain an orientation of type D_4 by Lemma C.2. The condition $a + |b|\sqrt{p} < 0$ is automatically satisfied because every quartic number field contained in $M_2(B_{p,\infty})$ is necessarily totally imaginary.

As for the primitivity, an easy calculation shows

$$\operatorname{disc} \mathbb{Z}[\sqrt{a + b\sqrt{p}}] = 256 b^4 p^2 (a^2 - b^2 p). \tag{16}$$

From [48, Thm. 1 and Tbl. B] we see that for $p \equiv 3 \bmod 4$ this equals $\operatorname{disc} K$ as soon as $b = \pm 1$ and $a^2 - p$ is square-free. So in these cases the order is maximal and our orientation is automatically primitive. A near-maximal example where $|b| = 4$ will be studied in Appendix F.

Example C.5 (R2-C4). We follow the same recipe to realize C_4. Lemma C.2 tells us that we must move away from characteristics $p \equiv 3 \bmod 4$. We therefore start from the elliptic curve $E_0 : y^2 = x^3 + 12x^2 + 18x$ over $\mathbb{F}_{p^2}$, where $p \equiv 5 \bmod 8$. We refer to Lemma B.4 for an explicit description of its endomorphism ring, which is isomorphic to the maximal order

$$\mathbb{Z}\left\langle 1, i, \frac{1+i+j}{2}, \frac{i+2j \pm k}{4} \right\rangle \subset B_{p,\infty} = \left(\frac{-2, -p}{\mathbb{Q}}\right),$$

with the sign depending on whether $p \equiv 5, 13 \bmod 48$ or $p \equiv 29, 37 \bmod 48$. For concreteness, we proceed under the assumption that $p \equiv 5, 13 \bmod 48$, corresponding to the $+$-sign. The trace-0 condition imposes

$$\begin{aligned} \alpha &= a_1 i + a_2 (i+j) + a_3 \frac{i+2j+k}{4}, \\ \beta &= \frac{1}{j}\left(b_1 + b_2 i + b_3 \frac{1+i+j}{2} - b_3 \frac{i+2j+k}{4}\right) \end{aligned}$$

for integers $a_1, a_2, a_3, b_1, b_2, b_3$, so we end up with expressions

$$a = -2a_1^2 - (p+2)a_2^2 - 4a_1a_2 - a_1a_3 - (p+1)a_2a_3 \\ - \frac{3p+1}{8}a_3^2 - b_1^2 - 2b_2^2 - b_1b_3 - b_2b_3 - \frac{p+3}{8}b_3^2,$$
$$b = -a_1b_3 + 2a_2b_1 + a_3b_1 - a_3b_2.$$

Again this will generically result in orientations of type D_4. To end up in the C_4 case, we write $p = x^2+y^2$ as a sum of two squares (possible because $p \equiv 1 \bmod 4$) and look for a non-trivial zero of the quadratic form

$$xa - pb \in \mathbb{Z}[a_1, a_2, a_3, b_1, b_2, b_3].$$

Then $0 = (xa - pb)(xa + pb) = (p - y^2)a^2 - p^2b^2 = p(a^2 - pb^2) - a^2y^2$ so that $p(a^2 - pb^2)$ is a square as wanted.

Controlling primitivity for R2-C4 is harder than for R2-D4. Indeed, by construction the factor $a^2 - pb^2$ is a square away from factors p, and following [48, Thm. 1 and Tbl. C] this makes our order non-maximal, typically. An example family of solutions is where $p = 1 + 4m^2$ for some integer $m \equiv 5 \bmod 6$ (so that $p \equiv 5 \bmod 8$ and $p \not\equiv 29, 37 \bmod 48$): we can take

$$a_1 = -m+1, a_2 = -1, a_3 = 2, b_1 = -1, b_2 = m, b_3 = 2 \quad \Rightarrow \quad a = -2p, b = -2$$

yielding $a^2 - pb^2 = 16pm^2$. So $\operatorname{disc} K$ is smaller than (16) by a factor m^2, at least.

Remark C.6. Many of the examples we found satisfy $p \mid a, b$ and consequently $p^3 \mid a^2 - b^2p$ (since $p(a^2 - b^2p)$ has to be a square), in which case one checks that

$$\gamma = \frac{\sqrt{a + b\sqrt{p}}}{\sqrt{p}}$$

is integral. If we want our orientation to extend to $\mathbb{Z}[\gamma] \subset \mathbb{Z}_K$ then

$$\iota(\sqrt{a + b\sqrt{p}}) = \begin{pmatrix} \alpha & \beta\pi_p \\ \pi_p\beta & \frac{1}{p}\pi_p\alpha\pi_p \end{pmatrix} = \begin{pmatrix} 0 & \pi_p \\ \pi_p & 0 \end{pmatrix} \begin{pmatrix} \beta & \frac{1}{p}\alpha\pi_p \\ \frac{1}{p}\pi_p\alpha & \frac{1}{p}\pi_p\beta\pi_p \end{pmatrix}$$

should factor through Π_p, i.e., the right-most factor should really be an element of $\operatorname{End}(A)$. However, this is not always the case. The bottleneck is the bottom-right entry, which is not always free from denominators; recall that we sample β from $(1/p)\operatorname{Hom}(E^{(p)}, E)\pi_p$. E.g., again for $p = 1 + 4m^2$ we can consider

$$a_1 = -2m^2 - m + 1, a_2 = -2m - 1, a_3 = 4m + 2, \\ b_1 = 2m - 1, b_2 = -2m^2 + 3m, b_3 = -4m + 2$$

(leading to $a = -2p^2, b = -2p$) and one checks that the coefficient at k of $j\beta j$ equals $b_2 + b_3/4 \equiv 2m + 1 \not\equiv 0 \bmod p$. So this example shows that $\mathcal{M}_k(R)$ can be non-empty even if $p \mid [\mathbb{Z}_K : R]$: this cannot occur in dimension $g = 1$.

D Proofs from Section 4 (except for Lemma 4.3)

D.1 Proof from Section 4.1

Proof (of Lemma 4.1). If $K = \mathrm{Frac}(R)$ is Galois (cases $C_2 \times C_2, C_4$) then ℓR necessarily factors into prime ideals of equal degree $4/r$ for $r \in \{1, 2, 4\}$ and the analysis is easy. In the D_4-case there is an exception to this: it may happen that $\ell R = \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}$ for degree-1 primes $\mathfrak{p}_1, \mathfrak{p}_2$ and a degree-2 prime $\mathfrak{p}$. But then necessarily $\mathfrak{p}$ is self-conjugate and this does not lead to a suitable factorization of ℓR. □

D.2 Proofs from Section 4.2

Proof (of Theorem 4.4). The product polarization on A_i^u gives that

$$\iota_{A_i}(\alpha)^\dagger = (\iota_i(\alpha_{k,j})^\dagger)_{1\le k,j\le u} = (\iota_i(\bar{\alpha}_{k,j}))_{1\le j,k\le u} = \iota_{A_i}(\alpha^*)$$

for any $\alpha \in M_u(R)$. Moreover, $\iota_{A_i}(\alpha)^\dagger \circ \iota_{A_i}(\alpha) = \iota_{A_i}(n\mathrm{Id}_u)$ because $\iota_{A_i} : M_u(R) \to \mathrm{End}(A_i^u)$ is an injective ring homomorphism and $\alpha^*\alpha = n\mathrm{Id}_u \in M_u(R)$. Then, $\Phi^\dagger = \begin{pmatrix} \varphi^\dagger \mathrm{Id}_u & -\alpha_{A_1}^\dagger \\ \alpha_{A_2} & \varphi\,\mathrm{Id}_u \end{pmatrix}$. Using $\alpha^*\alpha = \alpha\alpha^* = n\mathrm{Id}_u$, $\varphi^\dagger \circ \varphi = \iota_1(d)$, $\varphi \circ \varphi^\dagger = \iota_2(d)$, the four blocks of $\Phi^\dagger\Phi$ are

$$\begin{aligned}
&(1,1) : \varphi^\dagger\varphi\,\mathrm{Id}_u + \alpha_{A_1}^\dagger\alpha_{A_1} = \iota_1(d+n)\mathrm{Id}_u = \iota_1(N)\mathrm{Id}_u,\\
&(1,2) : \varphi^\dagger\alpha_{A_2}^\dagger - \alpha_{A_1}^\dagger\varphi^\dagger\mathrm{Id}_u = 0,\\
&(2,1) : \alpha_{A_2}\varphi\,\mathrm{Id}_u - \varphi\,\alpha_{A_1}\mathrm{Id}_u = 0,\\
&(2,2) : \alpha_{A_2}\alpha_{A_2}^\dagger + \varphi\varphi^\dagger\mathrm{Id}_u = \iota_2(n+d)\mathrm{Id}_u = \iota_2(N)\mathrm{Id}_u,
\end{aligned}$$

where in $(1,2)$ and $(2,1)$ we used that φ is oriented, hence $\varphi^\dagger \circ \iota_2(\alpha)^\dagger = \iota_1(\alpha)^\dagger \circ \varphi^\dagger$ and $\varphi \circ \iota_1(\alpha) = \iota_2(\alpha) \circ \varphi$. Since $\iota_i(N) = [N]$, it follows that $\Phi^\dagger\Phi = [N]$ and Φ is an N-isogeny of PPAVs of dimension $2ug$, of degree N^{2ug}.

(i) $\Phi(P, 0) = (\varphi\,\mathrm{Id}_u(P),\ -\alpha_{A_1}(P))$ for $P \in A_1^u$; taking $P = (P_1, 0, \ldots, 0)$ the first coordinate of the image is $\varphi(P_1)$.

(ii) For $P \in A_1^u[N]$,

$$\begin{aligned}
&\Phi(-\alpha_{A_1}^\dagger(P),\ \varphi\,\mathrm{Id}_u(P)) =\\
&\qquad (-\varphi\,\alpha_{A_1}^\dagger(P) + \alpha_{A_2}^\dagger\varphi(P),\ \alpha_{A_1}\alpha_{A_1}^\dagger(P) + \varphi^\dagger\varphi(P)) = (0,\ [N]P) = 0.
\end{aligned}$$

The map $P \mapsto (-\alpha_{A_1}^\dagger(P), \varphi\,\mathrm{Id}_u(P))$ is injective on $A_1^u[N]$: indeed, if both coordinates vanish, then $P \in A_1^u[nR_0] \cap A_1^u[dR_0]$; coprimality of nR_0 and dR_0 then implies that $xn + yd = 1$ for suitable $x, y \in R_0$, hence $P = 0$. Since N is coprime to the base field characteristic and $\#A_1^u[N] = N^{2ug} = \deg\Phi$, the image coincides with $\ker\Phi$. For the second equality, since $nR_0 + NR_0 = nR_0 + dR_0 = R_0$, then nR_0 and NR_0 are coprime, hence α_{A_1} acts as an automorphism on $A_1^u[N]$: indeed, $A_1[N]$ is a finite group and if $P \in A_1^u[N]$ and $\alpha_{A_1}(P) = 0$, then

$P \in A_1^u[nR_0] \cap A_1^u[NR_0]$, and again $P = 0$ by coprimality of nR_0 and NR_0. Then, substituting $P \mapsto \alpha_{A_1}(P)$ and using $-\iota_1(n)\mathrm{Id}_u = \iota_1(d)\mathrm{Id}_u$ on $A_1^u[N]$, and $\varphi \circ \iota_1(\alpha) = \iota_2(\alpha) \circ \varphi$, we obtain the stated subgroup. □

Remark D.1 (continuation of Remark 4.5). We have that $\Phi = \Psi_2 \circ \Psi_1$ with Ψ_1, Ψ_2 both $\sqrt{N}$-isogenies and $\ker \Psi_1 = \sqrt{N} \ker \Phi$, hence

$$\ker \Psi_1 = \{(-\alpha_{A_1}^\dagger(P),\ \varphi\,\mathrm{Id}_u(P)) \mid P \in A_1^u[\sqrt{N}]\}.$$

The Rosati dual $\Phi^\dagger$ is again of the shape of the theorem, up to sign flips in the α_{A_i}'s, so the same argument yields

$$\begin{aligned}\ker \Psi_2^\dagger &= \{(\alpha_{A_2}^\dagger(Q),\, \varphi^\dagger \mathrm{Id}_u(Q)) \mid Q \in A_2^u[\sqrt{N}]\} \\ &= \{(\alpha_{A_2}^\dagger \varphi\, \mathrm{Id}_u(P),\ \iota_1(d)\mathrm{Id}_u(P)) \mid P \in A_1^u[\sqrt{N}]\}.\end{aligned}$$

One computes Ψ_1 and $\Psi_2^\dagger$ from their kernels and glues them along an isomorphism between their codomains; the details of this gluing in the middle are technical, but are discussed in [33, App. F], for instance.

