Hash functions from superspecial genus-2 curves using Richelot isogenies

Wouter Castryck, Thomas Decru, Benjamin Smith

HASH FUNCTIONS FROM SUPERSPECIAL GENUS-2 CURVES USING RICHELOT ISOGENIES

WOUTER CASTRYCK, THOMAS DECRU, AND BENJAMIN SMITH

Abstract. Last year Takashima proposed a version of Charles, Goren and Lauter's hash function using Richelot isogenies, starting from a genus-2 curve that allows for all subsequent arithmetic to be performed over a quadratic finite field $\mathbb{F}_{p^2}$. In a very recent paper Flynn and Ti point out that Takashima's hash function is insecure due to the existence of small isogeny cycles. We revisit the construction and show that it can be repaired by imposing a simple restriction, which moreover clarifies the security analysis. The runtime of the resulting hash function is dominated by the extraction of 3 square roots for every block of 3 bits of the message, as compared to one square root per bit in the elliptic curve case; however in our setting the extractions can be parallelized and are done in a finite field whose bit size is reduced by a factor 3. Along the way we argue that the full supersingular isogeny graph is the wrong context in which to study higher-dimensional analogues of Charles, Goren and Lauter's hash function, and advocate the use of the superspecial subgraph, which is the natural framework in which to view Takashima's $\mathbb{F}_{p^2}$-friendly starting curve.

1. Introduction

After a cautious start with Couveignes' unpublished note [8] from 1997 and Stolbunov's master thesis [26] from 2004, the area of isogeny-based cryptography took a more visible turn in 2006 when Charles, Goren and Lauter [7] showed how to construct collision-resistant hash functions from deterministic walks in isogeny graphs of supersingular elliptic curves over finite fields. About five years later Jao and De Feo applied similar ideas to the design of a key exchange protocol [20, 10] now known as SIDH, after which isogenies became a very active topic of cryptographic research, largely due to their promise of leading to quantum resistant hard problems. Some of the recent constructions include non-interactive key exchange [11, 5], signatures [9, 13] and verifiable delay functions [12]. Last January it was announced that SIKE [1], which is an incarnation of SIDH, is one of the seventeen second-round contenders to become a NIST standard for post-quantum key establishment.[1]

While almost all of the ongoing research in isogeny-based cryptography is devoted to elliptic curves, there is a general awareness that many proposals should generalize to principally polarized abelian varieties (e.g., jacobians) of arbitrary dimension. This particularly applies to the supersingular isogeny walks on which SIDH and Charles, Goren and Lauter's hash function are based. In fact, in a follow-up paper [6, §6.2] the latter authors already hint at the possibility of a higher-dimensional analogue of their hash function. Last year Takashima [27, §4.2] made the concrete proposal of using jacobians of supersingular genus-2 curves and

[1] See https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions.

their 15 outgoing $(2,2)$-isogenies, which can be evaluated efficiently through Richelot's formulas. By disallowing backtracking he uses this to process one base-14 digit for each isogeny evaluation. Moreover he provides specific starting curves, such as $y^2 = x^5 + 1$ over $\mathbb{F}_p$ with $p \equiv 4 \bmod 5$, which allow for all computations to be done over $\mathbb{F}_{p^2}$, as was shown by himself and Yoshida about a decade ago [28]. Unfortunately Takashima's hash function is not collision-resistant due to the inherent presence of small cycles in the resulting isogeny graph, as was pointed out very recently by Flynn and Ti [16], who then proceeded with studying a genus-2 variant of SIDH.

The contributions of this paper are as follows. First, in Section 2 we argue that the full supersingular isogeny graph is the wrong arena for higher-dimensional analogues of Charles, Goren and Lauter's hash function, and promote the use of superspecial subgraphs. In doing so we give a natural explanation for why Takashima and Yoshida's starting curve indeed allows for all subsequent arithmetic to be carried out in $\mathbb{F}_{p^2}$. Second, some first properties of the $(2,2)$-isogeny graph of superspecial principally polarized abelian surfaces are gathered and proved in Section 4 and Appendix A. Third and foremost, we repair Takashima's hash function by showing that an extremely simple restriction (which still allows us to process one base-8 digit, i.e., 3 bits per isogeny) both prevents the Flynn–Ti attack and simplifies the reasoning on security; we also show that with high probability, the starting curve $y^2 = (x^2-1)(x^2-2x)(x-1/2)$ over $\mathbb{F}_p$ with $p \equiv 5 \bmod 6$ naturally avoids running into products of elliptic curves, which as we will see are technical nuisances. The details can be found in Section 6 and Section 7. In Sections 8 and 9 we report on an implementation in Magma and compare its performance with the elliptic curve case of Charles, Goren and Lauter.

Why generalizing? Besides scientific curiosity, we see a number of motivations for investigating higher-dimensional isogeny-based cryptography:

(1) There seem to exist some beneficial trade-offs between the larger computational cost of each isogeny evaluation and features such as larger graph sizes, higher numbers of outgoing isogenies, or arithmetic in smaller finite fields. As an illustration of this, we note that in Charles, Goren and Lauter's hash function one needs to compute one square root for each digested bit, while our proposal uses 3 square roots per 3 bits, which seems like no improvement at all, except that our square roots are to be extracted in finite fields of about one third of the bit size and can be handled in parallel. See Section 9 for some further comments on this.

(2) The fact that higher-dimensional abelian varieties have torsion subgroups of larger rank may allow for a symmetric set-up of SIDH in which Alice and Bob sample their secrets from the same space (but this is not touched upon in the current paper).

2. Supersingular versus superspecial

One apparent point of concern is that in the case of elliptic curves over a finite field of characteristic p, supersingularity has many equivalent characterizations whose natural generalizations to higher dimension become distinct notions. For instance, one such characterization reads that the trace t of Frobenius satisfies

$t \equiv 0 \bmod p$, which naturally generalizes to the notion of *superspeciality*.[2] An alternative characterization states that the Newton polygon is a straight line segment with slope $1/2$; this property makes sense in arbitrary dimension where it is still called *supersingularity*, but in dimension $g \geq 2$ this is a weaker condition than superspeciality. A third characterization is that there exists no non-trivial p-torsion. This also makes sense in arbitrary dimension but in dimension $g \geq 3$ it weakens the notion of supersingularity. A curve is called superspecial or supersingular if its accompanying jacobian is superspecial or supersingular respectively.

We refer to Brock's thesis [2] and the references therein for some general facts on supersingularity and superspeciality. Most notably, it can be shown that an abelian variety is supersingular if and only if it is isogenous to a product of supersingular elliptic curves, while it is superspecial if and only if it is isomorphic to such a product; moreover in dimension $g \geq 2$ all such products are isomorphic to each other, see e.g. [2, Thm. 2.1A] or [23, p. 13]. Here, isogenous and isomorphic should be understood in the context of abstract abelian varieties, regardless of the principal polarization with which they may come equipped: statements like 'all superspecial curves of genus $g \geq 2$ have isomorphic jacobians' are of course not true in general (see Theorem 1 below for a precise count in case $g = 2$). We will abbreviate principal polarization to p.p. from now on and will also assume that a product of elliptic curves always comes with the product polarization, unless stated otherwise.

We believe that the full graph of supersingular p.p. abelian varieties is the wrong context in which to study Charles–Goren–Lauter hash functions in dimension $g \geq 2$. Instead we argue for use of the superspecial subgraph. Indeed, the family of supersingular p.p. abelian varieties over $\overline{\mathbb{F}}_p$ is infinite-dimensional, whereas the superspecial subfamily is 0-dimensional. The latter implies that there is only a finite number of them and, furthermore, they all admit a model over $\mathbb{F}_{p^2}$ whose Frobenius endomorphism has characteristic polynomial $\chi(t) = (t \pm p)^{2g}$, in particular it acts as multiplication by $\pm p$; see [18]. Assuming that p is odd, this implies that all 2-torsion is $\mathbb{F}_{p^2}$-rational, hence so are all $(2, 2, \ldots, 2)$-isogenies and their codomains. By [2, Lem. 2.2A] these are again superspecial p.p. abelian varieties whose Frobenius has the same characteristic polynomial, so the argument repeats and we conclude that the full superspecial $(2, 2, \ldots, 2)$-isogeny graph is defined over $\mathbb{F}_{p^2}$. This explains the aforementioned observation by Takashima and Yoshida, whose starting curves are indeed superspecial. See [19], where several more examples of superspecial genus-2 curves over $\mathbb{F}_p$ can be found. These include $y^2 = x^5 - x$ which is superspecial if and only if $p \equiv 5$ or $7 \bmod 8$, and $y^2 = (x^2 - 1)(x^2 - 2x)(x - 1/2)$ which is superspecial if and only if $p \equiv 5 \bmod 6$. In characteristics 2 and 3 superspecial genus-2 curves do not exist. In general it seems unknown how to write down the equation of a random superspecial genus-2 curve.

Note that superspecial p.p. abelian varieties were also considered in Charles, Goren and Lauter's follow-up paper [6], albeit in a more theoretical context and using different edge and vertex sets for the associated graphs.

[2] In the case of jacobians, superspeciality amounts to the Hasse-Witt matrix $M \in \mathbb{F}_p^{g \times g}$ being zero, where we note that $M \equiv t \bmod p$ when $g = 1$. For arbitrary abelian varieties A being superspecial means that Frobenius acts as the zero map on $H^1(A, \mathcal{O}_A)$.

3. Further preliminaries

3.1. **Hyperelliptic curves of genus-2.** Let K be a field of characteristic $p > 5$. A *(hyperelliptic) curve of genus-2 over K* is an algebraic curve defined by an equation of the form $y^2 = f(x)$, where $f(x) \in K[x]$ is a squarefree polynomial of degree 5 or 6. Up to $\overline{K}$-isormorphism, any genus-2 curve has a representation with a monic polynomial of degree 6 and we will mostly work with these representations since it eases up the notation quite a bit. All formulas provided still work with a degree 5 polynomial if one sees the missing linear factor as '$0 \cdot x + 1$'. A genus-2 curve is determined (up to $\overline{K}$-isomorphism) by its Cardona–Quer invariants. The specific formulas for these invariants are discussed in [4], but for our purposes it suffices to know that they consist of an ordered triple $(j_1, j_2, j_3) \in K^3$.

