An Efficient Key Recovery Attack on SIDH

Wouter Castryck[1,2(✉)] and Thomas Decru[1]

[1] imec-COSIC, KU Leuven, Leuven, Belgium
wouter.castryck@esat.kuleuven.be

[2] Vakgroep Wiskunde: Algebra en Meetkunde, Universiteit Gent, Ghent, Belgium
thomas.decru@esat.kuleuven.be

Abstract. We present an efficient key recovery attack on the Supersingular Isogeny Diffie–Hellman protocol (SIDH). The attack is based on Kani's "reducibility criterion" for isogenies from products of elliptic curves and strongly relies on the torsion point images that Alice and Bob exchange during the protocol. If we assume knowledge of the endomorphism ring of the starting curve then the classical running time is polynomial in the input size (heuristically), apart from the factorization of a small number of integers that only depend on the system parameters. The attack is particularly fast and easy to implement if one of the parties uses 2-isogenies and the starting curve comes equipped with a non-scalar endomorphism of very small degree; this is the case for SIKE, the instantiation of SIDH that recently advanced to the fourth round of NIST's standardization effort for post-quantum cryptography. Our Magma implementation breaks `SIKEp434`, which aims at security level 1, in about ten minutes on a single core.

Keywords: isogeny-based cryptography · SIDH · elliptic curves · genus 2 curves

1 Introduction

We present a new and powerful key recovery attack on the Supersingular Isogeny Diffie–Hellman key exchange protocol (SIDH), proposed in 2011 by Jao and De Feo [25] and considered the flagship of isogeny-based cryptography. Its instantiation SIKE [24] recently advanced to the fourth round of the post-quantum cryptography standardization process, currently run by NIST [33].

The attack is based on a "reducibility criterion" from 1997 due to Kani [26, Theorem 2.6] for determining whether an isogeny emanating from a product of two elliptic curves takes us again to a product of elliptic curves, rather than

This work was supported in part by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement ISOCRYPT - No. 101020788) and by CyberSecurity Research Flanders with reference number VR20192203.

C. Hazay and M. Stam (Eds.): EUROCRYPT 2023, LNCS 14008, pp. 423–447, 2023.
https://doi.org/10.1007/978-3-031-30589-4_15

to the Jacobian of a genus 2 curve as one would expect. This heavily outperforms previous attack strategies, such as the ones discussed in [11,31,36,37, Sect. 5], both in theory and in practice. Run on a single core, the appended Magma code [2] breaks the Microsoft SIKE challenges `$IKEp182` and `$IKEp217` from [32] in about 55 s and 85 s, respectively. A run on the `SIKEp434` parameters, previously believed to meet NIST's quantum security level 1, took roughly 10 m, again on a single core. We also ran the code on random instances of `SIKEp503` (level 2), `SIKEp610` (level 3) and `SIKEp751` (level 5), which on average took about 20 m, 55 m and 3 h 15 m, respectively.

For the sake of exposition, we concentrate on the concrete set-up of SIKE and comment on more general parameter choices as we see fit. Our attack targets Bob's private key, which is a secret 3^b-isogeny $\varphi : E_{\text{start}} \to E$ between two supersingular elliptic curves E_{start}, E. The starting curve E_{start} is a system parameter and is endowed with two independent points $P_0, Q_0 \in E_{\text{start}}$ of order 2^a; the exponents a, b are system parameters too. Bob's public key consists of the codomain E and the image points $\varphi(P_0), \varphi(Q_0)$. As explained in Sect. 4, it follows from Kani's criterion that for any isogeny $\gamma : E_{\text{start}} \to C$ of degree $2^a - 3^b$ (assume for now that this is positive) the $(2^a, 2^a)$-isogeny from $C \times E$ with kernel generated by $(\gamma(P_0), \varphi(P_0))$, $(\gamma(Q_0), \varphi(Q_0))$ must again land on a product of elliptic curves. The idea behind our attack is that landing on a product is extremely unlikely if $E, \varphi(P_0), \varphi(Q_0)$ do *not* constitute a valid public key triple. In other words, we can use Kani's criterion as a decision tool. An easy search-to-decision reduction then allows to recover φ. The details of this reduction can be found in Sect. 6.

The main bottleneck is finding and evaluating the auxiliary isogeny γ; once this is done, the decision algorithm amounts to computing a length-a chain of $(2, 2)$-isogenies, which is very efficient (Richelot isogenies). Our focus lies on the cases where E_{start} is one of

$$y^2 = x^3 + x, \qquad\qquad y^2 = x^3 + 6x^2 + x, \tag{1}$$

which are supersingular in characteristic $p \equiv 3 \bmod 4$. The former was the starting curve of SIKE when it was submitted to the first round of the NIST standardization effort. The use of the latter curve was proposed from the second round onwards. Both curves come equipped with an explicit endomorphism $2\mathbf{i}$ satisfying $2\mathbf{i} \circ 2\mathbf{i} = [-4]$. As discussed in Sect. 5, this feature often lends itself to a very simple construction of γ, apart from the cost of factoring $2^a - 3^b$ (precomputable). In practice, the success probability is high enough for setting our search-to-decision reduction in motion, where now polynomially many integers of size $O(2^a)$ must be factored; concretely, these integers are all of the form $2^{a-j} - 3^{b-i}$. As the reader can tell from the above timings, the resulting attack on SIKE is devastating.

While the endomorphism $2\mathbf{i}$ is sufficient for a practical break of SIKE in all security levels, the asymptotic time complexity is only sub-exponential; more precisely, modulo the said factorizations, we expect it to run in time $L_p(1/4)$, see Sect. 10. In order to reach a polynomial runtime (heuristically and again modulo factorization), one must also resort to non-scalar endomorphisms of other very small degrees. Such endomorphisms may not exist on E_{start}, but in view of the work of Love and Boneh [29] one can easily find explicit isogenies to curves on which they do occur. The KLPT algorithm from [27] then allows to transform a degree $2^a - 3^b$ isogeny emanating from such a curve into the desired instance of $\gamma : E_{\text{start}} \to C$. In fact, for this approach, it is not required that E_{start} is among (1): any starting curve whose endomorphism ring is known will do.