Proof (of Proposition 4.6). Since $\alpha^* = \left(\begin{smallmatrix} X^* & -Y^* \\ Y & X \end{smallmatrix}\right)$, we have

$$\alpha^*\alpha = \begin{pmatrix} X^*X + Y^*Y & (YX - XY)^* \\ YX - XY & XX^* + YY^* \end{pmatrix}.$$

Under the hypotheses of (a) we have $X^*X = XX^* = x\mathrm{Id}_m$ and $Y^*Y = YY^* = y\mathrm{Id}_m$, so both diagonal blocks equal $(x+y)\mathrm{Id}_m$ and the off-diagonal blocks vanish. The conclusion for $\alpha\alpha^*$ follows similarly.
The hypotheses of (b) are exactly the conditions under which $\alpha^*\alpha$ equals $n\,\mathrm{Id}_{2m}$.
For (c), the claim follows by induction on k. The case $i = 0$ is proven by $\bar{\vartheta}_0\vartheta_0 = \vartheta_0\bar{\vartheta}_0 = N_{K/K_0}(\vartheta_0)$. For $i \geq 1$, a scalar matrix $\vartheta\mathrm{Id}_m$ satisfies $(\vartheta\mathrm{Id}_m)^*(\vartheta\mathrm{Id}_m) = (\vartheta\mathrm{Id}_m)(\vartheta\mathrm{Id}_m)^* = N_{K/K_0}(\vartheta)\,\mathrm{Id}_m$ and commutes with every matrix, so (a) applies at each step with $X = \alpha_{i-1}$ and $Y = \vartheta_i\,\mathrm{Id}_{2^{i-1}}$, adding the term $N_{K/K_0}(\vartheta_i)$ to n.
For (d), the claim again follows by induction and similarly as for (c), since the case $i = 0$ holds by the hypotheses on α_0 and the hypotheses of (a) apply at every step $i \geq 1$ with $X = \alpha_{i-1}$, $x = n_1 + \sum_{j=1}^{i-1} N_{K/K_0}(\vartheta_j')$, and $Y = \vartheta_i'\mathrm{Id}_{2^{k+i-1}}$. The last statement follows taking α_0 as the matrix from [85, Lemma 4.2]. □

E Proof of Lemma 4.3

This appendix is devoted to a detailed proof of Lemma 4.3. We also prove an easy corollary. Let us first recall the statement of the Lemma.

Lemma 4.3. *Assume that K is a $C_2 \times C_2$ CM field, and let K_0 be the real quadratic subfield and K_1, K_2 the imaginary quadratic subfields. If $\mathbb{Z}_K^\times$ contains*

no roots of unity other than ± 1, then

$$|\ker \kappa_{\mathbb{Z}_K}| \leq \frac{1}{2q(K)} \prod_{\ell \text{ prime}} e(\ell),$$
$$|\operatorname{coker} \kappa_{\mathbb{Z}_K}| \leq \frac{1}{4q(K)} |\operatorname{coker} N_{K/K_0}| \prod_{\ell \text{ prime}} e(\ell),$$

where $q(K) = [\mathbb{Z}_K^\times : \mathbb{Z}_{K_0}^\times] \in \{1, 2\}$, $e(\ell)$ is the ramification index of a prime ℓ in K, and $N_{K/K_0} : \mathrm{Cl}(\mathbb{Z}_K) \to \mathrm{Cl}^+(\mathbb{Z}_{K_0})$ denotes the ideal class norm map.

Corollary E.1. *Assume that K is a $C_2 \times C_2$ CM field with $\mathbb{Z}_K^\times$ containing no roots of unity other than ± 1. Let R be an order in K and $R_i = R \cap K_i$ for $i = 1, 2$. Then,*

$$|\ker \kappa_R| \leq \frac{C_1 C_2}{2q(K)} \prod_{\ell \text{ prime}} e(\ell)$$

where $C_i = f_i \prod_{p \text{ prime},\ p|f_i} \left(1 - \left(\frac{d_{K_i}}{p}\right) \frac{1}{p}\right)$, with $f_i = [\mathbb{Z}_{K_i} : R_i]$ the conductor of R_i, and d_{K_i} the discriminant of K_i, for $i = 1, 2$.

Proof. If $([\mathfrak{a}_1], [\mathfrak{a}_2]) \in \mathrm{Cl}(R_1) \times \mathrm{Cl}(R_2)$ is in the kernel of κ_R, then

$$([\mathfrak{a}_1 \mathbb{Z}_{K_1}], [\mathfrak{a}_2 \mathbb{Z}_{K_2}]) \in \mathrm{Cl}(\mathbb{Z}_{K_1}) \times \mathrm{Cl}(\mathbb{Z}_{K_2})$$

is in the kernel of $\kappa_{\mathbb{Z}_K}$. This yields a well-defined group homomorphism $\psi : \ker \kappa_R \to \ker \kappa_{\mathbb{Z}_K}$ whose kernel consists of the pairs in $\ker \kappa_R$ of classes of invertible ideals of R_1, R_2 that become principal when extended to $\mathbb{Z}_{K_1}$, $\mathbb{Z}_{K_2}$, respectively. Hence,

$$|\ker \psi| \leq \frac{h(R_1)}{h(\mathbb{Z}_{K_1})} \cdot \frac{h(R_2)}{h(\mathbb{Z}_{K_2})}$$

where h denotes the class number, and

$$\frac{h(R_i)}{h(\mathbb{Z}_{K_i})} = f_i \prod_{\substack{p \text{ prime} \\ p|f_i}} \left(1 - \left(\frac{d_{K_i}}{p}\right) \frac{1}{p}\right)$$

using [30, Theorem 7.24], and remembering the assumption that the only roots of unity in $\mathbb{Z}_K$ (and hence in $\mathbb{Z}_{K_i}$ and R_i) are ± 1. Together with Lemma 4.3, the inequality $|\ker \kappa_R| \leq |\ker \psi||\ker \kappa_{\mathbb{Z}_K}|$ concludes the proof. □

To prove the lemma we will make use of the following results from [64].

Lemma E.2 ([64, 1.3]). *Let K be a $C_2 \times C_2$ extension of $\mathbb{Q}$ and let K_0, K_1, K_2 be the three intermediate quadratic fields. The map*

$$j : \mathrm{Cl}(\mathbb{Z}_{K_0}) \times \mathrm{Cl}(\mathbb{Z}_{K_1}) \times \mathrm{Cl}(\mathbb{Z}_{K_2}) \to \mathrm{Cl}(K) : ([\mathfrak{a}_0], [\mathfrak{a}_1], [\mathfrak{a}_2]) \mapsto [\mathfrak{a}_0 \mathbb{Z}_K \mathfrak{a}_1 \mathbb{Z}_K \mathfrak{a}_2 \mathbb{Z}_K]$$

is a group homomorphism. Taking the restriction map $\widehat{j} : \widehat{C} \to \mathrm{Cl}(\mathbb{Z}_K)$ *where*

$$\widehat{C} = \{([\mathfrak{a}_0], [\mathfrak{a}_1], [\mathfrak{a}_2]) \in \mathrm{Cl}(\mathbb{Z}_{K_0}) \times \mathrm{Cl}(\mathbb{Z}_{K_1}) \times \mathrm{Cl}(\mathbb{Z}_{K_2}) \mid N_{K_0/\mathbb{Q}}(\mathfrak{a}_0) N_{K_1/\mathbb{Q}}(\mathfrak{a}_1) N_{K_2/\mathbb{Q}}(\mathfrak{a}_2) = 1\}$$

gives a surjection $\operatorname{coker} \widehat{j} \twoheadrightarrow \operatorname{coker} j$ *and*

$$|\operatorname{coker} \widehat{j}| = \frac{2^{d-2} \prod_{\ell \text{ prime}} e(\ell)}{[\mathbb{Z}^\times : H]},$$

where d *is the number of infinite primes that ramify in* $K/\mathbb{Q}$*,* $e(\ell)$ *is the ramification index of a prime* ℓ *in the extension* $K/\mathbb{Q}$*, and* $H \leq \mathbb{Z}^\times$ *is the subgroup of integer units that are in the image of the norm map* $N_{K/\mathbb{Q}}$*.*

Lemma E.3 ([64, 1.5]). *Let* K *be a CM* $C_2 \times C_2$ *field, with quadratic subfields* K_0, K_1, K_2*. Then*

$$h(\mathbb{Z}_K) = \frac{1}{2} q(K) h(\mathbb{Z}_{K_0}) h(\mathbb{Z}_{K_1}) h(\mathbb{Z}_{K_2})$$

where h *denotes the class number and* $q(K) = [\mathbb{Z}_K^\times : \mathbb{Z}_{K_0}^\times \mathbb{Z}_{K_1}^\times \mathbb{Z}_{K_2}^\times]$ *is the unit index.*

Moreover, we will need the following formulas for the narrow class number h^+ and the order of the Shimura class group.