3.2. **Richelot isogenies.** A *Richelot isogeny* is a $(2,2)$-isogeny between jacobians of genus-2 curves, i.e. the kernel of the isogeny is a group isomorphic to $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ that is maximal isotropic with regards to the 2-Weil pairing. The 2-torsion of the jacobian of the genus-2 curve $C : y^2 = f(x) = \prod_{i=1}^{6}(x - \alpha_i)$ is $\{0\} \cup \{[(\alpha_i, 0) - (\alpha_j, 0)] : i < j\}$, where the square brackets denote linear equivalence classes of divisors. A subgroup of the 2-torsion being maximal isotropic with regards to the 2-Weil pairing in this context simply means that the group contains exactly 3 non-trivial elements such that all α_i, $1 \leq i \leq 6$, occur exactly once in all the representations combined. Hence the Richelot isogenies can be represented by sets of quadratic factors of $f(x)$ that are pairwise coprime. More precisely, if we define

$$\begin{cases} G_1 = g_{1,3}x^2 + g_{1,2}x + g_{1,1} = (x - \alpha_1)(x - \alpha_2), \\ G_2 = g_{2,3}x^2 + g_{2,2}x + g_{2,1} = (x - \alpha_3)(x - \alpha_4), \\ G_3 = g_{3,3}x^2 + g_{3,2}x + g_{3,1} = (x - \alpha_5)(x - \alpha_6), \end{cases}$$

then the $(2,2)$-isogeny with kernel $\{0, [(\alpha_1, 0) - (\alpha_2, 0)], [(\alpha_3, 0) - (\alpha_4, 0)], [(\alpha_5, 0) - (\alpha_6, 0)]\}$ can be identified by the *quadratic splitting* $\{G_1, G_2, G_3\}$. While the above equalities force our quadratic factors to be monic, i.e. $g_{i,3} = 1$ for all i, we incorporate the leading coefficients in our discussion for the sake of generality (e.g., to cope with the degree 5 case where one of the $g_{i,3}$'s becomes zero). In any case quadratic splittings are only identified up to permutation and constant multiples of the three quadratics.

There are 15 possible ways of organizing the roots α_i into distinct quadratic splittings. It is possible that the resulting quadratics are only defined over an extension of the field over which our curve C is defined, in which case both the corresponding $(2,2)$-isogeny and its codomain also might be defined over this field extension. Nevertheless, if the splitting is fixed by Frobenius *as a set*, then the isogeny and codomain are defined over the ground field. As mentioned in Section 2, in the case of superspecial p.p. abelian surfaces, all domains, kernels, $(2,2)$-isogenies and associated codomains are defined over $\mathbb{F}_{p^2}$ up to isomorphism.

Proposition 1. *Let $C : y^2 = G_1(x) \cdot G_2(x) \cdot G_3(x)$ be a genus-2 curve, with $\{G_1, G_2, G_3\}$ the quadratic splitting associated with a maximal 2-Weil-isotropic subgroup $S \subset J_C[2]$, and let $\phi : J_C \to A \cong J_C/S$ be the quotient $(2,2)$-isogeny.*

Following the notation above, let

$$\delta := \det \begin{pmatrix} g_{1,3} & g_{1,2} & g_{1,1} \\ g_{2,3} & g_{2,2} & g_{2,1} \\ g_{3,3} & g_{3,2} & g_{3,1} \end{pmatrix} .$$

(1) *If $\delta \neq 0$, then A is isomorphic to the jacobian of the genus-2 curve*

$$C' : y^2 = \delta^{-1} H_1(x) \cdot H_2(x) \cdot H_3(x)$$

where

$$H_1 := G_2' G_3 - G_2 G_3' , \quad H_2 := G_3' G_1 - G_3 G_1' , \quad H_3 := G_1' G_2 - G_1 G_2' ,$$

where G_i' is the derivative of G_i with respect to x. Moreover, $\{H_1, H_2, H_3\}$ is a quadratic splitting corresponding to the dual isogeny $\hat{\phi} : J_{C'} \to J_C$.

(2) *If $\delta = 0$, then A is isomorphic to a product of elliptic curves $E_1 \times E_2$. The vanishing of the determinant δ implies that there exist s_1 and s_2 in $\mathbb{F}_{p^2}$ such that*

$$G_i = a_{i,1}(x - s_1)^2 + a_{i,2}(x - s_2)^2$$

for some $a_{i,1}$ and $a_{i,2}$ in $\mathbb{F}_{p^2}$ for $i = 1, 2, 3$. The elliptic curves forming the product isomorphic to A can be defined by the equations

$$E_1 : y^2 = \prod_{i=1}^{3} (a_{i,1} x + a_{i,2}) , \qquad E_2 : y^2 = \prod_{i=1}^{3} (a_{i,1} + a_{i,2} x) ,$$

and the isogeny ϕ is induced by $\phi_1 \times \phi_2$, where $\phi_1 : C \to E_1$ is $(x, y) \mapsto ((x-s_1)^2/(x-s_2)^2, y/(x-s_2)^3)$ and $\phi_1 : C \to E_2$ is $(x, y) \mapsto ((x-s_2)^2/(x-s_1)^2, y/(x-s_1)^3)$.

For a proof of this proposition and a more in-depth discussion about Richelot isogenies, see [25, Chapter 8].

3.3. **(2,2)-isogenies from products of elliptic curves.** Consider the p.p. abelian surface $E_1 \times E_2$ given by the equations

$$E_1 : y^2 = \prod_{i=1}^{3} (x - \alpha_i) , \qquad E_2 : y^2 = \prod_{i=1}^{3} (x - \beta_i) .$$

Just as in the case of jacobians of genus-2 curves, there are 15 outgoing $(2, 2)$-isogenies with domain $E_1 \times E_2$. Of these, 9 correspond to an isogeny that is the product of 2-isogenies on the respective elliptic curves, such that the image of this isogeny is again simply a product of elliptic curves. The other 6 determine an isogeny where the kernel is given by

$$\kappa = \{(\mathcal{O}_{E_1}, \mathcal{O}_{E_2}), (P_1, Q_{\sigma(1)}), (P_2, Q_{\sigma(2)}), (P_3, Q_{\sigma(3)})\},$$

with $\mathcal{O}_{E_1}$ and $\mathcal{O}_{E_2}$ the neutral element of E_1, respectively E_2, σ a permutation of $\{1, 2, 3\}$, and $P_i = (\alpha_i, 0)$, $Q_i = (\beta_i, 0)$. As long as κ is not the restriction of the graph of an isomorphism $E_1 \to E_2$, the image of the isogeny determined by κ is the jacobian of a genus-2 curve which can be constructed as follows. Define Δ_α and Δ_β

as the discriminants of the monic cubic polynomials $\prod_{i=1}^{3}(x-\alpha_i)$ and $\prod_{i=1}^{3}(x-\beta_i)$ respectively, and

$$\begin{aligned}
a_1 &= (\alpha_3-\alpha_2)^2/(\beta_3-\beta_2) + (\alpha_2-\alpha_1)^2/(\beta_2-\beta_1) + (\alpha_1-\alpha_3)^2/(\beta_1-\beta_3),\\
b_1 &= (\beta_3-\beta_2)^2/(\alpha_3-\alpha_2) + (\beta_2-\beta_1)^2/(\alpha_2-\alpha_1) + (\beta_1-\beta_3)^2/(\alpha_1-\alpha_3),\\
a_2 &= \alpha_1(\beta_3-\beta_2) + \alpha_2(\beta_1-\beta_3) + \alpha_3(\beta_2-\beta_1),\\
b_2 &= \beta_1(\alpha_3-\alpha_2) + \beta_2(\alpha_1-\alpha_3) + \beta_3(\alpha_2-\alpha_1).
\end{aligned}$$

It can be proved that $\Delta_\alpha, \Delta_\beta, a_1, b_1, a_2, b_2$ are all nonzero, such that $A = \Delta_\beta a_1/a_2$ and $B = \Delta_\alpha b_1/b_2$ are well defined and nonzero as well. With these notations in mind, the image of the $(2,2)$-isogeny with kernel κ is the jacobian of the genus-2 curve given by the equation

$$\begin{aligned}
y^2 = -\big(&A(\alpha_2-\alpha_1)(\alpha_1-\alpha_3)x^2 + B(\beta_2-\beta_1)(\beta_1-\beta_3)\big)\\
&\cdot\big(A(\alpha_3-\alpha_2)(\alpha_2-\alpha_1)x^2 + B(\beta_3-\beta_2)(\beta_2-\beta_1)\big)\\
&\cdot\big(A(\alpha_1-\alpha_3)(\alpha_3-\alpha_2)x^2 + B(\beta_1-\beta_3)(\beta_3-\beta_2)\big).
\end{aligned}$$

The three factors on the right hand side constitute a quadratic splitting for the dual isogeny back to $E_1 \times E_2$; note in particular that these factors are multiples of each other so that the corresponding value of δ is indeed 0.

If $E_1 \cong E_2$ we will have strictly fewer than six $(2,2)$-isogenies from $E_1 \times E_2$ to the jacobian of a genus-2 curve. The exact number in this case is given by the formula $6 - \#\mathrm{Aut}(E_1)/2$. If the j-invariant of E_1 is 0 or 1728 then this expression is 3 respectively 4 (under the assumption that $p > 3$). In all other cases this expression is 5 since the only automorphisms are ± 1.

The final case to consider is when we want to construct an isogeny with domain an abelian surface of the form $E_1 \times E_2$, with $E_1 \cong E_2$, and of which the kernel κ is the restriction of the graph of an isomorphism $\alpha : E_1 \to E_2$. The codomain is then the same as the domain and the $(2,2)$-isogeny is given by

$$\begin{aligned}
\phi : \quad E_1 \times E_2 &\to E_1 \times E_2\\
(P, Q) &\mapsto (P + \hat{\alpha}(Q), -Q + \alpha(P)),
\end{aligned}$$

which is clearly self-dual.

For a proof of these previous statements and a more in-depth discussion, see [17], [21] and [3].