Remark 1. If the endomorphism ring of E_{start} is unknown, then one can still construct γ efficiently in case $2^a - 3^b$ happens to be smooth. This event is highly unlikely, but as explained in Sect. 11 one can create more leeway by extending φ and by guessing how it acts on small-order torsion; as was pointed out to us by De Feo and Wesolowski (independently), the resulting attack runs heuristically in time $L_p(1/2 + \epsilon)$.

We finally note that our attack also breaks instantiations of SIDH that make other torsion choices for Alice and Bob. Indeed, the strategy can be used for the recovery of a secret ℓ_{B}^b-isogeny from ℓ_{A}^a-torsion point information for any small primes $\ell_{\text{A}}, \ell_{\text{B}}$, as long as $\ell_{\text{B}}^b = O(\ell_{\text{A}}^a)$; in particular, when applied to SIKE this also allows to find Alice's private key. It can even handle non-prime-power torsion, as used in for example B-SIDH [9]. Our claims on the asymptotic runtime still apply, but away from $\ell_A = 2$ implementing the attack is more cumbersome because one can no longer rely on fast Richelot isogenies; see Sect. 11 for a more elaborate discussion.

Follow-Up Work

After a first version of this paper was posted online, several improvements and extensions have made an appearance; for the sake of chronology, the remainder of this paper is free of references to these follow-up works, but let us give a quick overview. It was observed by Maino and Martindale [30],[1] Oudompheng [34], Petit (personal communication) and Wesolowski [45] that Kani's machinery also allows for a direct key recovery, which is considerably faster than our decisional approach. Various other speed-ups were found in an effort led by Oudompheng and Pope to reimplement the attack in SageMath [35,40], and in a parallel effort by Steel to fine-tune our Magma implementation. Notably, the Magma kernel was updated with improved $\mathbb{F}_{p^2}$-arithmetic, resulting in a faster execution of our code (the initial timings were slower by factors 4 to 8, roughly). In the case of a starting curve with known endomorphism ring, Wesolowski rigorously proved, assuming the generalized Riemann hypothesis, that the auxiliary isogeny γ can be constructed in polynomial time, without any need for factorizations [45]. The most remarkable follow-up work is due to Robert [38], who showed how to get rid, unconditionally, of all endomorphism ring assumptions by working with abelian eightfolds rather than surfaces (using an idea that is reminiscent of the Zarhin

[1] Right before posting our paper online, we learned that the authors of [30] had started pursuing related ideas.

trick). He also crushed the hope for secure higher-dimensional variants of SIDH. Fouotsa, Moriya and Petit have proposed an interesting (yet impractical) variant of SIDH that aims at thwarting the current attacks [18].

2 Impact and Non-impact on Isogeny-Based Cryptosystems

Our attack also impacts various cryptographic schemes that build on SIDH, or make use of similar hardness assumptions, such as B-SIDH [9], SHealS [19] and k-SIDH [1]. As discussed in Sect. 11, even in the case of a starting curve with unknown endomorphism ring, our attack lowers the security of all these schemes. Here, an interesting target is Séta [13], which allows much leeway for an attacker, coming from largely imbalanced torsion levels.[2] On the other hand, we stress that the attack relies crucially on the torsion point images exchanged by Alice and Bob, as well as on the knowledge of the degree of the secret isogeny. In particular, it cannot be adjusted in an obvious way to attack primitives that do not reveal this information, such as CRS/CSIDH [7,10,39] and SQISign [12], and the general supersingular isogeny path problem remains unaffected [44]. We forward the reader to an online project, initiated by De Feo, which attempts at organizing the most popular isogeny-based cryptographic protocols and their best classical and quantum attacks [14].

3 Concrete Set-Up

Concretely, we will describe an algorithm which, upon input of

(i) an SIDH prime p, i.e., $p = 2^a 3^b f - 1$ for integers $a \geq 2$, $b, f \geq 1$ with $2^a \approx 3^b$,
(ii) an elliptic curve $E_0/\mathbb{F}_{p^2}$ with $\#E_0(\mathbb{F}_{p^2}) = (p+1)^2$,
(iii) generators P_0, Q_0 of $E_0[2^a]$,
(iv) a 3^β-isogeny $\tau : E_0 \to E_{\text{start}}$ for some $\beta \geq 0$, where E_{start} is one of the two curves (1) that have served as starting curves in SIKE,
(v) the codomain $E/\mathbb{F}_{p^2}$ of a secret cyclic 3^b-isogeny $\varphi : E_0 \to E$,
(vi) the generators $P = \varphi(P_0)$ and $Q = \varphi(Q_0)$ of $E[2^a]$,

returns the isogeny φ. For simplicity we assume that φ is uniquely determined, which is true with overwhelming probability. If $2^{a-1} > 3^{b/2}$ then this is guaranteed by [43, Lemma 3.1]. A note on input (iv): when attacking SIKE, at the initial stage we will have $\beta = 0$ and $E_0 = E_{\text{start}}$, so the reader can keep this setting in mind for now. But our search-to-decision reduction will involve a recursion during which the value of β will grow, whence this more general formulation. Moreover, we also want to cope with larger values of β when discussing other starting curves with a known endomorphism ring.

[2] Séta is now fully broken in view of Robert's work [38].

4 Decision via Kani's Reducibility Criterion

We first study the following decision variant: we assume to be given (i), (ii), (iii) and an elliptic curve $E/\mathbb{F}_{p^2}$ satisfying $\#E(\mathbb{F}_{p^2}) = (p+1)^2$, along with generators P, Q of $E[2^a]$. The goal is to decide whether or not

(D) there is a 3^b-isogeny $\varphi : E_0 \to E$ such that $\varphi(P_0) = P$ and $\varphi(Q_0) = Q$.

We impose two technical conditions that will be discussed in more detail later on:

- We suppose that $2^a > 3^b$.
- Let $c = 2^a - 3^b$. We assume that we can compute the images $P_c = \gamma(P_0)$ and $Q_c = \gamma(Q_0)$ under an arbitrary c-isogeny $\gamma : E_0 \to C$ to some codomain curve C.