Lemma E.4. *Let* L *be a number field, then*

$$h^+(\mathbb{Z}_L) = \frac{2^r}{[\mathbb{Z}_L^\times : (\mathbb{Z}_L^\times)^+]} h(\mathbb{Z}_L),$$

where r *is the number of real embeddings* $\sigma_i : L \hookrightarrow \mathbb{C}$*,* $i \in \{1, \dots r\}$*.*

Proof. There is an exact sequence

$$1 \to \frac{\mathbb{Z}_L^\times}{(\mathbb{Z}_L^\times)^+} \xrightarrow{u \mapsto u} \frac{L^\times}{(L^\times)^+} \xrightarrow{u \mapsto [u\mathbb{Z}_L]} \mathrm{Cl}^+(\mathbb{Z}_L) \xrightarrow{[\mathfrak{a}] \mapsto [\mathfrak{a}]} \mathrm{Cl}(\mathbb{Z}_L) \to 1.$$

The formula follows from surjectivity of the sign map $\operatorname{sgn} : L^\times \to \{\pm 1\}^r$ sending elements $x \in L$ to $(\operatorname{sgn}(\sigma_1(x)), \dots, \operatorname{sgn}(\sigma_r(x))) \in \{\pm 1\}^r$, with kernel $(L^\times)^+$. □

Lemma E.5. *Let* K *be a CM* $C_2 \times C_2$ *field. If* $\mathbb{Z}_K^\times$ *contains no roots of unity other than* ± 1*, then*

$$|\mathfrak{C}(\mathbb{Z}_K)| = \frac{1}{2} h(\mathbb{Z}_{K_1}) h(\mathbb{Z}_{K_2}) |\operatorname{coker} N_{K/K_0}|$$

where $N_{K/K_0} : \mathrm{Cl}(\mathbb{Z}_K) \to \mathrm{Cl}^+(\mathbb{Z}_{K_0})$ *is the ideal class norm map.*

Proof. From the exact sequence in Section 2.2, it follows that

$$|\mathfrak{C}(\mathbb{Z}_K)| = [(\mathbb{Z}_{K_0}^\times)^+ : N_{K/K_0}(\mathbb{Z}_K^\times)] \frac{h(\mathbb{Z}_K)}{|\operatorname{im} N_{K/K_0}|}.$$

Using Lemma E.3, and Lemma E.4 applied to K_0, we can rewrite this as

$$\frac{1}{8} \frac{h^+(\mathbb{Z}_{K_0})}{|\operatorname{im} N_{K/K_0}|} h(\mathbb{Z}_{K_1}) h(\mathbb{Z}_{K_2}) q(K) [\mathbb{Z}_{K_0}^\times : (\mathbb{Z}_{K_0}^\times)^+] [(\mathbb{Z}_{K_0}^\times)^+ : N_{K/K_0}(\mathbb{Z}_K^\times)].$$

With the assumption that $\mathbb{Z}_K^\times$ contains no roots of unity other than ± 1, then $q(K) = [\mathbb{Z}_K^\times : \mathbb{Z}_{K_0}^\times \mathbb{Z}_{K_1}^\times \mathbb{Z}_{K_2}^\times] = [\mathbb{Z}_K^\times : \mathbb{Z}_{K_0}^\times]$, yielding

$$|\mathfrak{C}(\mathbb{Z}_K)| = \frac{1}{8} |\operatorname{coker} N_{K/K_0}| h(\mathbb{Z}_{K_1}) h(\mathbb{Z}_{K_2}) [\mathbb{Z}_K^\times : N_{K/K_0}(\mathbb{Z}_K^\times)].$$

Again under the assumption that $\mathbb{Z}_K^\times$ contains no roots of unity other than ± 1, it is not difficult to show that $[\mathbb{Z}_K^\times : N_{K/K_0}(\mathbb{Z}_K^\times)] = 4$, which ends the proof. □

We are now ready to prove Lemma 4.3.

Proof. First observe that if $([\mathfrak{a}_1], [\mathfrak{a}_2]) \in \operatorname{Cl}(\mathbb{Z}_{K_1}) \times \operatorname{Cl}(\mathbb{Z}_{K_2})$ is in the kernel of $\kappa_{\mathbb{Z}_K}$, then $([1], [\mathfrak{a}_1], [\mathfrak{a}_2]) \in \operatorname{Cl}(\mathbb{Z}_{K_0}) \times \operatorname{Cl}(\mathbb{Z}_{K_1}) \times \operatorname{Cl}(\mathbb{Z}_{K_2})$ is in the kernel of the map j from Lemma E.2, namely there is an injection $\ker \kappa_{\mathbb{Z}_K} \hookrightarrow \ker j$, from which we have

$$|\ker \kappa_{\mathbb{Z}_K}| \leq |\ker j| = \frac{h(\mathbb{Z}_{K_0}) h(\mathbb{Z}_{K_1}) h(\mathbb{Z}_{K_2})}{h(\mathbb{Z}_K)} |\operatorname{coker} j|.$$

Using Lemma E.3, then

$$|\ker j| = \frac{2}{q(K)} |\operatorname{coker} j|,$$

where one can see that $q(K) \in \{1, 2\}$, following e.g. [71, Proposition 5.12]. Hence,

$$|\ker j| \leq \frac{2}{q(K)} |\operatorname{coker} \widehat{j}| \leq \frac{1}{2q(K)} \prod_{\ell \text{ prime}} e(\ell)$$

using Lemma E.2 for the second inequality, with $d = 1$ (as the real infinite prime of $\mathbb{Q}$ ramifies in K) and $[\mathbb{Z}^\times : H] = 2$ (as -1 is not a norm in the totally imaginary field K). This proves the bound on $|\ker \kappa_{\mathbb{Z}_K}|$.

On the other hand,

$$|\operatorname{coker} \kappa_{\mathbb{Z}_K}| = \frac{|\mathfrak{C}(\mathbb{Z}_K)|}{h(\mathbb{Z}_{K_1}) h(\mathbb{Z}_{K_2})} |\ker \kappa_{\mathbb{Z}_K}| = \frac{1}{2} |\operatorname{coker} N_{K/K_0}| |\ker \kappa_{\mathbb{Z}_K}|$$

by Lemma E.5, which proves the bound on $|\operatorname{coker} \kappa_{\mathbb{Z}_K}|$. □

Example E.6. We consider the R2-D2 orientation introduced in Example 3.4 and used in Section F.2 below. Here $K = \mathbb{Q}(\sqrt{d}, \sqrt{-p})$ for a prime $p > 3$ and a square-free integer d coprime to p. Then $K_0 = \mathbb{Q}(\sqrt{d})$, and say $K_1 = \mathbb{Q}(\sqrt{-p})$, $K_2 = \mathbb{Q}(\sqrt{-dp})$. Assume that $d = q_1 q_2$, where $q_1, q_2 > 3$ are prime. Assume moreover $p \equiv d \equiv 3 \bmod 4$. Then $\operatorname{disc} \mathbb{Z}_K = 2^4 p^2 q_1^2 q_2^2$ by [98, Thm. 3]. There are then exactly four ramified primes in $K/\mathbb{Q}$, namely $2, p, q_1, q_2$, and since each of them is ramified in exactly two quadratic subfields, we have $e(2) = e(p) = e(q_1) = e(q_2) = 2$. From Lemma 4.3 it follows that $|\ker \kappa_{\mathbb{Z}_K}| \leq 8/q(K)$. Since the extension K/K_0 is ramified at a prime ideal of $\mathbb{Q}(\sqrt{d})$ above p, the norm map N_{K/K_0} is surjective on $\mathrm{Cl}^+(\mathbb{Z}_{K_0})$, hence $|\operatorname{coker} \kappa_{\mathbb{Z}_K}| \leq 4/q(K)$.

F Key exchange from R2

We now delve into the details of generating concrete parameters for a Diffie–Hellman style key exchange protocol built from Shimura class group actions, along the usual diagram depicted in Figure 1. We emphasize that the construction below was done for proof-of-concept purposes only: we leave considerable room for optimization, and the security remains to be assessed more thoroughly. For the appealing reasons mentioned in Section 4.4, we concentrate on the R2 orientations from Appendix C (see also Example 3.4), where the entire orbit is defined over $\mathbb{F}_p$. The "best" family to work with seems to be R2-D4. Indeed, in R2-C4 the parameters are quite constrained, hindering parameter generation, while R2-D2 has security issues [50].

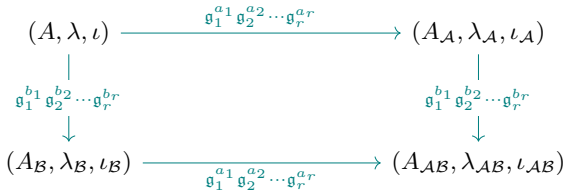

Fig. 1. Diffie–Hellman key exchange based on Shimura class group actions, for a set of "generators" $\mathbf{g} = \{\mathfrak{g}_1, \ldots, \mathfrak{g}_r\} \subset \mathfrak{C}(R)$ and exponents sampled from $\{-B, \ldots, B\}^r$.

F.1 Key exchange from R2-D4

Recall that R2-D4 is a (not necessarily primitive) orientation by $R = \mathbb{Z}[\gamma]$ with

$$\gamma = \sqrt{a + b\sqrt{p}}, \qquad a + |b|\sqrt{p} < 0.$$

Note that

$$\bar{\gamma}\gamma = d := -a - b\sqrt{p} \in R_0^+$$

so in order to represent $\iota(\gamma)$, as explained in Section 4.3 we can stay in dimension 2 as soon as we can write $N - d = \alpha^2$ for some $\alpha = s + t\sqrt{p} \in \mathbb{Z}[\sqrt{p}]$. This spells out as:

- $2st = b$,
- $s^2 + pt^2 = N + a$,

showing that we cannot let $b = \pm 1$, which is a necessary condition for $\mathbb{Z}[\gamma]$ to be maximal. We choose $b = 4$, and then we can assume that $s = 2$, $t = 1$ so that $a = p + 4 - N$.[21]

Enforcing complete splitting of small primes. For an efficient evaluation of the Shimura class group action in the **g**-REGA model, we want that $A[\ell]$ is "easily accessible", for every small odd prime ℓ involved in the construction of **g**. In view of Remark C.1 the most natural sufficient condition is $\ell \mid p - 1$, so that $A[\ell] \subset A(\mathbb{F}_{p^2})$. Then ℓ splits completely in R if and only if $a \pm 4$ are both non-zero quadratic residues modulo ℓ. Interestingly, since

$$a = p + 4 - N \equiv 5 - N \bmod \ell,$$

this condition does not explicitly depend on a and/or p. So the idea is to collect small prime numbers $\ell_1, \ell_2, \ldots, \ell_r$ modulo which both $9 - N$ and $1 - N$ are non-zero quadratic residues; for a fixed value of N we expect this to hold for a proportion $\frac{1}{4}$ of the primes (within the wiggle room one typically has when choosing N, one can give preference to the value giving rise to the smallest ℓ_i). We then choose

$$p = Mc\ell_1\ell_2 \cdots \ell_r + 1$$

with M a large factor of N and c a small cofactor ensuring that p is a prime number. Here, more precisely, we take $N = M^2$ so that $A[\sqrt{N}] \subset A(\mathbb{F}_{p^2})$ for the purpose of meet-in-the-middle.

Realization as an instance of R2-D4. For an incorporation into the R2-D4 family we need $|a| > p/4$ and consequently $N = p + 4 - a > 5p/4$. This comes about as follows. Recall from Example C.4 that our task is to find integers $a_1, a_2, a_3, b_1, b_2, b_3$ such that

$$\begin{aligned} a &= -a_1^2 - a_1a_2 - \frac{p+1}{4}a_2^2 - pa_3^2 - b_1^2 - b_1b_3 - \frac{p+1}{4}b_3^2 - b_2^2, \\ 4 = b &= a_2b_1 + (a_1 + a_2)b_3 - 2a_3b_2. \end{aligned}$$

These formulae can be rewritten as

$$\begin{aligned} -4a &= A^2 + B^2 + 4b_2^2 + p(a_2^2 + 4a_3^2 + b_3^2), \\ 8 &= Ab_3 + Ba_2 - 4a_3b_2, \end{aligned}$$

[21] The reason we skip $b = 2$ is that in this case $a = p+1-N$. Since we will choose p such that $p - 1$ and N share a large prime power M, implying $a^2 - b^2p \equiv (2 - N)^2 - 4 \equiv 0 \bmod M$, it follows that $\operatorname{disc} K$ plummets in view of [48, Thm. 1 and Tbl. B].

where $A = 2a_1 + a_2$ and $B = 2b_1 + b_3$. So the bound $|4a| > p$ is immediate. Typically this is not enough: we want some extra leeway in order for the following solving strategy to succeed: iterate over positive integers $n = 1, 2, 3, \ldots$ and over all vectors $(a_2, 2a_3, b_3) \in \mathbb{Z} \times 2\mathbb{Z} \times \mathbb{Z}$ of squared-Euclidean length n until we can find a corresponding vector $(A, B, 2b_2)$ of squared-Euclidean length $-4a - np$ in the affine plane

$$H : b_3 X + a_2 Y - 2a_3 Z = 8,$$

where moreover we want that $A \equiv a_2 \bmod 2$ and $B \equiv b_3 \bmod 2$, so that $a_1 = (A - a_2)/2$ and $b_1 = (B - b_3)/2$ are integers.