4. The superspecial $(2,2)$-isogeny graph

For each prime p, we define a directed multigraph $\mathcal{G}_p$ as follows.[3] The vertices of $\mathcal{G}_p$ represent the isomorphism classes of superspecial p.p. abelian surfaces defined

[3]Every $(2,2)$-isogeny $\phi : A_1 \to A_2$ has a unique dual $(2,2)$-isogeny $\hat{\phi} : A_2 \to A_1$, so one might think that we could easily treat $\mathcal{G}_p$ as an undirected graph. Unfortunately, this may fail if A_1 has automorphisms different from ± 1. Indeed, in that case it is possible that two non-isomorphic $(2,2)$-isogenies $\phi : A_1 \to A_2$ and $\psi : A_1 \to A_2$ are obtained from each other by pre-composition with such an automorphism, so that their duals are obtained from one another by *post*-composition with this automorphism (more precisely if $\phi = \psi \circ \alpha$ then $\hat{\phi} = \alpha^{-1} \circ \hat{\psi}$). So these duals have the same kernel, hence they are isomorphic. In the elliptic curve case, this technicality can be combated by choosing $p \equiv 1 \bmod 12$, since then the automorphisms of all curves are always ± 1. In the case of superspecial genus-2 curves, however, no such convenient restriction exists: there are jacobians with a different number of automorphisms for any prime p [19].

over $\overline{\mathbb{F}}_p$. The graph $\mathcal{G}_p$ has an edge from vertex A_1 to vertex A_2 for every $(2,2)$-isogeny from the superspecial p.p. abelian surface corresponding to A_1 to the one corresponding to A_2, again up to isomorphism. Here, isomorphisms of outgoing $(2,2)$-isogenies are commutative diagrams

$$\begin{array}{ccc} A_1 & \xrightarrow{\phi} & A_2 \\ & \searrow^{\phi'} & \downarrow{\scriptstyle \iota} \\ & & A_2' \end{array}$$

where ϕ and ϕ' are $(2,2)$-isogenies and ι is an isomorphism of superspecial p.p. abelian surfaces. Since the isomorphism class of an outgoing isogeny is uniquely determined by its kernel, this simply means that we have an outgoing edge for each $(2,2)$-subgroup of A_1, i.e., each subgroup that is isomorphic to $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ and maximal isotropic with regards to the 2-Weil pairing.

By construction, $\mathcal{G}_p$ is a 15-regular (multi)graph, since both types of superspecial p.p. abelian surfaces have 15 different $(2,2)$-isogenies. One might simplify the situation by combining parallel edges to turn $\mathcal{G}_p$ into a simple directed graph, but for our application we will need to distinguish between all 15 outgoing edges. In any case, for large p the number of parallel edges is expected to be negligible relative to the size of the graph (for very small p, where there are few superspecial p.p. abelian surfaces, the opposite holds—as we will see in §5).

The vertices of $\mathcal{G}_p$ fall into two classes:

$$V(\mathcal{G}_p) = \mathcal{E}_p + \mathcal{J}_p\,,$$

where $\mathcal{E}_p$ is the set of isomorphism classes corresponding to products of supersingular elliptic curves, and $\mathcal{J}_p$ is the set of isomorphism classes of superspecial genus-2 jacobians. Theorem 1 gives us the cardinalities of these subsets.

Theorem 1. *Let $\mathcal{G}_p$, $\mathcal{E}_p$, and $\mathcal{J}_p$ be defined as above.*

- *If $p = 2$ or 3, then $\#\mathcal{J}_p = 0$ and $\#\mathcal{E}_p = 1$.*
- *If $p = 5$, then $\#\mathcal{J}_p = 1$ and $\#\mathcal{E}_p = 1$.*
- *If $p > 5$, then*

$$\#\mathcal{J}_p = \frac{p^3 + 24p^2 + 141p - 346}{2880} + \delta_p$$

and

$$\#\mathcal{E}_p = \frac{1}{2}\left(\frac{p-1}{12} + \epsilon_p\right)\left(\frac{p-1}{12} + \epsilon_p + 1\right),$$

where $\delta_p \in [0, \frac{881}{720}]$ depends only on $p \bmod 120$ and $\epsilon_p \in \left[0, \frac{7}{6}\right]$ depends only on $p \bmod 12$.

Proof. The values for $\#\mathcal{J}_p$ appear in [2, Theorem 3.10(b)] or [19, Theorem 3.3]. The formulas for $\#\mathcal{E}_p$ follow from the fact that up to $\overline{\mathbb{F}}_p$-isomorphism, the number of supersingular elliptic curves over $\overline{\mathbb{F}}_p$ is $(p-1)/12 + \epsilon_p$, where $\epsilon_p \in \left[0, \frac{7}{6}\right]$ depends only on $p \bmod 12$ (see for example [24, Section V, Theorem 4.1(c)]). □

Theorem 1 implies that $\mathcal{G}_p$ is a finite graph, although this could already be derived from the fact that every isomorphism class of superspecial p.p. abelian

surfaces has a representative defined over $\mathbb{F}_{p^2}$. Asymptotically, we have

$$\#\mathcal{G}_p = O(p^3)\,, \qquad \#\mathcal{E}_p = O(p^2)\,, \qquad \#\mathcal{J}_p = O(p^3)\,.$$

In particular, the proportion of superspecial p.p. abelian surfaces that are the product of 2 supersingular elliptic curves is $O(1/p)$ relative to the total size of the graph: for p large, the number of vertices in $\mathcal{G}_p$ that are *not* in $\mathcal{J}_p$ is negligible.

Informally, when p is large, one could see $\mathcal{E}_p$ as the 'boundary' of the graph $\mathcal{G}_p$, and $\mathcal{J}_p$ as the 'interior'. A first reason is the size argument we just made. A second reason is the connectivity of the 2 types of superspecial p.p. abelian surfaces that we briefly touched on in the preliminaries. Indeed, every product of elliptic curves has at least 9 out of 15 $(2,2)$-isogenies that have a codomain that is a product of elliptic curves as well, hence this part of our graph is very well connected while only making up a fraction of our graph. Vice versa there is also no jacobian of a genus-2 curve that could be 'hiding' in between the products of elliptic curves, which we can make precise with the following theorem.

Theorem 2. *With the notation above:*

(1) *Suppose $p \neq 5$. If J is a vertex in $\mathcal{J}_p \subset \mathcal{G}_p$, then (counting multiplicity) at most 6 of the 15 edges out of J are to vertices in $\mathcal{E}_p$.*
(2) *If E is a vertex in $\mathcal{E}_p \subset \mathcal{G}_p$, then (counting multiplicity) at most 6 of the 15 edges out of E are to vertices in $\mathcal{J}_p$.*

Proof. The second part of this theorem was mentioned in the preliminaries and it follows from the fact that 9 out of 15 $(2,2)$-isogenies are simply a product of 2-isogenies on the separate elliptic curve factors. A proof of a more general formula can be found in [21]. For a proof of the first part, see Appendix A. □

A simple counting argument then tells us that for sufficiently large p, the chance of a vertex in $\mathcal{J}_p$ having a neighbour in $\mathcal{E}_p$ in our graph $\mathcal{G}_p$ is negligible. Intuitively this makes sense, since the δ in Proposition 1 is the determinant of a seemingly random 3×3 matrix for large p, and will therefore almost surely be nonzero.

We now state a pair of conjectures inspired by analogous theorems for the elliptic supersingular 2-isogeny graph.

Conjecture 1. *The graph $\mathcal{G}_p$ is connected.*

Conjecture 1 is the most natural, but will not necessarily be the one that we will base choices on. We mainly state it due to the analogy with the elliptic curve case.

Conjecture 2. *The subgraph of $\mathcal{G}_p$ supported on $\mathcal{J}_p$ is connected.*

Conjecture 2 (which is identical to Conjecture 1 in the elliptic case) is more relevant to our discussion. It implies that $\mathcal{E}_p$ not only makes no significant contribution to the size of $\mathcal{G}_p$ as $p \to \infty$, but it is also not essential for connectivity. (Thus, we consider $\mathcal{E}_p$ to be the "boundary" of $\mathcal{G}_p$.) Conjecture 2 implies Conjecture 1, since every vertex in $\mathcal{E}_p$ has at least 4 outgoing edges into $\mathcal{J}_p$ for $p > 3$ (as mentioned in the preliminaries). As a final note, one may wonder if all non-superspecial supersingular p.p. abelian surfaces also form a similar connected component (which is necessarily infinite). Given that we will not make use of these abelian surfaces, we will not explore that thought any further.

5. The graph $\mathcal{G}_{13}$

We now give a small example to show the possible case distinctions that can occur in the graphs $\mathcal{G}_p$. We take $p = 13$, since this yields a small graph that still exhibits most of the subtleties and pathologies that one might encounter in larger graphs.

Figure 1 shows $\mathcal{G}_{13}$. There are 3 superspecial genus-2 curves defined over $\overline{\mathbb{F}}_{13}$ up to isomorphism, say C_i for i in $\{1, 2, 3\}$; we denote their jacobians by J_{C_i}. There is only 1 supersingular elliptic curve defined over $\overline{\mathbb{F}}_{13}$ up to isomorphism, say E, so there is only one vertex in $\mathcal{G}_{13}$ that corresponds to a product of elliptic curves.

First of all it is easily verifiable that there are at most 6 outgoing edges from any J_{C_i} to $E \times E$, see Appendix A. Furthermore, since clearly $E \cong E$, there are strictly fewer than 6 outgoing edges from $E \times E$ to jacobians of genus-2 curves. Since the j-invariant of E is not in $\{0, 1728\}$, we know there are exactly 5 like that, so the remaining 10 edges must go to products of elliptic curves as well, which here (by lack of other options) means a loop with multiplicity 10.

This example also shows clearly why direction is important in the graph. There are 4 edges from J_{C_1} to J_{C_2}, but only 1 edge back. In other words $C_1 : y^2 = x^5 - x$ has 4 quadratic splittings whose associated Richelot isogenies have J_{C_2} as codomain,[4] while starting from any Weierstraß equation for C_2, only one quadratic splitting gives rise to a Richelot isogeny with J_{C_1} as codomain. This stems from the fact that the 4 corresponding $(2, 2)$-subgroups of J_{C_1} are mapped to each other by an automorphism of J_{C_1}. In other words the 4 resulting isogenies

$$\phi_1, \ldots, \phi_4 : J_{C_1} \to J_{C_2}$$

are obtained from one another by pre-composition with such an automorphism. But then their duals

$$\hat{\phi}_1, \ldots, \hat{\phi}_4 : J_{C_2} \to J_{C_1}$$

are obtained from each other by *post*-composition with an automorphism. In particular they have the same kernel or, equivalently, they correspond to the same quadratic splitting.

The only thing missing from the graph is a vertex corresponding to a product of non-isomorphic elliptic curves. Such a vertex always has 9 outgoing edges (possibly loops) to other vertices in $\mathcal{E}_p$, and 6 outgoing edges to vertices in $\mathcal{J}_p$. The smallest example of this phenomenon is in the graph $\mathcal{G}_{17}$, which already has double the number of vertices of $\mathcal{G}_{13}$.