We let $x \in \mathbb{Z}$ denote a multiplicative inverse of 3^b modulo 2^a. Note that $-x$ is then a multiplicative inverse of c modulo 2^a.

4.1 $(2^a, 2^a)$-Subgroups Built from Torsion Point Information

If (D) holds then we can consider the isogeny

$$\psi = [-1] \circ \varphi \circ \hat{\gamma} : C \to E,$$

where we note that $\psi(P_c) = -cP$ and $\psi(Q_c) = -cQ$. For all $R, S \in C[2^a]$ we have that

$$e_{2^a}(x\psi(R), x\psi(S)) = e_{2^a}(R, S)^{x^2 c 3^b} = e_{2^a}(R, S)^{-1}$$

or in other words the group homomorphism

$$[x] \circ \psi|_{C[2^a]} : C[2^a] \to E[2^a]$$

is a so-called "anti-isometry" with respect to the 2^a-Weil pairing. This implies that the group

$$\langle (P_c, x\psi(P_c)), (Q_c, x\psi(Q_c)) \rangle = \langle (P_c, P), (Q_c, Q) \rangle \qquad (2)$$

is maximally isotropic with respect to the 2^a-Weil pairing on the product $C \times E$ (equipped with the product polarization). Indeed,

$$e_{2^a}((P_c, x\psi(P_c)), (Q_c, x\psi(Q_c))) = e_{2^a}(P_c, Q_c) e_{2^a}(x\psi(P_c), x\psi(Q_c)) = 1$$

because the Weil pairing on $C \times E$ is just the product of the Weil pairings of the corresponding components.

Therefore it concerns the kernel of a $(2^a, 2^a)$-isogeny of principally polarized abelian surfaces. By writing this isogeny as a composition of $(2, 2)$-isogenies, it can be viewed as a walk of length a in the $(2, 2)$-isogeny graph of superspecial principally polarized abelian surfaces over $\overline{\mathbb{F}}_p$, all of whose vertices are defined over $\mathbb{F}_{p^2}$. These vertices come in two types: about $p^2/288$ products of supersingular elliptic curves and about $p^3/2880$ Jacobians of superspecial genus 2 curves, see e.g. [3]. Therefore it is to be expected that most isogenies in the chain are between Jacobians of genus 2 curves, and such isogenies can be computed efficiently using “classical” formulae due to Richelot [42]. But the first step is clearly an exception to this: with overwhelming probability, this is a “gluing” step, mapping the product $C \times E$ to a Jacobian (more precisely, by Theorem 1 below this can only fail if $C \cong E$). Formulae for this gluing step were derived in [23] and are recalled in Sect. 8.

4.2 Kani's Theorem

What is the role of the isogeny γ in all this? Its aim is to force us into the exceptional situation where the *last* step of the chain is split, i.e., the codomain of our $(2^a, 2^a)$-isogeny is again a product of elliptic curves. In that case the anti-isometry $x\psi|_{C[2^a]}$ and the group (2) are called “reducible”. This event is characterized by the theorem of Kani [26, Theorem 2.6]:

Definition 1. *Let C, E be two elliptic curves and $N \geq 2$ an integer. Let $\psi : C \to E$ be a separable isogeny and let $H_1, H_2 \subset \ker \psi$ be subgroups such that $H_1 \cap H_2 = \{0\}$, $\#H_1 \cdot \#H_2 = \deg \psi$ and $\#H_1 + \#H_2 = N$. Then the triplet (ψ, H_1, H_2) is called an* isogeny diamond configuration of order N *between C and E.*

Theorem 1. *Let (ψ, H_1, H_2) be an isogeny diamond configuration of order $N \geq 2$ between two elliptic curves C and E. Let $d = \gcd(\#H_1, \#H_2)$, let $n = N/d$ and let $k_i = \#H_i/d$ for $i = 1, 2$. Then ψ factors uniquely over $[d]$, i.e. $\psi = \psi' \circ [d]$ and there is a unique reducible anti-isometry $\iota : C[N] \to E[N]$ such that*

$$\iota(k_1 R_1 + k_2 R_2) = \psi'(R_2 - R_1) \text{ for all } R_i \in [n]^{-1} H_i \ (i = 1, 2) \ . \tag{3}$$

Moreover, if $N \leq p$ then every reducible anti-isometry $C[N] \to E[N]$ is of this form.

Remark 2. Kani allows for inseparable isogenies in Definition 1, in which case $\#H_i$ should be interpreted as the *degree* of the corresponding subgroup scheme. When doing so, the condition $N \leq p$ in Theorem 1 can be discarded; this was merely added to ensure that ψ is separable.

In our case, the kernel of ψ is a group of order $c3^b$, so it admits two (unique) subgroups H_1, H_2 of respective orders c and 3^b. We clearly have that $H_1 \cap H_2 = \{0\}$ and

$$\#H_1 + \#H_2 = 2^a, \quad \#H_1 \cdot \#H_2 = \deg \psi,$$

so the triplet (ψ, H_1, H_2) is an isogeny diamond configuration of order 2^a. Then Kani's theorem implies that our anti-isometry $x\psi|_{C[2^a]}$ is reducible. Indeed, let us check condition (3) explicitly: we need to verify that

$$x\psi(cR_1 + 3^b R_2) = \psi(R_2 - R_1)$$

for all points R_1, R_2 such that $2^a R_1 \in H_1$ and $2^a R_2 \in H_2$ (note that $d = 1$ in our case). But this is easy: since $\psi(R_1)$ and $\psi(R_2)$ are 2^a-torsion points, we can rewrite the left hand side as

$$\begin{aligned} xc\psi(R_1) + x3^b\psi(R_2) &= 3^{-b}(2^a - 3^b)\psi(R_1) + 3^{-b}3^b\psi(R_2) \\ &= \psi(R_2) - \psi(R_1) \\ &= \psi(R_2 - R_1) \end{aligned}$$

as wanted (recall that $x3^b \equiv -xc \equiv 1 \bmod 2^a$).