This can be turned into an efficient algorithm, but for generating the concrete parameters below we have used a simplified search by restricting to $a_3 = 0$ and $\gcd(a_2, b_3) = 1$; this is good enough for our proof-of-concept purposes, although it increases the expected number of n's to be tried and therefore requires a somewhat larger gap between $|4a|$ and p. To proceed, we use the extended Euclidean algorithm to compute $\lambda, \mu \in \mathbb{Z}$ such that $1 = \lambda b_3 + \mu a_2$. Then for any x, b_2 the tuple $(8\lambda + x a_2, 8\mu - x b_3, 2b_2)$ lies in H. Expressing that it has squared-Euclidean length $-4a - np$ amounts to

$$(nx + 8(\lambda a_2 - \mu b_3))^2 + 4nb_2^2 = -4na - n^2 p - 64.$$

This asks to represent the right-hand side by the positive definite binary quadratic form $X^2 + 4nY^2$. If this right-hand side is prime, then we expect it to be representable with probability $1/2h(\mathbb{Z}[\sqrt{-4n}]) \approx 1/\sqrt{n}$ by [30, Thm. 9.12] and in that case a corresponding representation can be found efficiently using Cornacchia's algorithm. Three remarks are in order:

- We model $-4na - n^2 p - 64$ as a random integer of size polynomial in p, so that we expect success after a number of tries that is polylogarithmic in p. However if $3 \mid a$ and $p \equiv 2 \bmod 3$ then using $3 \nmid n$ (indeed, n is a sum of coprime squares) one sees that $-4na - n^2 p - 64$ is always divisible by 3, so in these cases the modeling is clearly flawed.[22] But in our parameters below we have $3 \nmid a$ and $p \equiv 1 \bmod 3$ and this is not a concern.
- If this integer can indeed be represented as $c^2 + 4nd^2$ for certain $c, d \in \mathbb{Z}$, then there is no guarantee that $c = nx + 8(\lambda a_2 - \mu b_3)$ for some $x \in \mathbb{Z}$. This may seem a strong additional assumption but luckily it is not: one checks that $(\lambda a_2 - \mu b_3)^2 \equiv -1 \bmod n$ and therefore $c^2 \equiv (8(\lambda a_2 - \mu b_3))^2 \bmod n$. E.g., if n is prime then this implies $c \equiv \pm 8(\lambda a_2 - \mu b_3) \bmod n$ so by changing the sign of c if needed we can ensure that c is of the form $nx + 8(\lambda a_2 - \mu b_3)$. If n is not prime then this may fail but in any case, in view of the Erdös–Kac theorem, the inverse probability of success is in the order of $\log n$ rather than in the order of n as one might expect at first sight.
- If n is even then $-4na - n^2 p - 64 \equiv 0 \bmod 4$ so the primality assumption forces n to be odd. It is then easy to check that the requirements $A \equiv a_2 \bmod 2$ and $B \equiv b_3 \bmod 2$ are automatic.

[22] Other deviating features are that if n is even then $-4na - n^2 p - 64 \equiv 0 \bmod 4$, while if n is odd then $-4na - n^2 p - 64 \equiv 1 \bmod 4$. We did not observe a noticeable effect on the runtime of our method.

Artifact of $p \equiv 3 \bmod 4$: N cannot be a power of 2. Ideally, we would like N to be a large power of 2, but this conflicts with $M \mid p-1$ given the R2-D4 assumption $p \equiv 3 \bmod 4$. It should be clear to the reader that this assumption is not crucial: e.g., Example C.5 where $p \equiv 5 \bmod 8$ also produces instances of type D_4. Switching to $p \equiv 1 \bmod 4$ will be more technical (by lack of a uniform description of $B_{p,\infty}$) but should be perfectly doable. We will by-pass this by letting $M = 3^f$, $N = 3^{2f}$ for some exponent f.

Group size. Recall from (16) that the discriminant of $\mathbb{Z}[\gamma]$ equals $2^{16}p^2(a^2 - 16p)$. But $\mathbb{Z}[\gamma]$ is not a maximal order; in fact it has index $b^2 = 16$ in the order $\mathbb{Z}[\gamma, \sqrt{p}]$, by which our PPASs are still naturally oriented; consequently $\mathbb{Z}[\gamma, \sqrt{p}]$ has discriminant $2^8p^2(a^2 - 16p)$. This order is still not maximal, but almost: using that $a = p + 4 - 3^{2f}$ and $p \equiv 1 \bmod 3^f$ for some large f, one verifies

$$a^2 - 16p \equiv 4 \bmod 8, \qquad a^2 - 16p \equiv 9 \bmod 27,$$

and then from [48, Thm. 1 and Tbl. B] it follows that the discriminant of the maximal order is $2^6p^2(a^2 - 16p)/9$, if we assume that $(a^2 - 16p)/36$ is square-free (very likely true for the parameters below). We believe that the orientation by $\mathbb{Z}[\gamma, \sqrt{p}]$ is typically primitive, but recall from Remark 4.8 that it is not crucial to know this for sure, as long as the security is based on the maximal order in view of potential hidden stabilizer attacks, as well as to avoid potential "walking up the volcano" attacks (see Section 7). In any case the relevant discriminant is of order p^2a^2 and, since we expect $h^+(\mathbb{Z}[\sqrt{p}])$ to be very small (see the footnote to Corollary 5.15), in view of the Brauer–Siegel theorem it is fair to estimate the size of the Shimura class group as $p|a|$.

Remark F.1. In particular this is unbounded: by letting $|a|$ grow we can reach arbitrarily large group sizes while working exclusively with varieties defined over $\mathbb{F}_p$; this is impossible in dimension $g = 1$. This being said, as soon as $|a| \gtrapprox p^2$ we necessarily have $N \gtrapprox p^2$ and $p - 1$ can no longer host $M = \sqrt{N}$ as a factor, which creates trouble for representing the orientation ι using meet-in-the-middle. One can circumvent this by using smaller values of M and publishing auxiliary PPASs as part of the orientation data. E.g., one can let $N = M^3$ be a cube and choose parameters such that $M \mid p - 1$. To represent $\iota(\gamma)$, note that our N-isogeny $\iota(\gamma)+\alpha$ necessarily factors as a composition $\kappa_3 \circ \kappa_2 \circ \kappa_1$ of M-isogenies:

$$(A, \lambda) \xrightarrow{\kappa_1} (B, \lambda_B) \xrightarrow{\kappa_2} (C, \lambda_C) \xleftarrow{\hat{\kappa}_3} (A, \lambda).$$

If we provide interpolation data on $A[M]$ then this is enough for recovering the outer isogenies in this diagram. We can then append $\ker \kappa_2 \subset B[M]$ to our description of (A, λ, ι) as a PPAS. When evaluating the action, this whole diagram can be pushed through. This is somewhat reminiscent of the technique used in [46], but a full generalization of the method from [46] is left for future research.

Concrete parameters. We have generated parameters aiming for a Shimura class group size of roughly 2048 bits, which should be compared with CSIDH-4096, aiming for NIST-I security. We choose $|a| \approx N \approx p$ to maximize the room available in $p-1$ for packing small primes ℓ (but this also maximizes the field size, so other choices can be made here). We must in fact take N slightly larger than p to create leeway for the generation of corresponding R2-D4 parameters. Within our range of interest, it turns out that $N = 3^{2f}$ with $f = 333$ admits a more-than-average number of small primes ℓ modulo which $9-N$ and $1-N$ are non-zero squares. Concretely we can use the following 56 primes ℓ_i

$$\begin{aligned}&23, 43, 53, 61, 67, 97, 101, 103, 173, 197, 211, 227, 233, 257, 277, 283,\\ &\quad 293, 307, 313, 359, 373, 383, 397, 419, 433, 443, 457, 491, 521, 563, 577,\\ &\quad 601, 613, 673, 683, 691, 733, 743, 769, 787, 809, 811, 859, 883, 887, 911,\\ &\qquad 977, 983, 1069, 1151, 1163, 1187, 1193, 1237, 1249, 1259,\end{aligned}$$

let $p = 2 \cdot 3^{f+2} \cdot 7 \cdot \ell_1 \cdots \ell_{56} + 1$, and then set $a = p + 4 - N$ as required (observe the extra factor 3^2 which is convenient for gluing in the middle). The respective bit lengths of p and $|a|, N$ are 1021 and 1056, making for a Shimura class group estimate of 2077 bits. Using trial division we checked that $(a^2 - 16p)/36$ contains no squares of primes smaller than 2^{40}, so this discriminant factor is most likely square-free, but we could not check this rigorously.[23] Concrete R2-D4 parameters can be found with $a_2 = 1758$, $a_3 = 0$, $b_3 = 565$ so that $n = a_2^2 + b_3^2 = 3409789$, which is prime. The reader can then verify that also $-4na - n^2p - 64$ is prime and representable by the form $X^2 + 4nY^2$, eventually yielding

$$\begin{aligned}a_1 = \ & 72447886291797740848322932962051593473734748284190344\\ & 65405345610109441454835644829643664359635439561189660\\ & 15384207541540878628641645549368269105214510733042954\end{aligned}$$

$$\begin{aligned}b_1 = -& 23283876993666509430775004052081427936666742195999741\\ & 02931752144318449614324311336034511014331071303795311\\ & 71042137236046983177009402579859540412085437180984222\end{aligned}$$

$$\begin{aligned}b_2 = \ & 23122390115850551190499797644857286144606472585743281\\ & 92024567734772318935855806010161778248613700364778186\\ & 27408228857519388198365075195925178584136167770036 7\end{aligned}$$

for the remaining R2-D4 parameters. These integers are roughly of size 3^f.

Isogeny computations involved. Recall the count from (4): to reach a key space size having roughly 256 bits, we want

$$(2B+1)^{2 \cdot 56} \approx 2^{256},$$

[23] One way to get around this is by randomizing the cofactor c, used to construct p, until $a^2 - 16p$ is easy to factor (i.e., a prime number times a smooth cofactor).

showing that $B = 2$ suffices. This means that each arrow in Figure 1 amounts to at most $2 \cdot 56 \cdot 2 = 224$ actions by Shimura ideals of the form $(\mathfrak{a}, \ell_i)$. Each such action involves the computation of 2 isogenies of reduced degree 3^f that "meet in the middle", plus some linear algebra to determine the kernel, followed by an isogeny of reduced degree ℓ_i. All these isogenies are between PPAVs of dimension 2.

F.2 Key exchange from R2-D2

We mention once again that the vectorization problem from the R2-D2 orientation, as it was introduced in Example 3.4, can be reduced to vectorization problems for the ideal-class groups of the underlying imaginary quadratic rings $\mathbb{Z}[\sqrt{-p}]$ and $\mathbb{Z}[\sqrt{-dp}]$ and therefore does not reach the security level one would hope [50]. We nevertheless include a quick, high-level discussion. On one hand, we do this for the sake of context for [50]. On the other hand, it allows us to highlight an interesting connection between R2-D2 orientations and the "higher-degree" supersingular group actions studied by Chenu–Smith [22], used in Orient Express [16]. These concern actions by $\mathrm{Cl}(\mathbb{Z}[\sqrt{-dp}])$ on the set of $\mathbb{Z}[\sqrt{-dp}]$-oriented elliptic curves (E, ι).