6. A special class of paths in $\mathcal{G}_p$

We are interested in the kinds of isogenies that are represented by paths in $\mathcal{G}_p$: that is, the compositions of isogenies corresponding to adjacent edges.

First, fix a single edge $\phi_0 : A_0 \to A_1$ in $\mathcal{G}_p$. By definition, ϕ_0 represents (up to isomorphism) a $(2, 2)$-isogeny: that is, an isogeny whose kernel is a maximal 2-Weil isotropic subgroup of $A_0[2]$, hence isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$.

Now, consider the set of edges leaving A_1: these correspond to $(2, 2)$-isogenies that may be composed with ϕ_0. We know that (counting multiplicity) there are fifteen such edges. These edges fall naturally into three classes *relative to* ϕ_0,

[4] Up to isomorphism, that is: the resulting equations for the curve C_2 are in fact different, but the Cardona–Quer invariants are the same.

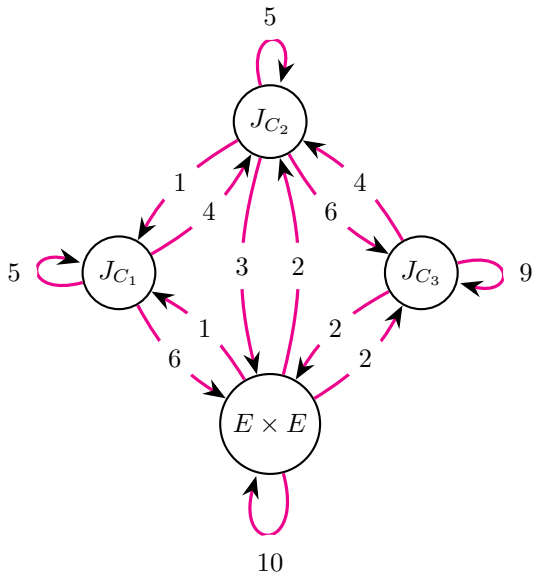


FIGURE 1. The graph $\mathcal{G}_{13}$. The vertices J_{C_i}, $i \in \{1,2,3\}$, correspond to jacobians of genus-2 curves, whereas the vertex $E \times E$ corresponds to a product of elliptic curves. The numbers indicate the multiplicities of the edges.

according to the structure of the kernel of the composed isogeny (which, in each case, is a maximal 4-Weil isotropic subgroup of $A_0[4]$).

Definition 1. Let $\phi_0 : A_0 \to A_1$ and $\phi_1 : A_1 \to A_2$ be edges in $\mathcal{G}_p$.

- We say that ϕ_1 is the (necessarily unique) **dual extension** of ϕ_0 if $\ker(\phi_1 \circ \phi_0) \cong (\mathbb{Z}/2\mathbb{Z})^4$, so $\phi_1 \circ \phi_0$ is a $(2,2,2,2)$-isogeny (hence isomorphic to $[2]_{A_0}$). In this case, $\ker \phi_1 = \phi_0(A_0[2])$.
- We say that ϕ_1 is a **bad extension** of ϕ_0 if $\ker(\phi_1 \circ \phi_0) \cong (\mathbb{Z}/4\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})^2$, so $\phi_1 \circ \phi_0$ is a $(4,2,2)$-isogeny. In this case $(\ker \phi_1) \cap \phi_0(A_0[2]) \cong \mathbb{Z}/2\mathbb{Z}$, so there are precisely 6 bad extensions of any given ϕ_0.
- We say that ϕ_1 is a **good extension** of ϕ_0 if $\ker(\phi_1 \circ \phi_0) \cong (\mathbb{Z}/4\mathbb{Z})^2$, so $\phi_1 \circ \phi_0$ is a $(4,4)$-isogeny. In this case $(\ker \phi_1) \cap \phi_0(A_0[2]) = 0$, so there are precisely 8 good extensions of any ϕ_0.

Remark 1. In [25, Definition 9.2.1], *good* extensions are called *cyclic* and *bad* extensions are called *acyclic*. We prefer the good/bad terminology here to avoid confusion with the notion of composing isogenies to form eventual cycles in $\mathcal{G}_p$; the reason why good is good and bad is bad will become clear in Section 7.

We have seen how the three kinds of extensions

$$A_0 \xrightarrow{\phi_1} A_1 \xrightarrow{\phi_2} A_2$$

can be distinguished by how the kernel of ϕ_2 intersects with the image of $A_0[2]$ under ϕ_1. We can make these criteria more explicit in terms of the Richelot isogeny formulas.

6.1. **Extensions of isogenies from $\mathcal{J}_p$ to $\mathcal{J}_p$.** Recall the construction of Richelot isogenies $\phi_1 : J_{C_0} \to J_{C_1}$ from Proposition 1: given the curve $C_0 : y^2 = G_1 \cdot G_2 \cdot G_3$, we set

$$H_1 := G_2'G_3 - G_3'G_2\,, \qquad H_2 := G_3'G_1 - G_1'G_3\,, \qquad H_3 := G_1'G_2 - G_2'G_1\,.$$

The curve C_1 is defined by $C_1 : y^2 = \delta^{-1} \cdot H_1 \cdot H_2 \cdot H_3$ where $\delta := \det(G_1, G_2, G_3)$. The kernel of ϕ_1 corresponds to $\{G_1, G_2, G_3\}$, and the subgroup $\phi_1(J_{C_0}[2]) \subset J_{C_1}[2]$ corresponds to $\{H_1, H_2, H_3\}$.

Proposition 2. *With the notation above: if*

$$H_1 = L_1 \cdot L_2\,, \qquad H_2 = L_3 \cdot L_4\,, \qquad H_3 = L_5 \cdot L_6\,,$$

with the L_i all linear (except possibly for one constant L_i in the case where $H_1H_2H_3$ is quintic), then the good extensions of ϕ_1 are the Richelot isogenies with kernels corresponding to one of the following factorizations of $H_1H_2H_3$:

$$\begin{aligned}
&(L_1L_3, L_2L_5, L_4L_6), \quad (L_1L_3, L_2L_6, L_4L_5),\\
&(L_1L_4, L_2L_5, L_3L_6), \quad (L_1L_4, L_2L_6, L_3L_5),\\
&(L_1L_5, L_2L_3, L_4L_6), \quad (L_1L_5, L_2L_4, L_3L_6),\\
&(L_1L_6, L_2L_3, L_4L_5), \quad (L_1L_6, L_2L_4, L_3L_5).
\end{aligned}$$

Proof. The quadratic splitting $\{H_1, H_2, H_3\}$ corresponds to the subgroup of $J_{C_1}[2]$ which is the kernel of the dual $\hat{\phi}_1$, and also the image $\phi_1(J_{C_0}[2])$. The good extensions of ϕ_1 are those whose kernel intersects trivially with $\phi_1(J_{C_2}[2])$; they therefore correspond to the quadratic splittings with no quadratics proportional to any of the H_i. The list of 8 splittings above follows from direct calculation. □

We now discuss the good extensions of isogenies involving products of elliptic curves. This is mainly for the sake of completeness, because in our proposed hash function below, these cases will not be implemented.

6.2. **Extensions of isogenies from $\mathcal{J}_p$ to $\mathcal{E}_p$.** Recall from the preliminaries that for a $(2,2)$-isogeny $\phi_1 : J_{C_0} \to E_1 \times E_2$, the domain can be written as the jacobian of a curve $C_0 : y^2 = G_1G_2G_3$, where

$$G_i = a_{i,1}(x - s_1)^2 + a_{i,2}(x - s_2)^2$$

for certain $s_1, s_2, a_{i,1}, a_{i,2} \in \mathbb{F}_{p^2}$ for $i = 1, 2, 3$. The elliptic curves determining the codomain can then be defined by the equations

$$E_1 : y^2 = \prod_{i=1}^{3}(a_{i,1}x + a_{i,2})\,, \qquad E_2 : y^2 = \prod_{i=1}^{3}(a_{i,1} + a_{i,2}x)\,.$$

For $i = 1, 2, 3$ we will write $\{\alpha_i, \alpha_i'\}$ for the roots of G_i, $P_i = (-a_{i,2}/a_{i,1}, 0)$ for the Weierstraß points of E_1, $Q_i = (-a_{i,1}/a_{i,2}, 0)$ for the Weierstraß points of E_2, and $\mathcal{O}_{E_1}$ and $\mathcal{O}_{E_2}$ for the neutral element of respectively E_1 and E_2.

Proposition 3. *With the notation above, the good extensions of ϕ_1 are the $(2,2)$-isogenies with kernel one of the 6 combinations*

$$\{(\mathcal{O}_{E_1}, \mathcal{O}_{E_2}), (P_i, \mathcal{O}_{E_2}), (\mathcal{O}_{E_1}, Q_j), (P_i, Q_j)\},$$

for $i \neq j$ in $\{1,2,3\}$, or one of

$$\{(\mathcal{O}_{E_1}, \mathcal{O}_{E_2}), (P_1, Q_2), (P_2, Q_3), (P_3, Q_1)\},$$
$$\{(\mathcal{O}_{E_1}, \mathcal{O}_{E_2}), (P_1, Q_3), (P_2, Q_1), (P_3, Q_2)\}.$$

Proof. The proof of the formulas in [25, Proposition 8.3.1] shows that, for $\{i,j,k\} = \{1,2,3\}$, the 2-torsion elements $[(\alpha_i, 0) - (\alpha_j, 0)]$, $[(\alpha_i, 0) - (\alpha'_j, 0)]$, $[(\alpha'_i, 0) - (\alpha_j, 0)]$, $[(\alpha'_i, 0) - (\alpha'_j, 0)]$ get mapped to (P_k, Q_k) in $E_1 \times E_2$. So the good extensions of ϕ_1 are the isogenies whose kernels intersect

$$\phi_1(J_{C_0}[2]) = \{(\mathcal{O}_{E_1}, \mathcal{O}_{E_2}), (P_1, Q_1), (P_2, Q_2), (P_3, Q_3)\}$$

trivially, which are exactly the ones listed. □

Note that in the previous proposition, the 6 good extensions of the first type always have a product of elliptic curves as codomain. The other 2 will typically be to a jacobian of a genus-2 curve, unless $E_1 \cong E_2$ and the given kernel is contained in the graph of an isomorphism $E_1 \to E_2$.