4.3 Decision Strategy

Our decision strategy amounts to testing whether or not quotienting out $C \times E$ by (2) takes us to a product of elliptic curves, as depicted in Fig. 1. As we have just argued, if (D) holds, then we pass the test.

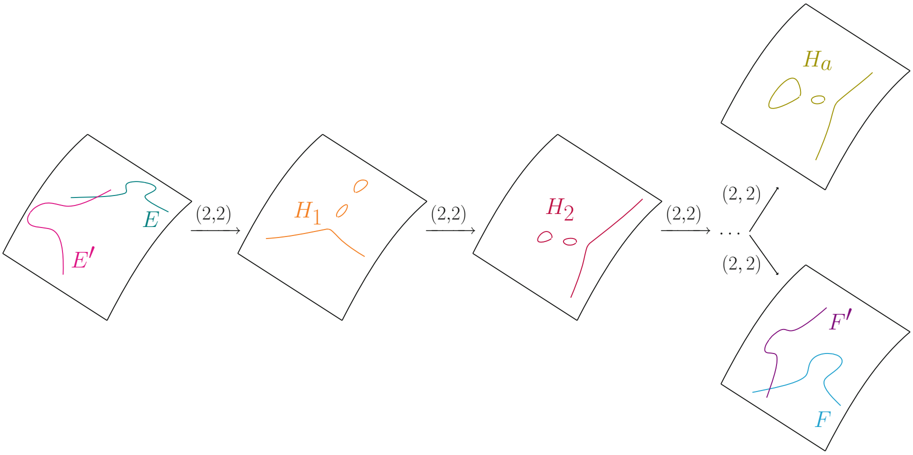

Fig. 1. Decision strategy based on Kani's reducibility criterion.

For now, we content ourselves with the loose heuristic that if (D) does not hold, then the test should fail with overwhelming probability because the proportion of products of elliptic curves among all vertices in the graph is only about $10/p$. We can actually be a bit more precise about this heuristic in the cases that are relevant for our attack, namely the "wrong guesses" in our search-to-decision reduction: see Remark 4.

5 Constructing and Evaluating the Auxiliary Isogeny γ

The assumption that we can (efficiently) compute the image points P_c and Q_c under a degree-c isogeny is non-trivial, and this is where we need the factorization of $c = 2^a - 3^b$. It is also here that we rely on the special nature of E_{start}: both options come with an endomorphism $2\mathbf{i}$ satisfying $2\mathbf{i} \circ 2\mathbf{i} = [-4]$. Indeed, on $E_{\text{start}} : y^2 = x^3 + x$ we have the automorphism $\mathbf{i} : (x, y) \mapsto (-x, \sqrt{-1}y)$ and we simply let $2\mathbf{i} = [2] \circ \mathbf{i}$. For $E_{\text{start}} : y^2 = x^3 + 6x^2 + x$ we can obtain $2\mathbf{i}$ as the composition of its outgoing 2-isogeny to $y^2 = x^3 + x$, the automorphism $\mathbf{i}$ on the latter curve, and the dual of the said 2-isogeny.

5.1 Construction

There is a reasonable chance that the prime factorization of c only involves prime factors that are congruent to 1 mod 4; this chance is inversely proportional to $\sqrt{a}$ by a theorem of Landau (see Sect. 10 for a more detailed discussion). As far as we are aware, the only known way to find out is by factoring c explicitly. Once this factorization is done and all prime factors are indeed congruent to 1 mod 4, we can efficiently write $c = u^2 + 4v^2 = (u + 2\mathbf{i}v)(u - 2\mathbf{i}v)$. Then

$$\gamma_{\text{start}} = [u] + [v] \circ 2\mathbf{i}$$

is an easy-to-evaluate degree-c endomorphism of E_{start}.

Remark 3. The method for finding u and v is classical: e.g., in the squarefree case, one computes

$$\prod_{\text{primes } \ell | c} \gcd(z_\ell + \mathbf{i}, \ell)$$

using Euclid's algorithm over the Gaussian integers; here z_ℓ is any integer such that $z_\ell^2 \equiv -1 \bmod \ell$. The outcome is among $\pm(u + 2\mathbf{i}v), \pm\mathbf{i}(u + 2\mathbf{i}v)$.

Then in order to find γ, we use the isogeny τ from input (iv). Let $\tilde{\tau} : E_{\text{start}} \to C$ be the isogeny with kernel $\gamma_{\text{start}}(\tau(E_0[3^\beta])) = \gamma_{\text{start}}(\ker \hat{\tau})$. Then $\tilde{\tau} \circ \gamma_{\text{start}} \circ \tau : E_0 \to C$ is a $3^{2\beta}c$-isogeny vanishing on $E_0[3^\beta]$, so it factors over $[3^\beta]$ and we can let

$$\gamma = \frac{\tilde{\tau} \circ \gamma_{\text{start}} \circ \tau}{3^\beta}.$$

It remains to see that γ is easy to evaluate on our 2^a-torsion points P_0 and Q_0. For this, we first discuss a special case.

5.2 Evaluation: Case $\beta \leq b$

This is the only relevant case when attacking SIDH with base curve $E_0 = E_{\text{start}}$, as in the case of SIKE: while β will grow during our search-to-decision reduction, it will never grow beyond b. But then we always have that $\ker \hat{\tau} \subset E_{\text{start}}[3^b] \subset E_{\text{start}}(\mathbb{F}_{p^2})$. So we can explicitly write down a generator $T \in E_{\text{start}}(\mathbb{F}_{p^2})$ of $\ker \hat{\tau}$ and compute the isogeny $\tilde{\tau}$ with kernel $\langle \gamma_{\text{start}}(T) \rangle$. Evaluating γ in our 2^a-torsion points P_0 and Q_0 is then simply done by feeding them to $\tilde{\tau} \circ \gamma_{\text{start}} \circ \tau$ and scalar-multiplying the outcome with a multiplicative inverse of 3^β modulo 2^a. (In fact, this evaluation will naturally simplify in the context of our search-to-decision reduction.)