Generators. Following Section 4.1 the generating set $\mathbf{g}$ is built from the group homomorphism $\kappa_R : \mathrm{Cl}(\mathbb{Z}[\sqrt{-dp}]) \times \mathrm{Cl}(\mathbb{Z}[\sqrt{-p}]) \to \mathfrak{C}(R)$. That is, we pick a set of "easy" ideal classes

$$[\mathfrak{m}_1], [\mathfrak{m}_2], \ldots, [\mathfrak{m}_s] \in \mathrm{Cl}(\mathbb{Z}[\sqrt{-dp}]), \qquad [\mathfrak{l}_1], [\mathfrak{l}_2], \ldots, [\mathfrak{l}_t] \in \mathrm{Cl}(\mathbb{Z}[\sqrt{-p}])$$

and let $\mathbf{g}$ be the collection of images of $([\mathfrak{m}_i], 1)$, $i = 1, \ldots, s$ and $(1, [\mathfrak{l}_i])$, $i = 1, \ldots, t$. Abusing language, we will say that we act with $[\mathfrak{m}_i]$ when in fact acting with $\kappa_R([\mathfrak{m}_i], 1)$, and likewise for $[\mathfrak{l}_i]$. The ideals $\mathfrak{l}_i$ are chosen as in CSIDH. That is, we set things up such that $\ell_1 \ell_2 \cdots \ell_t \mid p + 1$ for distinct small odd prime numbers ℓ_i and we let $\mathfrak{l}_i = (\ell_i, \sqrt{-p} - 1)$. Then $\bar{\mathfrak{l}}_i = (\ell_i, \sqrt{-p} + 1)$ is a representative of $[\mathfrak{l}_i]^{-1}$. For the ideals $\mathfrak{m}_i$, we consider distinct small odd prime numbers m_i such that $m_1 m_2 \cdots m_s \mid p + 1$ and such that for all $i = 1, 2, \ldots, s$ there exists $a_i \in \mathbb{Z}$ such that $d \equiv a_i^2 \bmod m_i$. We then let $\mathfrak{m}_i = (m_i, \sqrt{-dp} - a_i)$ so that $[\mathfrak{m}_i]^{-1}$ is represented by $\bar{\mathfrak{m}}_i = (m_i, \sqrt{-dp} + a_i)$. The sets $\{\ell_1, \ldots, \ell_t\}$ and $\{m_1, \ldots, m_s\}$ need not be disjoint.

Remark F.2. Recall from Lemma 4.3 that κ_R is "almost" an isomorphism. See Example E.6 for a specialization of this lemma to R2-D2.

Parameters. The data needed for setting up R2-D2 are in accordance with those needed for the "square-free" case of Orient Express [16]. There too, one looks for a supersingular elliptic curve $E/\mathbb{F}_{p^2}$ on which $\pi_p^2 = -p$, admitting a degree-d isogeny

$$\tau : E \to E^{(p)}$$

such that $\hat{\tau} \circ \pi_p = \pi_p \circ \tau$, and such that many small odd primes ℓ divide $p + 1$ and split in $\mathbb{Z}[\sqrt{-dp}]$. Therefore, we can recycle the parameters from [16, §5].

Acting on the starting surface. Interestingly, the higher-degree supersingular group action can be "embedded" in a Shimura class group action. Indeed, given an elliptic curve (E, ι) that is oriented by $\mathbb{Z}[\sqrt{-dp}]$, it is not hard to show that $\iota(\sqrt{-dp})$ factors as

$$E \xrightarrow{\tau} E^{(p)} \xrightarrow{\pi_p} E$$

for a degree-d isogeny τ satisfying $\hat{\tau} \circ \pi_p = \pi_p \circ \tau$ [16, Lem. 1], so we can build an R-oriented PPAS from this exactly as in R2-D2. In view of the Weil restriction involved, let us denote this surface as $W(E, \iota) \in \mathcal{M}_{\overline{\mathbb{F}}_p}(R)$. It is then easy to verify the rule

$$\kappa_R([\mathfrak{m}], 1) \star W(E, \iota) = W([\mathfrak{m}] \star (E, \iota)),$$

whence the said embedding. This means that during public-key generation, acting with the $[\mathfrak{m}_i]$'s can be done at the cost of an Orient Express action, as claimed. In particular, this evaluation only uses isogenies between elliptic curves over $\mathbb{F}_{p^2}$ (although Orient Express uses 2-isogenies between abelian surfaces for interpolating τ). Subsequently acting with the $[\mathfrak{l}_i]$'s is similar to CSIDH and also very efficient, since the orientation representation comes for free.

Acting on a public key. But acting with the $[\mathfrak{m}_i]$'s on a public key, which is not necessarily a Weil restriction, requires us to switch to the machinery from Section 4.3. The parameters in Orient Express are such that $2^{2a} - d = \alpha^2$ for certain integers a, α. This means that the d-isogeny $\iota(\sqrt{d})$ can be represented in dimension $2ug = 2 \cdot 1 \cdot 2 = 4$ using Theorem 4.4. Unfortunately, here the dimension-halving trick from Section 4.3 does not seem to work. The asymmetry between both types of actions is summarized in Figure 2.

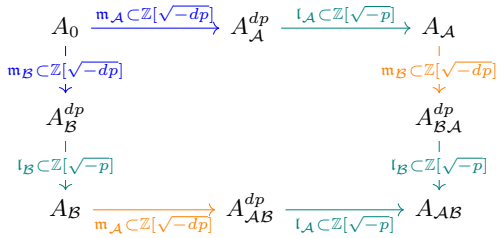

Fig. 2. An illustration of the R2-D2 key exchange. Blue arrows represent actions computed via 1-dimensional isogenies, with a 2D representation of the orientation; green arrows represent actions computed via 2-dimensional isogenies; orange arrows represent actions computed via 2-dimensional isogenies, with a 4D representation of the orientation. Polarizations and orientations are omitted to lighten the notation.

G Compressing interpolation data for R2

In this appendix we discuss the problem of compressing the interpolation data $P, \iota(\gamma)(P)$, with P iterating over a basis of $A[\sqrt{N}]$, in the context of the R2 orientations. To ease notation, let us write $M = \sqrt{N}$. We also assume $M \mid p \pm 1$ where the sign depends on whether we are working in the R2-D2 setting (+) or the R2-D4, R2-C4 settings (−). In view of Remark C.1 this ensures that all interpolation points are defined over $\mathbb{F}_{p^2}$.

Since $\iota(\gamma)$ is defined over $\mathbb{F}_p$ (indeed, it commutes with Π_p), it respects eigenvectors of Frobenius. In particular it maps the 1-eigenspace of Frobenius

$$A[M](\mathbb{F}_p) \cong \left(\frac{\mathbb{Z}}{M\mathbb{Z}}\right)^2$$

into itself, and likewise for the -1-eigenspace (which has the same group structure). To use this for compression purposes, let us first recall that a "symplectic basis" of $A[M]$ is a basis P_1, P_2, Q_1, Q_2 such that $e_M(P_i, Q_j)$ equals ζ_M if $i = j$ and 1 if $i \neq j$, where $\zeta_M \in \mathbb{F}_{p^2}$ denotes some fixed primitive Mth root of unity, whereas $e_M(P_i, P_j) = e_M(Q_i, Q_j) = 1$ for all i, j. Assuming M is odd (the even case is slightly more technical; note that M is odd in our protocol from Appendix F.1), we make the following case distinction:

- If $M \mid p + 1$ as in R2-D2 then it is possible to choose a symplectic basis of $A[M]$ such that $\langle P_1, P_2\rangle$ is the 1-eigenspace and $\langle Q_1, Q_2\rangle$ is the -1-eigenspace.
- If $M \mid p-1$ as in R2-C4 and R2-D4 then one can choose it such that $\langle P_1, Q_1\rangle$ is the 1-eigenspace and $\langle P_2, Q_2\rangle$ is the -1-eigenspace.

The second case where $\Pi^2 = p$ follows from [28, Lem. 3.17 and Lem. 5.1]. The first case where $\Pi^2 = -p$ is easier and well-known; e.g., it follows along the same lines as [28, Lem. 3.17]. Henceforth, we assume that P_1, P_2, Q_1, Q_2 is a deterministically generated basis of the above kind.

It now suffices to describe $\iota(\gamma)(P_i)$, $i = 1, 2$ by means of 4 coefficients $s, t, u, v \in \mathbb{Z}/M\mathbb{Z}$. E.g., in the case of R2-D2, these indicate that

$$\iota(\gamma)(P_1) = sP_1 + tP_2, \qquad \iota(\gamma)(P_2) = uP_1 + vP_2.$$

To expand the representation: we know that

$$\iota(\gamma)(Q_1) = xQ_1 + yQ_2, \qquad \iota(\gamma)(Q_2) = zQ_1 + wQ_2$$

for $x, y, z, w \in \mathbb{Z}/M\mathbb{Z}$ that can be determined through Weil pairing and discrete logarithm computations (in practice M is smooth so this can be done efficiently). Concretely we find

$$sx + ty = \log_{\zeta_M} e_M(\iota(\gamma)(P_1), \iota(\gamma)(Q_1)) = \log_{\zeta_M} e_M(P_1, \iota(d)(Q_1)),$$

where we have used that $\iota(\gamma)^\dagger \iota(\gamma) = \iota(\bar{\gamma}\gamma) = \iota(d)$ using the notation from Section 4.3, and similarly obtain $ux + vy$ from which we can extract x, y. We similarly find z, w and conclude.

H Proofs from Section 5

H.1 The Shimura class group action via matrix arithmetic

Let $g, a, b_1, b_2 \in M_2(\mathcal{O})$ be algebraic data associated with an oriented supersingular PPAS, as in Section 5. Let $(\mathfrak{a}, \mu)$ be a Shimura ideal of $R = \mathbb{Z}[a, b_1, b_2]$. For γ a generator of the left $M_2(\mathcal{O})$-ideal generated by $\mathfrak{a}$, we consider

$$g' := \gamma^{-*} g \mu \gamma^{-1}, \quad a' := \gamma a \gamma^{-1}, \quad b_i' := \gamma b_i \gamma^{-1}, i = 1, 2$$

as in the proof of Theorem 5.4. We now give a direct argumentation, essentially using matrix arithmetic only, for why this is a well-defined group action.

Why g' is a valid g-matrix. In Remark 5.2 we already explained why g' is positive-definite and self-conjugate of norm 1. So in order to prove that it concerns a valid g-matrix, we need to show that $g' \in M_2(\mathcal{O})$. This claim is equivalent to $g^{-1}g' \in M_2(\mathcal{O})$ as g is an invertible element in $M_2(\mathcal{O})$. This can be rewritten as $\gamma^{-\dagger}\mu \in M_2(\mathcal{O})\gamma$.

The ideal $\mathfrak{a}$ can be generated by $r \in R$ and μ with the property that $r^\dagger r = r_0 \mu$ where the norm of r_0 is coprime to the norm of $\mathfrak{a}$. Since γ is the generator of the extended ideal one has that $s\gamma = r$ for some $s \in M_2(\mathcal{O})$ which can be rewritten as $\gamma = s^{-1}r$. Note that $\gamma^{-\dagger}\mu \in M_2(\mathcal{O})$ by the Shimura property of the ideal. So we only need to show that $\gamma^{-\dagger}\mu$ vanishes on the kernel of $\mathfrak{a}$ as then by definition it belongs to the left ideal generated by γ. One has

$$\gamma^{-\dagger}\mu = s^\dagger r^{-\dagger}\mu = s^\dagger r \mu^{-1} r_0^{-1} \mu = s^\dagger r r_0^{-1} = s^\dagger r_0^{-1} r.$$

Here we used the identity $r^\dagger r = r_0 \mu$ as $r^{-\dagger} = r\mu^{-1}r_0^{-1}$. This shows that indeed the above element vanishes on the ideal kernel by the properties of r_0.