6.3. **Extensions of isogenies from $\mathcal{E}_p$ to $\mathcal{J}_p$.** Recall that every $(2,2)$-isogeny $\phi_1 : E_1 \times E_2 \to J_{C_1}$, with

$$E_1 : y^2 = \prod_{i=1}^{3} (x - \alpha_i)\,, \qquad E_2 : y^2 = \prod_{i=1}^{3} (x - \beta_i)\,,$$

always has codomain the jacobian of a genus-2 curve C_1 that can be defined by an equation of the form

$$\begin{aligned} y^2 = &- \big(A(\alpha_2 - \alpha_1)(\alpha_1 - \alpha_3)x^2 + B(\beta_2 - \beta_1)(\beta_1 - \beta_3)\big) \\ &\cdot \big(A(\alpha_3 - \alpha_2)(\alpha_2 - \alpha_1)x^2 + B(\beta_3 - \beta_2)(\beta_2 - \beta_1)\big) \\ &\cdot \big(A(\alpha_1 - \alpha_3)(\alpha_3 - \alpha_2)x^2 + B(\beta_1 - \beta_3)(\beta_3 - \beta_2)\big)\,, \end{aligned} \tag{1}$$

up to permutation of the roots β_i, for well-defined nonzero constants A and B that depend on α_i and β_i. We will denote the quadratic factors on the right hand side of Equation 1 on the first, second and third line by H_1, H_2 and H_3 respectively, such that $C_1 : y^2 = -H_1 \cdot H_2 \cdot H_3$.

Proposition 4. *With the notation above: if*

$$H_1 = L_1 \cdot L_2\,, \qquad H_2 = L_3 \cdot L_4\,, \qquad H_3 = L_5 \cdot L_6\,,$$

with the L_i all linear (except possibly for one constant L_i in the case where $H_1H_2H_3$ is quintic), then the good extensions of ϕ_1 are the Richelot isogenies with kernels corresponding to one of the following factorizations of $H_1H_2H_3$:

$$\begin{aligned} &(L_1L_3, L_2L_5, L_4L_6), \quad (L_1L_3, L_2L_6, L_4L_5), \\ &(L_1L_4, L_2L_5, L_3L_6), \quad (L_1L_4, L_2L_6, L_3L_5), \\ &(L_1L_5, L_2L_3, L_4L_6), \quad (L_1L_5, L_2L_4, L_3L_6), \\ &(L_1L_6, L_2L_3, L_4L_5), \quad (L_1L_6, L_2L_4, L_3L_5). \end{aligned}$$

Proof. The proof of Equation 1 in [17] constructs the dual isogeny $\hat{\phi}_1 : J_{C_1} \to E_1' \times E_2'$, where $E_1' \cong E_1$ and $E_2' \cong E_2$. More specifically, E_1' and E_2' are given by

$$\begin{aligned} E_1' : y^2 = &-(A(\alpha_2-\alpha_1)(\alpha_1-\alpha_3)x + B(\beta_2-\beta_1)(\beta_1-\beta_3)) \\ &\cdot (A(\alpha_3-\alpha_2)(\alpha_2-\alpha_1)x + B(\beta_3-\beta_2)(\beta_2-\beta_1)) \\ &\cdot (A(\alpha_1-\alpha_3)(\alpha_3-\alpha_2)x + B(\beta_1-\beta_3)(\beta_3-\beta_2)), \end{aligned}$$

$$\begin{aligned} E_2' : y^2 = &-(A(\alpha_2-\alpha_1)(\alpha_1-\alpha_3) + B(\beta_2-\beta_1)(\beta_1-\beta_3)x) \\ &\cdot (A(\alpha_3-\alpha_2)(\alpha_2-\alpha_1) + B(\beta_3-\beta_2)(\beta_2-\beta_1)x) \\ &\cdot (A(\alpha_1-\alpha_3)(\alpha_3-\alpha_2) + B(\beta_1-\beta_3)(\beta_3-\beta_2)x). \end{aligned}$$

Hence the quadratic splitting $\{H_1, H_2, H_3\}$ corresponds to the subgroup of $J_{C_1}[2]$ which is the kernel of the dual $\hat{\phi}_1$ and we can continue the proof just as in the Richelot isogeny case. □

6.4. Extensions of isogenies from $\mathcal{E}_p$ to $\mathcal{E}_p$.

Proposition 5. *Let $\phi_1 : E_1 \times E_2 \to E_1' \times E_2'$ be a $(2,2)$-isogeny. Denote by $\mathcal{O}_{E_1}$, $\mathcal{O}_{E_2}$, $\mathcal{O}_{E_1'}$, $\mathcal{O}_{E_2'}$ the identity elements of respectively E_1, E_2, E_1' and E_2'. For $i = 1,2,3$ we write P_i, Q_i, P_i', Q_i' for the Weierstraß points of respectively E_1, E_2, E_1', E_2'. If*

$$\ker(\phi_1) = \{(\mathcal{O}_{E_1}, \mathcal{O}_{E_2}), (P_1, \mathcal{O}_{E_2}), (\mathcal{O}_{E_1}, Q_1), (P_1, Q_1)\},$$

and $\phi_1|_{E_1}(P_2) = \phi_1|_{E_1}(P_3) = P_1'$, $\phi_1|_{E_2}(Q_2) = \phi_1|_{E_2}(Q_3) = Q_1'$, then the good extensions of ϕ_1 are the isogenies with kernel one of the 4 combinations

$$\{(\mathcal{O}_{E_1'}, \mathcal{O}_{E_2'}), (P_i', \mathcal{O}_{E_2'}), (\mathcal{O}_{E_1'}, Q_j'), (P_i', Q_j')\},$$

where $i \neq 1$ and $j \neq 1$, or one of

$$\begin{aligned} &\{(\mathcal{O}_{E_1'}, \mathcal{O}_{E_2'}), (P_1', Q_2'), (P_2', Q_3'), (P_3', Q_1')\}, \\ &\{(\mathcal{O}_{E_1'}, \mathcal{O}_{E_2'}), (P_1', Q_3'), (P_2', Q_1'), (P_3', Q_2')\}, \\ &\{(\mathcal{O}_{E_1'}, \mathcal{O}_{E_2'}), (P_1', Q_2'), (P_2', Q_1'), (P_3', Q_3')\}, \\ &\{(\mathcal{O}_{E_1'}, \mathcal{O}_{E_2'}), (P_1', Q_3'), (P_2', Q_2'), (P_3', Q_1')\}. \end{aligned}$$

Proof. The good extensions are determined by the $(2,2)$-isogenies that intersect

$$\{(\mathcal{O}_{E_1'}, \mathcal{O}_{E_2'}), (P_1', \mathcal{O}_{E_2'}), (\mathcal{O}_{E_1'}, Q_1'), (P_1', Q_1')\}$$

trivially, so the proof is immediate. □

6.5. Connectedness.

Conjecture 3. *For every 2 vertices A and A' in $\mathcal{J}_p \subset \mathcal{G}_p$, there exists a path*

$$A = A_0 \xrightarrow{\phi_0} A_1 \xrightarrow{\phi_1} \cdots \xrightarrow{\phi_{k-1}} A_k = A'$$

of k edges, for some $k \geq 0$, such that all of the A_i are in $\mathcal{J}_p$ and each $\phi_i, i \neq 0$, is a good extension of ϕ_{i-1}. (The composed isogeny is then a $(2^k, 2^k)$-isogeny.)

Conjecture 3 is our strongest conjecture. It differs from Conjecture 2 in that at each step in a path, the number of choices is reduced from all 15 isogenies to the 8 good isogenies. Conjecture 3 is easy to verify for small p using the formulas for Richelot isogenies and the exact formula from Theorem 1. We verified this part of the conjecture for $p \leq 1013$ using Magma, but from then onward the computations

become slow since we work with graphs of several hundred thousands of vertices already. Nonetheless, this is a first indication that Conjecture 3 might hold.

7. Hash functions from Richelot isogenies

Turning the graph $\mathcal{G}_p$ into a hash function happens analogously to the elliptic curve case with some small caveats. We will first describe the function, thereby repairing Takashima's proposal from [27], and then work out some small remarks and argue why certain choices were made.

We start by choosing a large prime p (in function of some security parameter λ) such that $p \equiv 5 \bmod 6$. We start at the vertex corresponding to the jacobian of the genus-2 curve C_0 defined over $\mathbb{F}_{p^2}$, given by the equation $y^2 = x(x-1)(x+1)(x-2)(x-1/2)$. The hash function starts by multiplying the input by 8^{10}, or equivalently, padding it with 30 zeroes. The hashing will happen 3 bits at a time, with each three bits determining a choice of one of the eight good extensions relative to the previous step. So for our starting vertex we will need to make an initial choice as if we performed a step prior to starting. The quadratic splitting we will choose for C_0 is

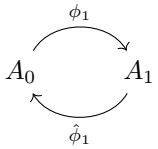
$$\left\{x^2 - 1, x^2 - 2x, x - \frac{1}{2}\right\}.$$

The 8 quadratic splittings that we will consider are those that have *no* quadratic factor in common with the one that was obtained from the previous step. These splittings are then ordered according to some lexicographical order of the roots. In practice this means we just need to fix a quadratic equation that determines the field extension $\mathbb{F}_p \subseteq \mathbb{F}_{p^2}$. Next we process 3 bits of our input according to the order of the 8 chosen edges. If the chosen edge leads to a vertex corresponding to the product of elliptic curves, the function stops and outputs an error. If the chosen edge leads to a vertex corresponding to a jacobian of a genus-2 curve then we rinse and repeat for the next 3 bits. Once the entire message has been processed we output the Cardona–Quer invariants of the genus-2 curve corresponding to the vertex we ended up in.

7.1. **Avoiding trivial cycles.**

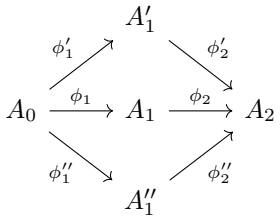
A hash function should be collision resistant, so we need to at least avoid trivial cycles in our graph. In the elliptic curve case, this is simply done by disallowing the edge associated to the dual isogeny from where we just came.