5.3 Evaluation: General Case

If $\beta > b$ then we cannot simply evaluate γ_{start} in a generator of $\ker \hat{\tau}$, unless we base-change to a potentially very large and costly extension of $\mathbb{F}_{p^2}$. But note that the isogeny $\tilde{\tau}$ is precisely the pushforward isogeny $[\gamma_{\text{start}}]_* \hat{\tau}$ that was studied in [12, Sect. 4]. This suggests the following alternative method for computing $\tilde{\tau}$. Note that the specific choice of E_{start} comes with an explicit isomorphism

$$\iota : \operatorname{End}(E_{\text{start}}) \to \mathcal{O}_{\text{start}}$$

where $\mathcal{O}_{\text{start}}$ is a maximal order in the quaternion algebra $B_{p,\infty} = \langle 1, \mathbf{i}, \mathbf{j}, \mathbf{ij} \rangle_{\mathbb{Q}}$ with $\mathbf{i}^2 = -1$ and $\mathbf{j}^2 = -p$. Then:

1. First, one converts the isogeny $\hat{\tau} : E_{\text{start}} \to E_0$ into a left ideal $I_{\hat{\tau}} \subset \mathcal{O}_{\text{start}}$ of norm 3^{β}, e.g. following [20, Algorithm 3]. In fact, in the main use cases of this general method, a large component of the isogeny $\hat{\tau}$ will arise *from* its corresponding left $\mathcal{O}_{\text{start}}$-ideal; so in those cases this step can be simplified.
2. Next, one computes the left ideal $I_{\tilde{\tau}} = [(\iota(\gamma_{\text{start}}))]_* I_{\hat{\tau}}$ using the formula from [12, Lemma 3]; this ideal again has norm 3^{β}.
3. Finally, one converts the ideal $I_{\tilde{\tau}}$ into a length-β chain of 3-isogenies emanating from E_{start}, e.g. using [20, Algorithm 2]. Then $\tilde{\tau}$ is the composition of these 3-isogenies.

Then, here too, evaluating γ in P_0 and Q_0 is done by applying $\tilde{\tau} \circ \gamma_{\text{start}} \circ \tau$ and scalar-multiplying with an inverse of 3^{β} modulo 2^a.

5.4 Away from the Endomorphism 2i

We conclude by remarking that there are many other candidate-ways for constructing the isogeny γ. Just to give one similar example, decompositions of the form $c = u^2 + 3v^2$ are useful as soon as one knows an explicit path to $y^2 = x^3 + 1$, because this curve comes equipped with an endomorphism ω such that $\omega^2 = -3$. This type of examples will reappear in Sect. 10. A different kind of example is the case where c is very smooth: in that case one can construct the desired c-isogeny $\gamma : E_0 \to C$ as a composition of small degree isogenies *without* knowing a path to some special-featured curve. Even though this event is highly unlikely, there are tricks to create more leeway; see Sect. 11 for a more elaborate discussion.

6 Key Recovery Algorithm: Basic Version

We resume with the set-up from Sect. 3. The previous sections suggest the following iterative approach to full key recovery. We assume for simplicity that $\beta = 0$, so that the base curve E_0 coincides with E_{start}. Recall that this is the case in SIKE. In the general case, one should just replace the maps $\hat{\kappa}_1 : E_1 \to E_0$, $\widehat{\kappa_2 \kappa_1} : E_2 \to E_0$, ... below with their compositions with τ.

6.1 Iteration

For the first iteration, choose $\beta_1 \geq 1$ minimal such that there exists some $\alpha_1 \geq 0$ for which

$$c_1 = 2^{a-\alpha_1} - 3^{b-\beta_1}$$

is of the form $u_1^2 + 4v_1^2$. Write $\varphi = \varphi_1 \circ \kappa_1$ with κ_1 a 3^{β_1}-isogeny. To an attacker, there are a priori 3^{β_1} options for κ_1 (this assumes knowledge of an "incoming isogeny", otherwise there are $4 \cdot 3^{\beta_1 - 1}$ options). For each of these options, we can run our decision algorithm on

(ii) the curve $E_1 = \kappa_1(E_0)$,
(iii) the generators $P_1 = \kappa_1(2^{\alpha_1} P_0)$ and $Q_1 = \kappa_1(2^{\alpha_1} Q_0)$ of $E_1[2^{a-\alpha_1}]$,
(iv) the 3^{β_1}-isogeny $\hat{\kappa}_1 : E_1 \to E_0$,
(v) the codomain E; if the guess is correct then it is connected to E_1 via the unknown isogeny φ_1 of degree $3^{b-\beta_1}$,
(vi) the generators $2^{\alpha_1} P, 2^{\alpha_1} Q$ of $E[2^{a-\alpha_1}]$

where the numbering (ii)–(vi) is chosen to be consistent with our set-up from Sect. 3. According to our heuristic assumption discussed in Sect. 4.3, we expect that only the correct guess for κ_1 will pass the test; see also Remark 4 below.

Let us discuss in more detail what "running the test" amounts to in this case. First, one must compute the images P_{c_1}, Q_{c_1} of P_1, Q_1 under the isogeny

$$\gamma_1 = \frac{\tilde{\hat{\kappa}}_1 \circ \gamma_{\text{start}} \circ \hat{\kappa}_1}{3^{\beta_1}} \tag{4}$$

where $\tilde{\hat{\kappa}}_1 : E_{\text{start}} \to C_1$ is the isogeny with kernel $\gamma_{\text{start}}(\ker \kappa_1)$, with $\gamma_{\text{start}} = [u_1] + 2\mathbf{i} \circ [v_1]$. Observe that this simplifies: all one should do is compute

$$P_{c_1} = 2^{\alpha_1} \tilde{\hat{\kappa}}_1 \gamma_{\text{start}}(P_0), \quad Q_{c_1} = 2^{\alpha_1} \tilde{\hat{\kappa}}_1 \gamma_{\text{start}}(Q_0). \tag{5}$$

Once these points have been computed, one checks whether the quotient of $C_1 \times E$ by the $(2^{a-\alpha_1}, 2^{a-\alpha_1})$-subgroup

$$\langle (P_{c_1}, 2^{\alpha_1} P), (Q_{c_1}, 2^{\alpha_1} Q) \rangle \tag{6}$$

is again a product of elliptic curves. This is done by computing the corresponding chain of $(2, 2)$-isogenies. With overwhelming probability, the first $a - \alpha_1 - 1$ steps in this chain amount to one gluing step followed by $a - \alpha_1 - 2$ Richelot isogenies between Jacobians of genus 2 curves. An easy "$\delta = 0$ test" then checks whether or not the last step splits (see Sect. 8 for algorithmic details).