Why g', a', b_1', b_2' are valid algebraic data. Since conjugation by any invertible element is an automorphism of $M_2(B_{p,\infty})$, one has that $\gamma a \gamma^{-1}$, $\gamma b_1 \gamma^{-1}$, $\gamma b_2 \gamma^{-1}$ generate a CM order isomorphic to $R = \mathbb{Z}[a, b_1, b_2]$ but we also have to show that Rosati acts as conjugation on them and we also have to show that both lie in $M_2(\mathcal{O})$. We will do the calculation for a but the calculation for the b_i's is identical.

We need that $(\gamma^{-*} g \mu \gamma^{-1})^{-1} (\gamma a \gamma^{-1})^* (\gamma^{-*} g \mu \gamma^{-1}) = \gamma a \gamma^{-1}$. The left-hand side can be rewritten as

$$(\gamma \mu^{-1} g^{-1} \gamma^*)(\gamma^{-*} a^* \gamma^*)(\gamma^{-*} g \mu \gamma^{-1}) = \gamma \mu^{-1}(g^{-1} a^* g) \mu \gamma^{-1}.$$

Finally, $\gamma \mu^{-1}(g^{-1}a^*g)\mu\gamma^{-1} = \gamma a \gamma^{-1}$ as a was symmetric with respect to Rosati and a and μ commute.

What remains is that $\gamma a \gamma^{-1}$ and $\gamma b_i \gamma^{-1}$ are indeed in $M_2(\mathcal{O})$. We show it again only for a as the other cases are very similar. The condition $\gamma a \gamma^{-1} \in M_2(\mathcal{O})$ is equivalent to $\gamma a \in M_2(\mathcal{O})\gamma$ which is the left ideal generated by γ. The

left ideal generated by γ consists of exactly those endomorphisms that vanish on the kernel of $\mathfrak{a}$, so we need to show that γa vanishes on the kernel of $\mathfrak{a}$. Now $\gamma = s^{-1}r$, so $\gamma a = s^{-1}ra = s^{-1}ar$. Now since the reduced norm of s is coprime to the norm of $\mathfrak{a}$ and r is in $\mathfrak{a}$, γa will indeed vanish on the kernel of $\mathfrak{a}$ (the proof carries over to b, as we have only used the fact that a and r commute).

H.2 Omitted proofs

Proof (of Lemma 5.3). Writing

$$g = \begin{pmatrix} s & r \\ \bar{r} & t \end{pmatrix}$$

with $s, t \in \mathbb{Q}_{>0}$ and $r \in B_{p,\infty}$, we can decompose this matrix as

$$g = \gamma \begin{pmatrix} s & 0 \\ 0 & t' \end{pmatrix} \gamma^*, \qquad \gamma = \begin{pmatrix} 1 & 0 \\ s^{-1}\bar{r} & 1 \end{pmatrix},$$

with $t' = t - s^{-1}\,\mathrm{n}(r)$. We have $t' > 0$ by positive-definiteness. By the Hasse–Schilling theorem [94, Thm. 14.7.4] we can write $s = \sigma\bar{\sigma}$ and $t' = \tau\bar{\tau}$ for certain quaternions σ, τ. It follows that

$$h = \gamma \begin{pmatrix} \sigma & 0 \\ 0 & \tau \end{pmatrix}$$

is of the requested kind. □

Proof (of Lemma 5.8). First we compute what the centralizer of a is. Since the centralizer is a purely algebraic object, it is enough to prove it for matrices of the form

$$\begin{pmatrix} 0 & s \\ 1 & 0 \end{pmatrix} \tag{17}$$

where $s \in \mathbb{Z}_{>0}$. Indeed, quadratic fields can be generated by trace-0 elements and all pairwise isomorphic quadratic subfields of $M_2(B_{p,\infty})$ are conjugate to each other by the Noether-Skolem theorem. The centralizer of (17) will be the set of matrices of the form

$$\begin{pmatrix} x & sy \\ y & x \end{pmatrix}$$

where $x, y \in B_{p,\infty}$. This naturally decomposes as the tensor product of diagonal matrices (isomorphic to $B_{p,\infty}$) and $\mathbb{Q}(a)$.

Then we observe that if $b \in M_2(B_{p,\infty})$ commutes with a, then so does $b^\dagger$ (its Rosati involution). Indeed, if $ab = ba$, then $b^\dagger a^\dagger = a^\dagger b^\dagger$ and then $b^\dagger a = ab^\dagger$ as $a^\dagger = a$. This shows that Rosati induces an involution on $C = \mathbb{Q}(a) \otimes B_{p,\infty}$ but we also need a stronger statement. Observe that for any element $c \in C$, the element $c^\dagger c$ is symmetric with respect to Rosati. Due to the positivity of the Rosati involution, $c^\dagger c$ will generate a totally real quadratic subfield of C. However, since C is isomorphic to $\mathbb{Q}(a) \otimes B_{p,\infty}$, the only such subfield is the

center. Thus Rosati defines a standard involution on C (i.e., for every $c \in C$, the element $c^\dagger c$ lies in the center [94, Rmk. 3.2.5]) and for quaternion algebras there exists a unique such involution which is quaternion conjugation. □

Proof (of Lemma 5.10). The decision version is folklore (e.g., it follows from [52, §6]), hence we focus on the embedding problem. First, we compute an explicit isomorphism between $C \otimes_K K(\sqrt{d})$ and $M_2(K(\sqrt{d}))$ in quantum polynomial time [60]. Using this isomorphism we can find an element $s \in C \otimes_K K(\sqrt{d})$ for which $s^2 = 0$. Now if we write

$$s = (s_1 + r_1\sqrt{d}) + (s_2 + r_2\sqrt{d})i + (s_3 + r_3\sqrt{d})j + (s_4 + r_4\sqrt{d})k$$

then we get that $s_1 = r_1 = 0$ as its trace is 0 and if we square it and look at the rational part we get

$$\mathrm{n}(s_2 i + s_3 j + s_4 k) = -d\,\mathrm{n}(r_2 i + r_3 j + r_4 k)$$

which implies that the norm of $(s_2 i + s_3 j + s_4 k)/(r_2 i + r_3 j + r_4 k)$ equals $-d$. Now we only have to show that the trace of this element is 0. The trace of this element is

$$-2\frac{s_2 r_2 i^2 + s_3 r_3 j^2 + s_4 r_4 k^2}{\mathrm{n}(r_2 i + r_3 j + r_4 k)}$$

which is 0 as the numerator equals the coefficient of $\sqrt{d}$ in s^2. This proves our lemma. □

Proof (of Lemma 5.14). We first note that ξ, γ indeed exist, respectively because of Lemma 5.3 and the fact that $M_2(\mathcal{O})$ is a principal ideal ring. Next we claim that the conductor ideal can be characterized as

$$\mathfrak{c} = \{\, \alpha \in K \mid \alpha\gamma^* \in \gamma^{-1} M_2(\mathcal{O}) \}.$$

To see this: it is clear that $\gamma^{-1}M_2(\mathcal{O})$ can be characterized as the set of matrices $x \in M_2(B_{p,\infty})$ such that $\mathbb{Z}_K x \subset M_2(\mathcal{O})$. Let $\alpha \in K$. Using that $\gamma^* M_2(\mathcal{O}) = (M_2(\mathcal{O})\gamma)^* = (M_2(\mathcal{O})\mathbb{Z}_K)^* = \mathbb{Z}_K M_2(\mathcal{O})$ along with the commutativity of K we see that $\mathbb{Z}_K \alpha\gamma^* M_2(\mathcal{O}) = \alpha\mathbb{Z}_K^2 M_2(\mathcal{O}) = \alpha\mathbb{Z}_K M_2(\mathcal{O})$, whence the claim.

As a consequence we have $\kappa := \gamma\mu\gamma^* \in M_2(\mathcal{O})$. Below we show that in fact $\kappa \in \mathrm{GL}_2(\mathcal{O})$. Assuming this, the lemma follows fairly easily. Indeed, let $\Lambda' = \gamma^{-1}M_2(\mathcal{O})\gamma$ be the right order of $M_2(\mathcal{O})\gamma = M_2(\mathcal{O})\mathbb{Z}_K$ and observe using $\mathbb{Z}_K^2 = \mathbb{Z}_K$ that this right order contains $\mathbb{Z}_K$. Then its conjugate Λ contains $\xi^{-1}\mathbb{Z}_K\xi$ and the computation

$$\Lambda^* = \xi^*\gamma^* M_2(\mathcal{O})\gamma^{-*}\xi^{-*} = \xi^{-1}\gamma^{-1}\kappa M_2(\mathcal{O})\kappa^{-1}\gamma\xi = \xi^{-1}\gamma^{-1}M_2(\mathcal{O})\gamma\xi = \Lambda$$

shows that it is closed under conjugate-transpose. The fact that $\iota(\bar{\alpha}) = \iota(\alpha)^*$ for all $\alpha \in \mathbb{Z}_K$ follows from $(\xi^{-1}\alpha\xi)^* = \xi^*\alpha^*\xi^{-*} = \xi^{-1}\alpha^*\xi$ for all $\alpha \in K$, where in turn the last equality follows from the fact that $\xi\xi^* = \mu$ commutes with α.

Proving that $\kappa \in \mathrm{GL}_2(\mathcal{O})$ amounts to showing $\mathrm{n}(\gamma\mu\gamma^*) = 1$. To establish this, it suffices to produce for every prime number $\ell \mid \mathrm{n}(\mu)$ some element $\mu_\ell \in \mathfrak{c}$

such that $\ell \nmid \mathrm{n}(\gamma\mu_\ell\gamma^*)$. It is indeed allowed to skip the primes $\ell \nmid \mathrm{n}(\mu)$ (which include p by assumption) because in those cases $\mathbb{Z}_K \cap M_2(\mathcal{O})$ is locally maximal at ℓ, implying that also $\mathrm{n}(\gamma)$ must be an ℓ-adic unit, and therefore $\ell \nmid \mathrm{n}(\gamma\mu\gamma^*)$.

Instead of finding μ_ℓ directly, we will produce an element $\nu_\ell \in K_\ell$ such that $\gamma\nu_\ell\gamma^* \in \mathrm{GL}_2(\mathcal{O}_\ell)$, where $K_\ell = K \otimes_{\mathbb{Z}} \mathbb{Z}_\ell, \mathcal{O}_\ell = \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$. This is good enough indeed, because if we then pick $\nu'_\ell \in K$ with the same $\mathfrak{l}$-adic valuation as ν_ℓ for every non-zero prime ideal $\mathfrak{l} \mid \ell\mathbb{Z}_K$ (which is always possible by the Chinese remainder theorem), then also $\gamma\nu'_\ell\gamma^* \in \mathrm{GL}_2(\mathcal{O}_\ell)$. It then suffices to let $\mu_\ell = n\nu'_\ell$ for a suitable integer n; e.g., we can consider any positive integer m such that $m\nu'_\ell$ is an integral multiple of μ and choose n such that $n \equiv 1 \bmod \ell$ and $n \equiv m$ modulo all prime divisors of m different from ℓ.