Similarly, we must avoid using dual isogenies when walking in $\mathcal{G}_p$, to avoid extremely easy cycles:



But there is an additional subtlety in genus-2, as noted in [16]. If we follow a $(2,2)$-isogeny $A_0 \to A_1$ with a bad extension $A_1 \to A_2$, then we end up with a $(4,2,2)$-isogeny; but then, for every $(4,2,2)$-isogeny $A_0 \to A_2$ there are 3 distinct ways to split it up into the concatenation of two $(2,2)$-isogenies as in the following

diagram.



Luckily all these cases are easy to distinguish, as we saw in Section 6.

The eight $(2,2)$-isogenies corresponding to cyclic extensions do not result in trivial cycles. In practice this means that, after a choice for our initial $(2,2)$-isogeny corresponding to one of 15 possible edges, we are left with only 8 options at every next step along the way. This implies that we should not only keep track of our current vertex by some form of equation, but also by some order of the roots of that equation (or more precisely, by a quadratic splitting).

This observation means we can hash up to 3 bits at every step in our hash function and that a hash will always correspond to computing a $(2^k, 2^k)$-isogeny.

7.2. **Products of elliptic curves.** For our hash function, the vertices corresponding to products of elliptic curves are a nuisance for the following reasons.

- There is no clear candidate invariant that is similar to the ordered triple in case of the genus-2 Cardona–Quer invariants. So ideally, we would prefer not to end the hash function in a vertex like this.
- The formulas involving products of elliptic curves are a lot more involved than the Richelot isogenies, and their simplicity was one of the main reasons for the restriction to $(2,2)$-isogenies.

In the way we presented our hash function, we simply use Richelot isogenies only and let our hash function break down whenever we pass at a vertex corresponding to a product of elliptic curves. Given that this only occurs with probability $O(1/p)$, this only happens with negligible probability for practical values of p.

An alternative way of dealing with this is as follows. Assume we try to process a step in our hash function that corresponds to a $(2,2)$-isogeny between a jacobian of a genus-2 curve and a product of elliptic curves. Then (in the same step) we immediately choose one edge corresponding to a $(2,2)$-isogeny from the product of elliptic curves back to a jacobian of a genus-2 curve. This has to be done in a deterministic way and we should avoid the dual and bad extensions since they would result in small cycles in $\mathcal{G}_p$. Unfortunately Proposition 3 tells us that we can only find 2 good extensions that possibly have the jacobian of a genus-2 curve as codomain. In the case of $E \times E$, with E having j-invariant 0 or 1728, these kernels may both be to a product of elliptic curves again. Solving this issue can be done by either choosing $p \equiv 1 \bmod 12$ (such that elliptic curves with j-invariant 0 and 1728 never occur), or by (deterministically) using the results from Proposition 5 to add an extra step in this specific case.

A third option is to keep working with all the formulas for products of elliptic curves as well. This means we should find a way to merge the Cardona-Quer invariants and (unordered) pairs of j-invariants into one output type, which is only an issue when *ending* in a product of elliptic curves.

7.3. **Initial choices.** As mentioned earlier, there is no known way to generate the equation of a random superspecial genus-2 curve that is defined over $\mathbb{F}_{p^2}$. Some specific examples such as $y^2 = x^5 - x$ with $p \equiv 5$ or $7 \bmod 8$ are listed in [19]. Unfortunately, the examples that are easiest to represent all have some $(2,2)$-isogenies with codomain the product of 2 supersingular elliptic curves. This seems to imply that we cannot avoid having to deal with vertices corresponding to products of elliptic curves, instead of ignoring them.

However, another initial choice to make is whether we start by picking one of 15 possible edges or already restrict ourselves to 8, since this is needed for every subsequent step anyway. We will take only 8 which means we need to choose an initial quadratic splitting instead of just an initial curve.[5] Fortunately this solves our problem of finding an appropriate starting curve in a way. Consider C_0, the genus-2 curve given by $y^2 = x(x-1)(x+1)(x-2)(x-1/2)$ defined over $\mathbb{F}_p$ with $p > 5$. Then C_0 is superspecial if and only if $p \equiv 5 \bmod 6$ [19]. Now the vertex corresponding to the jacobian of C_0 has 4 neighbours that are products of supersingular elliptic curves. However, if we take the initial quadratic splitting $\left\{x^2 - 1, x^2 - 2x, x - \frac{1}{2}\right\}$, then the 8 allowed outgoing $(2,2)$-isogenies all have the jacobian of a superspecial genus-2 curve as codomain. The only restriction this puts on our hash function is that we need to work with a prime p such that $p \equiv 5 \bmod 6$, but this is easy to enforce.

An issue that arises with this curve C_0 however, is that its jacobian has many automorphisms and hence has multiple outgoing isogenies with the same codomain.[6] More precisely, starting from the given splitting of C_0, the 8 good extensions only have 3 distinct codomains up to isomorphism, one of which even occurs with multiplicity 5, which leads to trivial cycles in our graph. An easy way to fix this is to simply take a (relatively short) deterministic path prior to starting to hash our input, or equivalently, pad the input with some zeroes from the right. For other possible starting curves, this padding can be used to additionally avoid products of elliptic curves. In the end, choosing a starting curve seems to be a choice between either a very compact representation, or not having to do any extra computations that avoid trivial collisions.

7.4. **Security.** The security of our hash function depends on the hardness of finding isogenies between certain p.p. abelian surfaces. A lot of the choices discussed in the previous subsections make slight alterations to the underlying mathematical hard problems. We will formulate them in a general form to keep them succinct since we don't think any of the changes would impact the hardness of the problems. In essence they are equivalent to the hard problems from the elliptic curve hash function in [7].

Problem 1. *Given two superspecial genus-2 curves C_1 and C_2 defined over $\mathbb{F}_{p^2}$, find a $(2^k, 2^k)$-isogeny between their jacobians.*

Problem 2. *Given any superspecial genus-2 curve C_1 defined over $\mathbb{F}_{p^2}$, find*

[5] Remark that in this case, Conjecture 3 is no longer strong enough to prove that we can reach all vertices in $\mathcal{G}_p$, since it relies on having all 15 initial $(2,2)$-isogenies present. However, there is no clear reason to assume that only allowing 8 out of 15 possible edges for our initial choice all of a sudden would disallow us to reach certain vertices.

[6] Remark that in the elliptic curve case the same thing happens with for example $y^2 = x^3 + x$ with $p \equiv 3 \bmod 4$.

(1) *a curve* C_2 *and a* $(2^k, 2^k)$*-isogeny* $J_{C_1} \to J_{C_2}$,
(2) *a curve* C_2' *and a* $(2^{k'}, 2^{k'})$*-isogeny* $J_{C_1} \to J_{C_2'}$,

such that C_2 *and* C_2' *are* $\overline{\mathbb{F}}_p$*-isomorphic. Here, it is allowed that* $k = k'$ *but in this case the kernels should be different.*

They are related to our hash function in the following way.

- *Preimage resistance:* Finding a preimage in our hash function implies a solution to Problem 1 with $C_1 = C_0$ as follows. Let C_2 be a representative of the isomorphism class of the output of the hash function. A preimage for that output corresponds to a path of length k in our graph, or equivalently, a $(2^k, 2^k)$-isogeny between the jacobians of C_0 and C_2.
- *Collision resistance:* Finding a collision in our hash function implies a solution to Problem 1 with $C_1 = C_0$ as follows. A collision in our hash function corresponds to two distinct paths in our graph with the same ending vertex. Equivalently this amounts to a pair of isogenies

$$\phi : J_{C_0} \to J_{C_2} \quad \text{and} \quad \phi' : J_{C_0} \to J_{C_2'}$$

of type $(2^k, 2^k)$ resp $(2^{k'}, 2^{k'})$ such that $C_2 \cong C_2'$, and with different kernels.

To our knowledge, there are no known ways to find isogenies of the said kinds between jacobians of (superspecial) genus-2 curves which perform better than the generic attacks. In the classical case the best known such attack is Pollard-rho, which can find a collision or preimage in time complexity the square root of the number of possibilities times the amount of time that one step computation takes. In our case we have a graph of size $O(p^3)$ and one step is simply a polynomial computation with some constants, which we can perform in time complexity $\log p$. Hence a Pollard-rho attack could find a solution to Problem 1 or Problem 2 in time $\tilde{O}(p^{3/2})$.

With quantum computers in mind, the best known attack is a claw finding algorithm to find a collision or preimage in the graph $\mathcal{G}_p$ [29]. An attack like that has time complexity the third root of the size of the graph we work over, instead of the square root in the classical case. This implies we could find a solution to Problem 1 or Problem 2 in time $\tilde{O}(p)$.

8. Implementation and timings

We have implemented our hash function in Magma, taking into account all the choices made from the previous section. The pseudocode can be found below; the Magma code can be found in Appendix B. The subroutine `Factorization` is defined as follows: when the input is a quadratic polynomial, `Factorization` returns its two linear factors (which, in this application, are guaranteed to exist over the ground field). When the input is a linear polynomial, it returns that polynomial and 1.

Remark 2. We don't keep track of the leading coefficient of the polynomial determining the genus-2 curve, for the reason that a twist of a curve does not change its Cardona–Quer invariants anyway. Similarly we never need to know the exact value of $\delta = \det(G_1, G_2, G_3)$. We are only interested in whether or not δ equals 0, and with the formulas from the preliminaries, this condition can be easily verified to be equivalent to all H_i being a multiple of one another. Hence it suffices to check if $\operatorname{rank}(H_1, H_2) < 2$ instead.

Algorithm 1: Hashing a message m using Richelot isogenies, with λ bits of security on a classical computer

Data: Message m and security parameter λ
Result: The hash of m using Richelot isogenies in a graph $\mathcal{G}_p$, or $\perp$ (failure)

$S \leftarrow [(\{1,3\},\{2,5\},\{4,6\}),\ldots,(\{1,6\},\{2,4\},\{3,5\})]$
$p \leftarrow$ the smallest prime such that $p > 2^{\lceil 2\lambda/3 \rceil}$ and $p \equiv 5 \bmod 6$
$(L_1, L_2, L_3, L_4, L_5, L_6) \leftarrow (x-1, x+1, x, x-2, x-1/2, 1) \in \mathbb{F}_{p^2}[x]^6$
$m \leftarrow 2^{30}m$
while $m > 0$ **do**
  $i \leftarrow m \bmod 8$
  $m \leftarrow (m-i)/8$
  $[G_1, G_2, G_3] \leftarrow$ pairwise products of the L_j according to $S[i]$
  $(L_1, L_2) \leftarrow$ `Factorization`(H_1) **where** $H_1 := G_2'G_3 - G_2G_3'$
  $(L_3, L_4) \leftarrow$ `Factorization`(H_2) **where** $H_2 := G_3'G_1 - G_3G_1'$
  if $\mathrm{rank}(H_1, H_2) = 1$ **then**
    return $\perp$ `// We have hit a vertex in` $\mathcal{E}_p$
  $(L_5, L_6) \leftarrow$ `Factorization`(H_3) **where** $H_3 := G_1'G_2 - G_1G_2'$
return *invariants of genus-2 curve defined by the equation* $y^2 = \prod_{j=1}^{6} L_j$

The deterministic choice of ordering the edges depends on 2 things. First, there's the (arbitrary) way we hardcoded the set S, which denotes the pairs of indices of the allowed quadratic splittings. Secondly, the subroutine `Factorization` automatically orders the roots of the polynomial in some way. In this statement we silently assumed that this happens deterministically by the used software, which is the case for Magma.