If the test fails, then we try again with a different guess for κ_1. We remark that, even in the case of a wrong guess, the subgroup (6) is always maximally isotropic with respect to the Weil pairing, so this is *not* the way in which one can detect having taken the wrong direction: one really has to perform the gluing and its successive Richelot walk. (The failure of detecting wrong steps using the Weil pairing is well-known, see e.g. [21, Sect. 7.2]; with some imagination, our attack can be viewed as a refinement of this approach.) If the test passes, then very likely we have found the correct instance of κ_1.

Remark 4. If a wrong guess for κ_1 passes the test, then in view of Kani's theorem the points P_{c_1}, Q_{c_1} must be connected to $2^{\alpha_1}P$, $2^{\alpha_1}Q$ via an anti-isometry coming from an isogeny $\psi : C_1 \to E$ fitting in an isogeny diamond configuration of order $2^{a-\alpha_1}$. It is easy to see that the natural candidate for ψ, namely the degree $3^{b+\beta_1}(2^{a-\alpha_1} - 3^{b-\beta_1})$-isogeny

$$\varphi \circ \hat{\kappa}_1 \circ \hat{\gamma}_1 : C_1 \to E,$$

does *not* fit in such an isogeny diamond. Indeed, if it would, then we would have

$$3^{b+\beta_1}(2^{a-\alpha_1} - 3^{b-\beta_1}) = k(2^{a-\alpha_1} - k) \tag{7}$$

for some natural number

$$k \in [1, 2^{a-\alpha_1} - 1]. \tag{8}$$

Modulo $2^{a-\alpha_1}$ the Eq. (7) implies $3^{2b} \equiv k^2$, so that k is congruent to one of

$$3^b, \quad -3^b, \quad 3^b + 2^{a-\alpha_1-1}, \quad -3^b + 2^{a-\alpha_1-1}.$$

In particular, k and $2^{a-\alpha_1} - k$ must be of the form $\pm 3^b + \lambda 2^{a-\alpha_1-1}$. On the other hand, (7) implies that either k or $2^{a-\alpha_1} - k$ is divisible by $3^{b+\beta_1}$. This can only happen if the corresponding λ is non-zero and divisible by 3^b, but then (unless we are in the trivial boundary case $\alpha_1 = a, \beta_1 = b$) we necessarily fall outside the interval (8): a contradiction.

Remark 5. We did not manage to fully rule out the existence of instances of ψ other than $\varphi \circ \hat{\kappa}_1 \circ \hat{\gamma}_1$. However, at least heuristically, the odds are strongly against this. Indeed, loosely speaking, these instances would need to act on $C_1[2^{a-\alpha_1}]$ in essentially the same way as $\varphi \circ \hat{\kappa}_1 \circ \hat{\gamma}_1$ does, and a variation on [43, Lemma 3.1] shows that there is typically no room for another such isogeny.

Once we have found the correct κ_1 we continue from E_1. That is, we let $\beta_2 > \beta_1$ be minimal such that there is some $\alpha_2 \geq 0$ for which $c_2 = 2^{a-\alpha_2} - 3^{b-\beta_2}$ is of the form $u_2^2 + 4v_2^2$. Now one tries to recover the $3^{\beta_2-\beta_1}$-component $\kappa_2 : E_1 \to E_2$ such that $\varphi_1 = \varphi_2 \circ \kappa_2$. In this case, for each guess for κ_2 one computes

$$P_{c_2} = 2^{\alpha_2} \widetilde{\widehat{\kappa_2\kappa_1}} \gamma_{\text{start}}(P_0), \quad Q_{c_2} = 2^{\alpha_2} \widetilde{\widehat{\kappa_2\kappa_1}} \gamma_{\text{start}}(Q_0)$$

with $\widetilde{\widehat{\kappa_2\kappa_1}} : E_{\text{start}} \to C_2$ the isogeny with kernel $\gamma_{\text{start}}(\ker \kappa_2\kappa_1)$ and $\gamma_{\text{start}} = [u_2] + 2\mathbf{i} \circ [v_2]$. One then checks whether

$$\langle (P_{c_2}, 2^{\alpha_2}P), (Q_{c_2}, 2^{\alpha_2}Q) \rangle \ \subset \ C_2 \times E$$

is reducible or not. By continuing in this way, one eventually retrieves all of φ.

6.2 Step Sizes

The gaps between the consecutive integers $0, \beta_1, \beta_2, \beta_3, \ldots, \beta_r = b$ should be as small as possible, because this reduces the number of possible guesses in each iteration. More concretely, the expected number of $(2,2)$-chains that need to be computed is

$$\frac{1}{2}\left(3^{\beta_1} + 3^{\beta_2-\beta_1} + 3^{\beta_3-\beta_2} + \ldots + 3^{b-\beta_{r-1}}\right). \tag{9}$$

A necessary condition on each β_i is that $b - \beta_i$ is odd, except in the last iteration where we have $\beta_r = b$. Indeed, if $b - \beta_i > 0$ is even then

$$c_i = 2^{a-\alpha_i} - 3^{b-\beta_i} \equiv 3 \bmod 4$$

cannot be of the form $u_i^2 + 4v_i^2$. Therefore the best we can hope is that the sequence grows by steps of two, in which case the estimate (9) becomes about $9b/4$. Asymptotically, this hope is too good to be true, but for the concrete SIKE parameters experiment shows that this optimal estimate lies close to reality, with the only exceptions corresponding to small β_i. This makes sense: as β_i grows, the amount of leeway (i.e., the number of candidate α_i's) grows as well, and moreover the probability of success increases as c_i is allowed to get smaller. Example: for the parameters of `SIKEp434` where we have $a = 216$ and $b = 137$, one quickly finds suitable α_i for every even β_i in $\{0, 1, \ldots, b\} \setminus \{4\}$.