Recall that $\mathbb{Z}_K$ is contained in the right order of $M_2(\mathcal{O})\gamma$, which implies that $\gamma^{-1}M_2(\mathcal{O})$ is closed under left-multiplication with elements from $\mathbb{Z}_K$. It is then also true that $\gamma^{-1}M_2(\mathcal{O}_\ell)$ is closed under left-multiplication by $\mathbb{Z}_{K,\ell} = \mathbb{Z}_K \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$. Fix an isomorphism $B_{p,\infty} \otimes_{\mathbb{Z}} \mathbb{Z}_\ell \cong M_2(\mathbb{Q}_\ell)$, which through component-wise application also yields an isomorphism $M_2(B_{p,\infty} \otimes_{\mathbb{Z}} \mathbb{Z}_\ell) \cong M_4(\mathbb{Q}_\ell)$. We can then consider the $\mathbb{Z}_\ell$-lattice

$$L = \gamma^{-1}L_0, \qquad L_0 = \mathbb{Z}_\ell^4 \tag{18}$$

by viewing the elements of L_0 as column vectors. Observe that this is again closed under left-multiplication by $\mathbb{Z}_{K,\ell}$. This gives L the structure of a $\mathbb{Z}_{K,\ell}$-module. A crucial observation is that this module is free of rank 1. To see why, consider the module

$$\gamma^{-1}\mathbb{Q}_\ell^4 = \bigoplus_{\mathfrak{l} \mid \ell\mathbb{Z}_K} M_\mathfrak{l} \qquad \text{over} \qquad K_\ell \cong \prod_{\substack{\text{non-zero prime} \\ \text{ideals } \mathfrak{l} \mid \ell\mathbb{Z}_K}} K_\mathfrak{l}$$

where $K_\mathfrak{l}$ denotes the completion of K at $\mathfrak{l}$ and the decomposition on the left is obtained by acting on the whole module with subsets of K_ℓ of the form $\{0\} \cdots \times K_\mathfrak{l} \times \cdots \times \{0\}$. Each $M_\mathfrak{l}$ is a vector space over $K_\mathfrak{l}$ and, since both K_ℓ and $\gamma^{-1}\mathbb{Q}_\ell^4$ have dimension 4 as vector spaces over $\mathbb{Q}_\ell$, we have $\dim_{K_\mathfrak{l}} M_\mathfrak{l} = 1$ for all $\mathfrak{l}$, i.e., $\gamma^{-1}\mathbb{Q}_\ell^4$ is a free K_ℓ-module of rank 1. Switching to the "integral side", namely the module

$$L = \gamma^{-1}\mathbb{Z}_\ell^4 = \bigoplus_{\mathfrak{l} \mid \ell\mathbb{Z}_K} L_\mathfrak{l} \qquad \text{over} \qquad \mathbb{Z}_{K,\ell} \cong \prod_{\mathfrak{l} \mid \ell\mathbb{Z}_K} \mathbb{Z}_{K,\mathfrak{l}},$$

and using that the $\mathbb{Z}_{K,\mathfrak{l}}$'s are PIDs, we can deduce the same desired statement: L is a free rank-1 module over $\mathbb{Z}_{K,\ell}$. In other words

$$L = \mathbb{Z}_{K,\ell}v \tag{19}$$

for some $v \in \mathbb{Q}_\ell^4$.

We now consider a $\mathbb{Q}_\ell$-bilinear form $\langle\cdot,\cdot\rangle$ on $\mathbb{Q}_\ell^4$ such that $\langle\mu v, w\rangle = \langle v, \mu^* w\rangle$ for all $v, w \in \mathbb{Q}_\ell^4$ and all $\mu \in M_2(B_{p,\infty})$. While such a bilinear form exists for general reasons [55, §1], let us make this explicit by noting that

$$\bar{a} = H^{-1}a^T H, \qquad H = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

independently of the chosen isomorphism: this merely uses that $\bar{a} = \mathrm{tr}(a) - a$. We can then let $\langle v, w \rangle = v^T \,\mathrm{diag}(H, H) w$. Then using $\langle \gamma^* v, \gamma^{-1} w \rangle = \langle v, \gamma\gamma^{-1} w \rangle = \langle v, w \rangle$ for all $v, w \in \mathbb{Q}_\ell^4$ we find that the dual lattice of L with respect to this form is given by

$$L^* = \{\, v \in \mathbb{Q}_\ell^4 \mid \langle v, L \rangle \subset \mathbb{Z}_\ell \,\} = \gamma^* L_0. \tag{20}$$

But on the other hand, through (19) the form also induces a bilinear map

$$K_\ell \times K_\ell \to \mathbb{Q}_\ell : (r, s) \mapsto \langle rv, sv \rangle = \langle v, \bar{r}sv \rangle.$$

Note that the right-hand side is a $\mathbb{Q}_\ell$-linear expression in $\bar{r}s$. It is a standard property of the trace function that there must exist some $\delta \in K_\ell^\times$ such that

$$\langle v, \bar{r}sv \rangle = \mathrm{Tr}_{K/\mathbb{Q}}(\delta \bar{r} s),$$

where we have extended the domain of $\mathrm{Tr}_{K/\mathbb{Q}}$ to K_ℓ by $\mathbb{Z}_\ell$-linearity. Indeed, this only requires that the trace pairing remains non-degenerate on K_ℓ, which is immediate. Thus from this perspective we can also write

$$L^* = \delta^{-1} \mathfrak{D}^{-1} v \tag{21}$$

with $\mathfrak{D}$ the ideal of $\mathbb{Z}_{K,\ell}$ generated by the different ideal of K. Now $\mathbb{Z}_{K,\ell}$ is a principal ideal ring, again because $\mathbb{Z}_{K,\ell}$ is a product of PIDs, so we can rewrite this as $L^* = \mathbb{Z}_{K,\ell} \nu_\ell^{-1} v$ for some generator ν_ℓ of $\delta\mathfrak{D}$. As the notation suggests, this is the element we are after. Indeed, comparing expressions (18), (19), (20) and (21) we see that $\gamma^{-1} L_0 = L = \nu_\ell L^* = \nu_\ell \gamma^* L_0$ from which it follows that $\gamma \nu_\ell \gamma^* \in \mathrm{GL}_2(\mathcal{O}_\ell)$, as wanted. □

I On the one endomorphism problem in dimension 2

The machinery from Section 5 allows us to prove 2-dimensional analogues of the first two statements in Section 5.4, which we recall here:

- Given one non-scalar endomorphism on a supersingular elliptic curve, one can compute the rest of the endomorphism ring in subexponential time [45].
- Given an endomorphism of small degree (polylogarithmic in p) on a supersingular elliptic curve in characteristic p, one can compute the full endomorphism ring in polynomial time [66].

Our result strongly relies on Corollary 5.15, which for technical reasons was restricted to cases where $R = \mathbb{Z}_K$ is a maximal order; for this reason we make the same restriction here.

Theorem I.1. *Let (A, λ, ι) be a superspecial PPAS oriented by $\mathbb{Z}_K$ for a quartic CM field K and assume that $h^+(K_0)$ is polynomially bounded. There exists a heuristic sub-exponential quantum algorithm that computes the endomorphism ring of A. If the discriminant of K is polynomial in the input size, then the endomorphism ring of A can be computed heuristically in classical polynomial time.*

Proof. As a first step we compute a $\mathbb{Z}_K$-oriented superspecial PPAV (A', λ', ι') along with corresponding algebraic data g, a, b_1, b_2 using Corollary 5.15, which is where the assumption on $h^+(K_0)$ comes into play. Then one can solve vectorization using Kuperberg's algorithm between (A, λ, ι) and (A', λ', ι'). This has no solution if these oriented PPASs lie in different orbits. In that case one can retry with ι' replaced with $\iota' \circ \tau$ for some automorphism τ of K (which by Remark 6.14 amounts to taking the Frobenius conjugate in the case of complex conjugation), but if (A, λ, ι) and (A', λ', ι') come from inequivalent CM types then this brings no solace. In those cases we retry with a new (A', λ', ι') generated by Corollary 5.15: it is a heuristic assumption that this reaches all CM types up to equivalence. Note that this step is never needed for C_4-fields, because here all CM types are equivalent by Lemma 6.8. In any case, after $O(1)$ tries we find $[(\mathfrak{a}, \mu)] \in \mathfrak{C}(\mathbb{Z}_K)$ such that $(A, \lambda, \iota) = [(\mathfrak{a}, \mu)] \star (A', \lambda', \iota')$. Evaluating this action algebraically as described in Section 5.2, we end up with a g-matrix for A, from which its endomorphism ring can be extracted using [73]. If the size of the Shimura class group is polynomial in the input size, then instead of running Kuperberg one can simply compute this group and apply the action of every group element to (A', λ', ι') until (A, λ, ι) is reached. □

Remark I.2. The reader might wonder why we have a specific statement for small discriminant CM orientations. The reason is that if one obtains an oriented superspecial PPAS by using CM reduction, then one immediately knows an orientation by a relatively small CM order. The above theorem then shows that in this case one can compute the full endomorphism ring in polynomial time, generalizing a result of Love and Boneh [66] to dimension 2.

J The Gorenstein and Bass conditions

There are two standard ring-theoretic conditions on a CM order $R \subset \mathbb{Z}_K$ that enter the Shimura class group action: Gorenstein and Bass. Gorenstein is the (mild) hypothesis under which the characteristic-0 Shimura–Taniyama transitivity theorem is expected to hold (see the footnote to Theorem 6.7 in Section 6.2). Bass is the hypothesis under which the ascending-walk picture mentioned in Section 7 is available [3]. Bass implies Gorenstein, but not conversely. This appendix records both notions, two sufficient conditions for Bass, and the index bound on the integer-polarized subgroup $\mathfrak{C}(R)_{\mathbb{Z}} \subseteq \mathfrak{C}(R)$.

Definition J.1. *An order S in a number field is* Gorenstein *if it has Cohen–Macaulay type* 1 *at every maximal ideal $\mathfrak{m} \subset S$, that is, $\dim_{S/\mathfrak{m}}((S : \mathfrak{m})/S) = 1$ for every $\mathfrak{m}$. Equivalently, every fractional S-ideal $\mathfrak{a}$ with $(\mathfrak{a} : \mathfrak{a}) = S$ is invertible.*

The equivalence between the Cohen–Macaulay type condition and the invertibility of every proper fractional ideal is described in Marseglia [67, Proposition 3.4], where the two formulations go back to Bass [4].

If R is Gorenstein, every proper fractional R-ideal is invertible; this is the setting in which $\mathfrak{C}(R)$ acts on proper-ideal classes and in which Theorem 6.7 describes orbits.

If R is not Gorenstein, then there exist proper fractional R-ideals $\mathfrak{a}$ with $(\mathfrak{a} : \mathfrak{a}) = R$ that are not invertible. In this case, the invertible-ideal action of $\mathfrak{C}(R)$ does not cover these additional proper classes. Further, lifting non-transitivity on proper classes to a statement on PPAVs requires in addition that the extra classes admit a principal polarization of the prescribed CM type; the bimodule framework of [3] treats the non-Gorenstein case.

Lemma 2.6 shows that for any CM order R, every square in $\mathfrak{C}(R)$ is integer-polarized, so $[\mathfrak{C}(R) : \mathfrak{C}(R)_{\mathbb{Z}}]$ divides $\#\mathfrak{C}(R)[2]$. The Bass hypothesis gives this bound a graph-theoretic meaning on the crater and underlies the volcano picture used in Section 7.

Definition J.2. *An order R in a number field L is called* Bass *if every over-order of R is Gorenstein. Equivalently, every fractional R-ideal is generated as an R-module by at most two elements.*

Remark J.3. The equivalence between the two formulations is treated in [67, Prop. 4.6]. This covers the classical results of Bass [4] and Drozd–Kirichenko–Roiter [37]. Two elementary sufficient conditions place R in the Bass-class:

- *Index-2 subfield.* If R contains the maximal order of a subfield $K' \subset K$ of index 2, then R is Bass. Maximal real multiplication ($R \supset \mathbb{Z}_{K_0}$) is the main case we use, see [3, §2] and [67, Prop. 4.6].
- *Fourth-power-free discriminant.* If $\mathrm{disc}(R)$ is fourth-power-free in $\mathbb{Z}$, then R is Bass, see [44, Thm. 3.6].

We give an example of a non-Bass CM order.