Note that we do not claim this code is optimized in any way. For example we simply pick the smallest prime p possible that satisfies our needs, whereas better choices may speed up the arithmetic in the field we work over. Additionally, we did not implement any proper padding schemes, nor did we make the function constant time (though this is not required for public inputs). The main goal of the implementation is to see what the order of magnitude is for the speed of the hash function and we leave possible optimizations for future work.

As a final remark we want to point out that the output of the hash function is dependent on the security level required. The output is a triple in a quadratic field extension of a finite field of characteristic roughly $2\lambda/3$ bits in case of classical security. This means our output has bit length 4λ, even though the number of possible hash values is only 2λ bits. It may be possible to compress this but we leave this discussion for future research, too.

The implementation of our genus two CGL hash function algorithm was done in Magma (version 2.32-2) on an Intel(R) Xeon(R) CPU E5-2630 v2 @ 2.60GHz with 128 GB memory. For every prime size we averaged the speed over 1000 random inputs of 100 bits. A summary of our timed results can be found in the following table.

	$p \approx 2^{86}$	$p \approx 2^{128}$	$p \approx 2^{171}$	$p \approx 2^{256}$
bits of classical security	128	192	256	384
bits of quantum security	86	128	170	256
time per bit processed	5.01ms	6.52ms	9.33ms	15.70ms
output bits	516	768	1026	1536

9. Comparison to Charles–Goren–Lauter, and concluding remarks

The computational cost of each iteration of the main loop in Algorithm 1 is dominated by the costs of the three square roots required to factor the H_i in Lines 9, 10, and 13. At first glance, this would appear to give no advantage over the Charles–Goren–Lauter hash function: we compute essentially one expensive square root per bit of hash input. However, there are two important remarks to be made here:

(1) The entropy in the Charles–Goren–Lauter hash function is linear in p, whereas in our case it is cubic in p. This implies that for the same security parameters we can work over much smaller finite fields, so the square roots are substantially easier to compute.
(2) The square roots, along with the H_i, can be computed completely independently. The algorithm therefore lends itself well to three-well parallelization, as well as to vectorization techniques on suitable computer architectures.

From this point of view, our proposal is a conjecturally secure version of an ill-constructed hash function that we could call 3CGL, where the message m is split up in 3 chunks m_1, m_2, m_3. Each of these m_i is then hashed using Charles, Goren and Lauter's hash function into a supersingular j-invariant j_i, resulting in a combined hash value $(j_1, j_2, j_3) \in \mathbb{F}_{p^2}$. Note that, here too, the number of possible outcomes is $O(p^3)$. However, the security of 3CGL clearly reduces to the problem of finding collisions or pre-images for one of the chunks, which Pollard-rho can do in time $\tilde{O}(p^{1/2})$, compared to $\tilde{O}(p^{3/2})$ in our case.

While this convinces us that genus 2 hash functions deserve their place in the arena of isogeny-based cryptography, more research is needed to have a better assessment of their security and performance. One potentially interesting track is to adapt Doliskani, Pereira and Barreto's recent speed-up to Charles, Goren and Lauter's hash function from [14], which has the appearance of an orthogonal improvement that may also apply to genus 2. From a security point of view, it would be interesting to understand to what extent the discussion from [22, 15], transferring the elliptic curve analogs of Problems 1 and 2 to questions about orders in non-commutative algebras and raising some concerns about using special starting curves, carries over to genus 2.

Acknowledgements. We are grateful to Yan Bo Ti for sharing with us a preliminary copy of [16] and to Frederik Vercauteren for helpful feedback with regards to this paper. This work was supported in part by the Research Council KU Leuven grants C14/18/067 and STG/17/019.

References

[1] Reza Azarderakhsh, Brian Koziel, Matt Campagna, Brian LaMacchia, Craig Costello, Patrick Longa, Luca De Feo, Michael Naehrig, Basil Hess, Joost Renes, Amir Jalali, Vladimir Soukharev, David Jao, and David Urbanik. Supersingular isogeny key encapsulation, 2017.

[2] Bradley W Brock. *Superspecial curves of genera two and three*. PhD thesis, Princeton University, 1994.
[3] Nils Bruin and Kevin Doerksen. The arithmetic of genus two curves with (4, 4)-split jacobians. *Canadian Journal of Mathematics*, 63(5):992–1024, 2011.
[4] Gabriel Cardona and Jordi Quer. Field of moduli and field of definition for curves of genus 2. In *Computational aspects of algebraic curves*, pages 71–83. World Scientific, 2005.
[5] Wouter Castryck, Tanja Lange, Chloe Martindale, Lorenz Panny, and Joost Renes. CSIDH: An efficient post-quantum commutative group action. In Thomas Peyrin and Steven Galbraith, editors, *Advances in Cryptology – ASIACRYPT 2018, Part III*, pages 395–427. Springer International Publishing, 2018.
[6] Denis X Charles, Eyal Z Goren, and Kristin E Lauter. Families of Ramanujan graphs and quaternion algebras. *Groups and symmetries: from Neolithic Scots to John McKay*, 47:53–63, 2009.
[7] Denis X Charles, Kristin E Lauter, and Eyal Z Goren. Cryptographic hash functions from expander graphs. *Journal of Cryptology*, 22(1):93–113, 2009.
[8] Jean-Marc Couveignes. Hard homogeneous spaces. Cryptology ePrint Archive, Report 2006/291, 2006.
[9] Luca De Feo and Steven Galbraith. SeaSign: Compact isogeny signatures from class group actions. In *Advances in Cryptology – EUROCRYPT 2019*, 2019. To appear.
[10] Luca De Feo, David Jao, and Jérôme Plût. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. *Journal of Mathematical Cryptology*, 8(3):209–247, 2014.
[11] Luca De Feo, Jean Kieffer, and Benjamin Smith. Towards practical key exchange from ordinary isogeny graphs. In Thomas Peyrin and Steven Galbraith, editors, *Advances in Cryptology – ASIACRYPT 2018, Part III*, pages 365–394. Springer International Publishing, 2018.
[12] Luca De Feo, Simon Masson, Christophe Petit, and Antonio Sanso. Verifiable delay functions from supersingular isogenies and pairings. Cryptology ePrint Archive, Report 2019/166, 2019.
[13] Thomas Decru, Lorenz Panny, and Frederik Vercauteren. Faster SeaSign signatures through improved rejection sampling. In *PQCrypto 2019*, 2019. To appear.
[14] Javad Doliskani, Geovandro C. Pereira, and Paulo S. Barreto. Faster cryptographic hash function from supersingular isogeny graphs. Cryptology ePrint Archive, Report 2017/1202, 2017.
[15] Kirsten Eisenträger, Sean Hallgren, Kristin Lauter, Travis Morrison, and Christophe Petit. Supersingular isogeny graphs and endomorphism rings: reductions and solutions. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in cryptology—EUROCRYPT 2018. Part III*, pages 329–368. Springer International Publishing, 2018.
[16] E Victor Flynn and Yan Bo Ti. Genus two isogeny cryptography. In *PQCrypto 2019*, 2019. To appear.
[17] Everett W Howe, Franck Leprévost, and Bjorn Poonen. Large torsion subgroups of split jacobians of curves of genus two or three. In *Forum Mathematicum*, volume 12.3, pages 315–364. Berlin; New York: De Gruyter, c1989-, 2000.
[18] Tomoyoshi Ibukiyama and Toshiyuki Katsura. On the field of definition of superspecial polarized abelian varieties and type numbers. *Compositio Mathematica*, 91(1):37–46, 1994.
[19] Tomoyoshi Ibukiyama, Toshiyuki Katsura, and Frans Oort. Supersingular curves of genus two and class numbers. *Compositio Mathematica*, 57(2):127–152, 1986.
[20] David Jao and Luca De Feo. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In *International Workshop on Post-Quantum Cryptography*, pages 19–34. Springer, 2011.
[21] Ernst Kani. The number of curves of genus two with elliptic differentials. *Journal für die reine und angewandte Mathematik*, 485:93–122, 1997.
[22] David Kohel, Kristin Lauter, Christophe Petit, and Jean-Pierre Tignol. On the quaternion ℓ-isogeny path problem. *LMS J. Comput. Math.*, 17(suppl. A):418–432, 2014.
[23] Ke-Zheng Li and Frans Oort. *Moduli of supersingular abelian varieties*, volume 1680 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998.
[24] Joseph H Silverman. *The arithmetic of elliptic curves*, volume 106. Springer Science & Business Media, 2009.
[25] Benjamin Smith. *Explicit endomorphisms and correspondences*. PhD thesis, University of Sydney, 2005.

[26] Anton Stolbunov. Public-key encryption based on cycles of isogenous elliptic curves. Master's thesis, Saint-Petersburg State Polytechnical University, 2004. In Russian.

[27] Katsuyuki Takashima. *Mathematical Modelling for Next-Generation Cryptography: CREST Crypto-Math Project*, chapter Efficient Algorithms for Isogeny Sequences and Their Cryptographic Applications, pages 97–114. Springer Singapore, Singapore, 2018.

[28] Katsuyuki Takashima and Reo Yoshida. An algorithm for computing a sequence of Richelot isogenies. *Bull. Korean Math. Soc*, 46(4):789–802, 2009.

[29] Seiichiro Tani. Claw finding algorithms using quantum walk. *Theoretical Computer Science*, 410(50):5285–5297, 2009.