6.3 Rephrasing in Terms of Bob's Secret Key

In practice, SIDH comes with public generators $P_{\mathrm{Bob}}, Q_{\mathrm{Bob}}$ of $E_0[3^b]$ and Bob's secret isogeny φ is encoded as the integer

$$\mathrm{sk}_{\mathrm{Bob}} \in [0, 3^b)$$

for which $\ker\varphi = \langle P_{\mathrm{Bob}} + \mathrm{sk}_{\mathrm{Bob}}\, Q_{\mathrm{Bob}}\rangle$. Upon expanding

$$\mathrm{sk}_{\mathrm{Bob}} = k_1 + k_2 3^{\beta_1} + \ldots + k_r 3^{\beta_{r-1}}, \qquad k_i \in [0, 3^{\beta_i - \beta_{i-1}} - 1)$$

(where we let $\beta_0 = 0$), we observe that

$$\ker\kappa_1 = \langle 3^{b-\beta_1} P_{\mathrm{Bob}} + k_1 3^{b-\beta_1} Q_{\mathrm{Bob}}\rangle. \tag{10}$$

So the first iteration amounts to

- guessing k_1,
- determining the 3^{β_1}-isogeny $\tilde{\kappa}_1 : E_{\mathrm{start}} \to C_1$ with kernel $\gamma_{\mathrm{start}}(\ker\kappa_1)$, with $\ker\kappa_1$ as in (10),
- computing the points $P_{c_1}, Q_{c_1} \in C_1$ as in (5),
- checking whether or not the subgroup (6) is reducible.

After finding k_1, we proceed with

$$\ker\kappa_2 = \langle 3^{b-\beta_2} P_{\mathrm{Bob}} + (k_1 + k_2 3^{\beta_1}) 3^{b-\beta_2} Q_{\mathrm{Bob}}\rangle$$

in order to determine k_2 via trial-and-error, and so on. So the attack determines $\mathrm{sk}_{\mathrm{Bob}}$ digit by digit. If all the gaps are of size two, then this amounts to determining one base-9 digit of $\mathrm{sk}_{\mathrm{Bob}}$ at a time.

6.4 Walking Backwards

As was pointed out to us by De Feo, it may be simpler to reconstruct Bob's secret isogeny φ starting from its tail. That is: using the same $c_1 = 2^{a-\alpha_1} - 3^{b-\beta_1}$, one instead writes $\varphi = \kappa_1 \circ \varphi_1$ and one makes a guess for $\hat{\kappa}_1$. Now writing

$$E_1 = \hat{\kappa}_1(E),\ P_1 = \hat{\kappa}_1(2^{\alpha_1} P),\ Q_1 = \hat{\kappa}_1(2^{\alpha_1} Q),$$

letting γ_{start} be our degree-c_1 endomorphism on $E_0 = E_{\text{start}}$, and writing

$$P_{c_1} = 2^{\alpha_1}\gamma_{\text{start}}(P_0),\ Q_{c_1} = 2^{\alpha_1}\gamma_{\text{start}}(Q_0),$$

one now should check whether the subgroup

$$\langle (P_{c_1}, yP_1), (Q_{c_1}, yQ_1)\rangle \subset E_0 \times E_1$$

is reducible, with y a multiplicative inverse of 3^{β_1} modulo $2^{a-\alpha_1}$. The advantage of this approach is that one can work (and keep working throughout the iteration) with γ_{start} directly, i.e., one avoids the need for transformations of the kind (4).

7 Speed-Ups

We can speed up key recovery as follows:

7.1 Take α_i as Large as Possible

If for a given β_i there indeed exists some $\alpha_i \geq 0$ such that $c_i = 2^{a-\alpha_i} - 3^{b-\beta_i}$ is positive and free of prime factors congruent to 3 mod 4, then usually α_i is not the unique integer with that property, so there is some freedom. The larger we choose α_i, the smaller will be the length $a - \alpha_i$ of our chain of $(2, 2)$-isogenies. Therefore, it is more efficient to take larger α_i's.

7.2 Use a Precomputed Table

We have precomputed a table which for all $s \in \{1, 3, 5, \ldots, 239\}$ stores the smallest integer $t(s)$ such that $2^{t(s)} - 3^s$ is a product of primes congruent to 1 modulo 4. It also stores corresponding values for u and v. The table is available as `uvtable.m` and can be used as follows: for every candidate-β_i such that $b - \beta_i$ is odd, one checks whether or not $t(b - \beta_i) \leq a$. If not, then we proceed to the next candidate. If yes, then we can use this instance of β_i, and we choose $a - t(b - \beta_i)$ as a corresponding value for α_i. This makes sure that α_i is as large as possible, and moreover we have u_i, v_i readily available, without the need for factoring. Our table is sufficiently large to be used for each of the proposed parameter sets for SIKE, up to `SIKEp751` targeting NIST's security level 5.

7.3 Extend Bob's Secret Isogeny Where Useful

Imagine that some candidate-β_i does not admit an integer $\alpha_i \geq 0$ such that $2^{a-\alpha_i} - 3^{b-\beta_i}$ is a product of primes congruent to 1 mod 4 (e.g., because $b-\beta_i > 0$ is even). But imagine that $\beta_i - 1$ does. Then one can prolong Bob's secret isogeny with an arbitrary 3-isogeny φ' and let $P' = \varphi'(P)$ and $Q' = \varphi'(Q)$. Treating $\varphi' \circ \varphi$ as the new secret isogeny, the relevant expression now becomes $2^{a-\alpha_i} - 3^{b+1-\beta_i}$, and we know that there exists some $\alpha_i \geq 0$ for which this *is* a product of primes congruent to 1 mod 4. We can now use our attack to determine Bob's secret key modulo 3^{β_i} and proceed.