Example J.4. Let $K = \mathbb{Q}(\zeta_5)$, a quartic CM field with $K_0 = \mathbb{Q}(\sqrt{5})$. Consider the two CM orders of conductor $2\mathbb{Z}_K$

$$R = \mathbb{Z} + 2\mathbb{Z}_K, \;\; R' = \mathbb{Z}_{K_0} + 2\mathbb{Z}_K.$$

Complex conjugation sends ζ_5 to $\zeta_5^{-1} = \zeta_5^4$, so $\mathbb{Z}_K = \mathbb{Z}[\zeta_5]$ is stable under conjugation. The real subfield $K_0 = \mathbb{Q}(\sqrt{5})$ is pointwise fixed. Hence both R and R' are closed under complex conjugation. The prime 2 is inert in $\mathbb{Z}_K$, so $\mathbb{Z}_K/2\mathbb{Z}_K \cong \mathbb{F}_{16}$, $R/2\mathbb{Z}_K \cong \mathbb{F}_2$ and $R'/2\mathbb{Z}_K \cong \mathbb{F}_4$.

Nakayama's lemma (see e.g. [38]) gives

$$\mu_R(\mathbb{Z}_K) = \dim_{\mathbb{F}_2}(\mathbb{Z}_K/2\mathbb{Z}_K) = 4 > 2, \;\; \mu_{R'}(\mathbb{Z}_K) = \dim_{\mathbb{F}_4}(\mathbb{Z}_K/2\mathbb{Z}_K) = 2.$$

By Definition J.2, R is not Bass and R' is Bass. We write

$$2\mathbb{Z}_K = R \cdot 2 + R \cdot 2\zeta_5 + R \cdot 2\zeta_5^2 + R \cdot 2\zeta_5^3$$

and explain why all four generators are necessary to represent $2\mathbb{Z}_K$. Write

$$\mathbb{Z}_K = \langle 1, \zeta_5, \zeta_5^2, \zeta_5^3 \rangle_{\mathbb{Z}}, \text{ and } 2\mathbb{Z}_K = \langle 2, 2\zeta_5, 2\zeta_5^2, 2\zeta_5^3 \rangle_{\mathbb{Z}}.$$

Using $R = \mathbb{Z} + 2\mathbb{Z}_K$, we obtain

$$\begin{aligned} R \cdot 2 + R \cdot 2\zeta_5 + R \cdot 2\zeta_5^2 + R \cdot 2\zeta_5^3 &= \sum_{i=0}^{3} (\mathbb{Z} + 2\mathbb{Z}_K) \cdot 2\zeta_5^i \\ &= 2\sum_{i=0}^{3} \mathbb{Z}\zeta_5^i + 4\sum_{i=0}^{3} \mathbb{Z}_K\zeta_5^i \\ &= 2\mathbb{Z}_K + 4\mathbb{Z}_K \\ &= 2\mathbb{Z}_K. \end{aligned}$$

Here $\sum_{i=0}^{3} \mathbb{Z}\zeta_5^i = \mathbb{Z}_K$, and $\mathbb{Z}_K\zeta_5^i = \mathbb{Z}_K$ because each ζ_5^i is a unit of $\mathbb{Z}_K$. Let $\mathfrak{m} = 2\mathbb{Z}_K$, the maximal ideal of R above 2. Then

$$(2\mathbb{Z}_K)/\mathfrak{m}(2\mathbb{Z}_K) = (2\mathbb{Z}_K)/(4\mathbb{Z}_K) \xrightarrow{\sim} \mathbb{Z}_K/(2\mathbb{Z}_K) \cong \mathbb{F}_{16}, \text{where } 2x \longmapsto x.$$

Since $R/\mathfrak{m} \cong \mathbb{F}_2$, the field $\mathbb{F}_{16}$ has dimension 4 as an $\mathbb{F}_2$-vector space. By Nakayama's lemma, the minimal number of generators of $(2\mathbb{Z}_K)_\mathfrak{m}$ over the local ring $R_\mathfrak{m}$ equals this dimension $\mu_{R_\mathfrak{m}}((2\mathbb{Z}_K)_\mathfrak{m}) = \dim_{\mathbb{F}_2} \frac{2\mathbb{Z}_K}{4\mathbb{Z}_K} = 4$. Therefore the four displayed generators are necessary. Thus the fractional R-ideal $2\mathbb{Z}_K$ is not generated by at most two elements, and Definition J.2 shows that R is not Bass. For R', we have $\mathbb{Z}_{K_0} \subset R'$ and $[K : K_0] = 2$, so the index-2 subfield criterion shows that R' is Bass. The equality $\mu_{R'}(\mathbb{Z}_K) = 2$ follows from $\dim_{\mathbb{F}_4}(\mathbb{F}_{16}) = 2$.

The structure of $\mathfrak{C}(R)$ is given by the relative-norm presentation of equation (2) in Section 2.2. Hasse's norm theorem for the quadratic extension K/K_0 in [74, Cor. VI.4.5] describes the elements of $K_0^\times/N_{K/K_0}(K^\times)$ in terms of local Hasse symbols at the places of K_0 ramifying in K/K_0. This is the input that controls the integer-polarized classes on a elements level. For $R = \mathbb{Z}_K$, Streng [93, Lem. I.8.4, Thm. III.2.2] shows that the cokernel of the type norm map into $\mathfrak{C}(\mathbb{Z}_K)$ has exponent dividing 2; the analogous result in dimension $g = 3$ is [36], where the exponent is 2 or 4 depending on the Galois type.

In the integer-polarized version we expect a sharper statement, for which we have computational evidence in the maximal-order case.

Conjecture J.5. Let K be a CM field and let $R \subseteq \mathbb{Z}_K$ be a CM order. Then

$$[\mathfrak{C}(R) : \mathfrak{C}(R)_\mathbb{Z}] \ \in \ \{1,\ \#\mathfrak{C}(R)[2]\}.$$

Remark J.6. Our computational evidence is for maximal orders across the tested Galois types for $g \in \{2, 3\}$. Computations for non-maximal orders and higher dimensions are in progress and already provide further evidence for our claim.

K Superspecial reduction

Reduction. We recall the superspecial reduction results Goren established for CM abelian surfaces in the primitive quartic case [42]. Goren did not consider the

biquadratic $C_2 \times C_2$ case because his setup only allowed for so-called "primitive" CM types (see Section 7 for a definition). His argument, however, carries over almost verbatim and gives Proposition K.2 below, which we record as a new result even though its proof is analogous to Goren's. As in [42], we restrict throughout this appendix to $R = \mathbb{Z}_K$ and to primes p unramified in K; the ramified case is left for future work (see the end of this appendix).

Proposition K.1. *Let K be a cyclic quartic CM field, let p be a prime number not ramifying in K, and let $(A, \lambda, \iota) \in \mathcal{M}_k(\mathbb{Z}_K)$ with $k = \overline{\mathbb{Q}}$. Let $L \subset k$ be a number field over which A is defined. Let $\bar{\mathfrak{p}} \subset \mathbb{Z}_L$ be a prime ideal above p. Assuming the reduction $A \bmod \bar{\mathfrak{p}}$ is good, it is superspecial if and only if p neither splits completely nor is inert in K.*

Proof. See [42, Thm. 1]. □

In other words, in the C_4 case we have superspecial reduction if and only if $p\mathbb{Z}_K = \mathfrak{p}_1\mathfrak{p}_2$ for two degree-2 prime ideals $\mathfrak{p}_i$ of K.

In the cases $C_2 \times C_2$ and D_4, the factorization pattern of p in K alone does not distinguish superspecial from non-superspecial and the CM type Φ comes into play. It should be clear, however, that the reduction type can only depend on Φ up to equivalence.

Proposition K.2. *We inherit the assumptions from Proposition K.1 but now assume that K is biquadratic. We can assume, by precomposing ι with complex conjugation if needed, that the CM type of (A, λ, ι) is of the form $\Phi = \{\mathrm{id}, \tau\}$ for some $\tau \in \mathrm{Gal}(K, \mathbb{Q})$. Let K_1 be the fixed field of the subgroup Φ. Assuming the reduction $A \bmod \bar{\mathfrak{p}}$ is good, it is superspecial if and only if p is inert in K_1.*

Proof. We can assume that L contains K and is Galois over $\mathbb{Q}$. Ekedahl's product criterion [39, Prop. 2.4], as discussed in [42, §2], says that $A \bmod \bar{\mathfrak{p}}$ is superspecial if and only if

$$\sigma\bar{\Phi}^{\mathrm{ext}} = \Phi^{\mathrm{ext}}$$

where $\sigma \in \mathrm{Gal}(L, \mathbb{Q})$ denotes a Frobenius element of $\bar{\mathfrak{p}}$ (i.e., an automorphism such that $\sigma(\bar{\mathfrak{p}}) = \bar{\mathfrak{p}}$ which reduces to the p-th power Frobenius automorphism of $\mathbb{Z}_L/\bar{\mathfrak{p}}$ over $\mathbb{F}_p$) and Φ^{ext} denotes the set of elements of $\mathrm{Gal}(L, \mathbb{Q})$ whose restriction to K is contained in Φ (and likewise for $\bar{\Phi}^{\mathrm{ext}}$). Now let σ^{res} be the Frobenius element of $\bar{\mathfrak{p}} \cap K$ (unique since p does not ramify). This equals the restriction of σ to K, so Ekedahl's criterion becomes: $A \bmod \bar{\mathfrak{p}}$ is superspecial if and only if $\sigma^{\mathrm{res}} \notin \{\mathrm{id}, \tau\}$. The result follows. □

In other words, in the $C_2 \times C_2$ case we have superspecial reduction if and only if $p\mathbb{Z}_K = \mathfrak{p}_1\mathfrak{p}_2$ for two degree-2 prime ideals $\mathfrak{p}_i$ of K and the splitting is complete in K_0 or in the imaginary quadratic subfield different from K_1. As a sanity check, consider our recurring orientations by

$$\mathbb{Z}\left[\frac{1+\sqrt{-p-4}}{2}, \sqrt{\frac{p+1}{4}}\right]$$

in characteristic $p \equiv 3 \bmod 4$ from Examples 3.1 and 3.2, shown to come from two inequivalent CM types in Examples 6.4 and 6.11. By Proposition K.2 combined with the lifting statement from [43, Lem. 4.4.1] (which was used in our proof of Theorem 6.9), this must mean that p splits in K_0 and is inert in the two imaginary quadratic subfields K_1, K_2. This is indeed easily checked.

Proposition K.3. *Let K be a quartic CM field with Galois closure M such that $\mathrm{Gal}(M, \mathbb{Q}) = \langle x, y \,|\, x^2, y^4, (xy)^2 \rangle \cong D_4$ and that $K = M^{\langle x \rangle}$. Let $\Phi = \{\mathrm{id}, y\}$ be a fixed CM type on K. Under the remaining assumptions of Proposition K.1, where we can assume $M \subset L$ w.l.o.g., and with p unramified in M, the reduction $A \bmod \bar{\mathfrak{p}}$ is superspecial if and only if the Frobenius element of $\bar{\mathfrak{p}} \cap M$ equals y^2 or xy.*

Proof. This is [42, Thm. 2] in cases (3) and (6). The two cases together are again seen to be equivalent to the product criterion of Ekedahl. □

This closes the unramified cases for dimension 2 (assuming $R = \mathbb{Z}_K$). The ramified case, where $p \mid \mathrm{disc}(K/\mathbb{Q})$, is not covered by [42, Thm. 1 and Thm. 2]. In forthcoming work we treat this case by replacing the unramified product criterion with a Lie-type Dieudonné module calculation. The Shimura–Taniyama formula gives the Newton polygon, and the Lie type gives the a-number. Superspeciality is then the condition that all slopes are $1/2$ and that $a = g$.