Appendix A. Proof of Theorem 2

We now settle part (1) of Theorem 2, as an immediate consequence to:

Theorem 3. *Let C be a genus-2 curve over a field K of characteristic different from 2 and 5. Then the number of outgoing $(2,2)$-isogenies with codomain a product of elliptic curves is at most 6.*

Proof. We can assume that K is algebraically closed, so that C admits a model of the form $y^2 = \prod_{i=1}^{6}(x-\alpha_i)$ for roots $\alpha_i \in K$ satisfying

$$\prod_{1\leq i<j\leq 6}(\alpha_i - \alpha_j) = 1.$$

Due to the formulas for Richelot isogenies, the number of $(2,2)$-isogenies with codomain a product of elliptic curves is determined by how many among the 15 different equations of the form

$$\det \begin{pmatrix} 1 & \alpha_{\sigma(1)} + \alpha_{\sigma(2)} & \alpha_{\sigma(1)}\alpha_{\sigma(2)} \\ 1 & \alpha_{\sigma(3)} + \alpha_{\sigma(4)} & \alpha_{\sigma(3)}\alpha_{\sigma(4)} \\ 1 & \alpha_{\sigma(5)} + \alpha_{\sigma(6)} & \alpha_{\sigma(5)}\alpha_{\sigma(6)} \end{pmatrix} = 0\,, \tag{2}$$

where σ is a permutation of $\{1,2,3,4,5,6\}$, can be simultaneously satisfied.

To show that no more than 6 can occur we work with Gröbner bases. The permutations of equation (2) determine, up to sign, 15 different polynomials $f_1, \ldots, f_{15}$ in $\mathbb{F}[\alpha_1, \ldots, \alpha_6]$, where $\mathbb{F}$ is the prime subfield of K. We pick a subset of 7 of these equations and form the ideal $I \subset \mathbb{F}[\alpha_1, \ldots, \alpha_6]$ generated by them. We then add the polynomial $\rho = \prod_{i,j}(\alpha_i - \alpha_j) - 1$ as a generator of I as well. Now we determine a Gröbner basis G for I. If $G = \{1\}$ then the variety defined by I is empty and hence those 7 equations we chose can not be satisfied simultaneously, under the assumption that all α_i are different. If we repeat this process for all possible subsets of 7 equations and find $G = \{1\}$ in all cases, then we are done. There are $\binom{15}{7} = 6435$ possible ways of selecting such a subset, but this is not a problem for Magma.[7]

When running the algorithm we choose $\mathbb{F} = \mathbb{Q}$, for which we indeed find $G = \{1\}$ in each of the cases. This only shows that there are no solutions if K is of characteristic 0, while we typically want to work over a field with prime characteristic. If the Gröbner basis G equals $\{1\}$ however, we can write 1 as linear combination of that particular choice of polynomials f_i, say for example $1 = h_1 f_1 + \ldots + h_7 f_7 + h_8 \rho$. If we then multiply both sides of the equations by the lowest common multiple of the denominators of the coefficients of the h_i, we obtain an equation with coefficients in $\mathbb{Z}[\alpha_1, \ldots, \alpha_6]$. So as long as the characteristic p of the field we work over does not divide the lowest common multiple of the denominators of the coefficients of those h_i, we still find a contradictory system. Hence it suffices to keep track of the primes that divide the denominators. The resulting primes are 2, 3, 5, 7 and 11. It then suffices to rerun the Gröbner basis computations for $\mathbb{F} = \mathbb{F}_p$ with $p = 3, 7, 11$, leading to the desired conclusion. □

The following is the Magma code that was used. The specific cases $p \in \{3, 7, 11\}$ can be checked by replacing `Rationals()` by `GF(p)` for any one value of p, and by removing the innermost loop that starts with `for coord in c do` completely.

[7] Remark that by using the symmetry in the variables, it is possible to reduce the number of case distinctions needed, but we see no need to optimize this since it is a one time computation.

```
Q<a1,a2,a3,a4,a5,a6> := PolynomialRing(Rationals(),6);
S := {1,2,3,4,5,6};
I := {};

for sub1 in Subsets(S,2) do
  subseq1 := SetToSequence(sub1);
  for sub2 in Subsets(S diff sub1, 2) do
    subseq2 := SetToSequence(sub2);
    subseq3 := SetToSequence(S diff (sub1 join sub2));
    M := Matrix(Q,3,3,
               [ 1, Q.subseq1[1] + Q.subseq1[2], Q.subseq1[1]*Q.subseq1[2],
                 1, Q.subseq2[1] + Q.subseq2[2], Q.subseq2[1]*Q.subseq2[2],
                 1, Q.subseq3[1] + Q.subseq3[2], Q.subseq3[1]*Q.subseq3[2] ] );
    eqn := Determinant(M);
    if -eqn notin I then
      I join:= {Determinant(M)};
    end if;
  end for;
end for;

disc := Q ! 1;
for sub in Subsets(S,2) do
  subseq := SetToSequence(sub);
  disc *:= Q.subseq[1] - Q.subseq[2];
end for;

groebnerboolean := true;
badprimes := {};
for j in Subsets(I,7) do
    J := {disc-1};
    J join:= j;
    if GroebnerBasis(Ideal(J)) ne [1] then groebnerboolean := false; end if;
    J := IdealWithFixedBasis(SetToSequence(J));
    c := Coordinates(J, Q ! 1);
    for coord in c do
        for coeff in Coefficients(coord) do
            badprimes join:= SequenceToSet(PrimeDivisors(Denominator(coeff)));
        end for;
    end for;
end for;
print groebnerboolean; badprimes;
```

Theorem 3 cannot be proved in this way for $p = 2$, because equations for hyperelliptic curves are a lot more involved in fields of even characteristic. Theorem 2 remains true in this case however, since there are no superspecial genus-2 jacobians in fields of even characteristic.

The following example shows why Theorem 2 is not true for $p = 5$, and also provides an example to show that the bound of 6 is sharp.

Example 1. Let C be the genus-2 curve given by $y^2 = x^5 - x$ over $\mathbb{F}_p$ (which is superspecial when $p \equiv 5 \bmod 8$), and let $i \in \mathbb{F}_{p^2}$ be a square root of -1. Of the

fifteen quadrating splittings of $x^5 - x$, the six splittings

$$\{x, x^2 - (i+1)x + i, x^2 + (i+1)x + i\}, \quad \{x, x^2 + (i-1)x - i, x^2 - (i-1)x - i\}$$
$$\{x - 1, x^2 + 1, x^2 + x\}, \quad \{x + 1, x^2 + 1, x^2 - x\},$$
$$\{x - i, x^2 - 1, x^2 + ix\}, \quad \{x + i, x^2 - 1, x^2 - ix\}$$

all have $\delta = 0$, so they are always singular. The quadratic splitting $\{x, x^2+1, x^2-1\}$ has $\delta = \pm 2$ (the sign of δ may change with the order of the factors), and so is never singular. There are eight splittings remaining. The four splittings

$$\{x - 1, x^2 - ix, x^2 + (i+1)x + i\}, \quad \{x - i, x^2 + x, x^2 + (i-1)x - i\},$$
$$\{x + 1, x^2 + ix, x^2 - (i+1)x + i\}, \quad \{x + i, x^2 - x, x^2 - (i-1)x - i\}$$

all have $\delta = \pm(3i + 1)$, while their "conjugates", the four splittings

$$\{x - 1, x^2 + ix, x^2 - (i-1)x - i\}, \quad \{x + i, x^2 + x, x^2 - (i+1)x + i\},$$
$$\{x + 1, x^2 - ix, x^2 + (i-1)x - i\}, \quad \{x - i, x^2 - x, x^2 + (i+1)x + i\}$$

have $\delta = \pm(3i - 1)$.

Now, when $p = 5$, we may take $i = 2$ or $i = 3$. If $i = 2$ then $3i - 1 = 0$, so the last set of four become singular (and the penultimate set of four have $\delta = \pm 2$), while if $i = 3$ then $3i + 1 = 0$, so the penultimate set of four become singular (and then the last set of four have $\delta = \pm 2$). In either case, for $p = 5$ we have exactly four additional singular splittings, making ten in total; and we cannot have $i = 2$ or 3 in any other characteristic, so if $p \neq 5$ then there are only six singular splittings.

Appendix B. Hash function

The following is the Magma code that implements the hash function we described with the specific choices we have made.

```
function G2CGLhash(lambda, message)

splits := [ [{1,3},{2,5},{4,6}], [{1,3},{2,6},{4,5}],
[{1,4},{2,5},{3,6}], [{1,4},{2,6},{3,5}],
[{1,5},{2,3},{4,6}], [{1,5},{2,4},{3,6}],
[{1,6},{2,3},{4,5}], [{1,6},{2,4},{3,5}] ];

p:= 2^Ceiling(lambda*2/3); repeat p := NextPrime(p); until p mod 6 eq 5;
F<a> := GF(p^2);
R<x> := PolynomialRing(F);
factors := [x-1, x+1, x, x-2, x-1/2, 1];
mbase8 := []; message := message*2^30;
while message gt 0 do Append(~mbase8, message mod 8); message div:= 8; end while;

function fac(pol)
r := [ rt[1] : rt in Factorization(pol)];
if #r eq 1 then Append(~r,1); end if;
return r;
end function;

for i := 1 to #mbase8 do
  split := splits[mbase8[i]+1];
  G1 := &*[ factors[j] : j in split[1]];
  G2 := &*[ factors[j] : j in split[2]];
  G3 := &*[ factors[j] : j in split[3]];
  h1 := Derivative(G2)*G3 - G2*Derivative(G3); r1 := fac(h1);
  h2 := Derivative(G3)*G1 - G3*Derivative(G1); r2 := fac(h2);
  if Rank(Matrix(F, 2, 3, [ [Coefficient(h1,j) : j in [0..2]],
                            [Coefficient(h2,j) : j in [0..2]] ])) eq 1 then
    print "No hash for this value possible."; return 0;
  end if;
  h3 := Derivative(G1)*G2 - G1*Derivative(G2); r3 := fac(h3);
  factors := r1 cat r2 cat r3;
end for;

return G2Invariants(HyperellipticCurve(&*factors));

end function;
```

Section of Algebra, Department of Mathematics, KU Leuven
E-mail address: wouter.castryck@esat.kuleuven.be

imec-COSIC, Department of Electrical Engineering, KU Leuven
E-mail address: thomas.decru@esat.kuleuven.be

Inria *and* Laboratoire d'Informatique de l'École polytechnique, Université Paris–Saclay, Palaiseau, France
E-mail address: smith@lix.polytechnique.fr