In practice, this means that most step sizes drop from 2 to 1, or in other words that we are determining one base-3 digit of $\mathrm{sk}_{\mathrm{Bob}}$ at a time. The only possibly larger step occurs at the beginning of the iteration. For instance, in the case of `SIKEp751`, the smallest β_1 such that $2^a - 3^{b-\beta_1} > 0$ is $\beta_1 = 6$, so we cannot hope for a smaller first gap. This implies a rather costly start of the algorithm: of the 3 h15 m that we spent on breaking `SIKEp751`, almost 2 h were needed for determining the first 6 out of 239 ternary digits of $\mathrm{sk}_{\mathrm{Bob}}$.

Remark 6. If 2^a is considerably smaller than 3^b, then it probably makes more sense to attack Alice's private key instead of Bob's, using chains of $(3, 3)$-isogenies; see Sect. 11. Of course, if 2^a gets much smaller than 3^b, then one enters the regime of the torsion-point attacks from [36,37].

Remark 7. There is a 1/4 probability that the random isogeny φ' matches with the dual of the last degree-3 component of φ. In this case, the wrong guesses are also at distance $3^{b-\beta_i}$ from E, so this creates false positives, leaving us clueless about which is the correct guess. However, this is easy to fix: if multiple guesses pass the test, then all one needs to do is change φ', and then we have identified the dual direction once and for all. If this happens, then it will be discovered when trying to determine the ternary digit at position $\beta_2 = \beta_1 + 1$ (and this does not affect the correctness of the first β_1 digits, as these were determined without the use of φ').

8 Computing Chains of (2, 2)-isogenies

In this section we explain how to determine whether or not a $(2^a, 2^a)$-subgroup $\langle (P_c, P), (Q_c, Q) \rangle$ of a product of elliptic curves $C \times E$ is reducible. Throughout, we avoid dealing with certain exceptional cases, e.g. every genus 2 curve $H : y^2 = h(x) = c_6x^6 + c_5x^5 + \ldots + c_0$ encountered is assumed to satisfy $c_6 \neq 0$, so that it has two places ∞_1, ∞_2 at infinity, and all points on its Jacobian J_H that we deal with are assumed to be representable as $(\alpha_1, \beta_1) + (\alpha_2, \beta_2) - \infty_1 - \infty_2$ with $\alpha_1 \neq \alpha_2$, so that they have a Mumford representation of the form $[x^2 + u_1x + u_0, v_1x + v_0]$. Moreover, all our chains of $(2, 2)$-isogenies are assumed to start off by gluing $C \times E$ into a Jacobian, after which we never run into a product of elliptic curves again, except possibly at the a-th and last step. The exceptions to these assumptions are expected to occur with probability $O(p^{-1})$, so we see no need to discuss nor implement them.

8.1 Gluing Elliptic Curves into a Jacobian

In the first step we want to glue the curves C and E into the Jacobian of a genus 2 curve H via the $(2,2)$-subgroup $\langle(2^{a-1}P_c, 2^{a-1}P), (2^{a-1}Q_c, 2^{a-1}Q)\rangle$. We also need to push the points (P_c, P), (Q_c, Q) through the corresponding isogeny. The relevant equations are as follows. We refer to [23, Proposition 4] and its proof for further details.

Proposition 1. *Let $C/K : y^2 = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3)$ and $E : y^2 = (x-\beta_1)(x-\beta_2)(x-\beta_3)$ be elliptic curves over a field K of characteristic different from two. Write Δ_α for the discriminant of $(x-\alpha_1)(x-\alpha_2)(x-\alpha_3)$ and Δ_β for the discriminant of $(x-\beta_1)(x-\beta_2)(x-\beta_3)$. Furthermore, define*

$$\begin{aligned}
a_1 &= (\alpha_3-\alpha_2)^2/(\beta_3-\beta_2) + (\alpha_2-\alpha_1)^2/(\beta_2-\beta_1) + (\alpha_1-\alpha_3)^2/(\beta_1-\beta_3),\\
b_1 &= (\beta_3-\beta_2)^2/(\alpha_3-\alpha_2) + (\beta_2-\beta_1)^2/(\alpha_2-\alpha_1) + (\beta_1-\beta_3)^2/(\alpha_1-\alpha_3),\\
a_2 &= \alpha_1(\beta_3-\beta_2) + \alpha_2(\beta_1-\beta_3) + \alpha_3(\beta_2-\beta_1),\\
b_2 &= \beta_1(\alpha_3-\alpha_2) + \beta_2(\alpha_1-\alpha_3) + \beta_3(\alpha_2-\alpha_1),\\
A &= \Delta_\beta a_1/a_2, \quad B = \Delta_\alpha b_1/b_2,\\
h(x) &= -\big(A(\alpha_2-\alpha_1)(\alpha_1-\alpha_3)x^2 + B(\beta_2-\beta_1)(\beta_1-\beta_3)\big)\\
&\quad \cdot \big(A(\alpha_3-\alpha_2)(\alpha_2-\alpha_1)x^2 + B(\beta_3-\beta_2)(\beta_2-\beta_1)\big)\\
&\quad \cdot \big(A(\alpha_1-\alpha_3)(\alpha_3-\alpha_2)x^2 + B(\beta_1-\beta_3)(\beta_3-\beta_2)\big).
\end{aligned}$$

Then the $(2,2)$-isogeny with domain $C \times E$ and kernel

$$\big\langle ((\alpha_1,0),(\beta_1,0)),\ ((\alpha_2,0),(\beta_2,0)) \big\rangle$$

has as codomain the Jacobian of a genus 2 curve H defined by $y^2 = h(x)$. The degree-2 morphisms of the dual isogeny are given by

$$\begin{aligned}
\varphi_1 &: H \to C\\
&(x,y) \mapsto (s_1/x^2 + s_2, (\Delta_\beta/A^3)(y/x^3)),\\
\varphi_2 &: H \to E\\
&(x,y) \mapsto (t_1x^2 + t_2, (\Delta_\alpha/B^3)y),
\end{aligned}$$

where

$$\begin{aligned}
s_1 &= -(B/A)(a_2/a_1),\\
s_2 &= \frac{1}{a_1}\left(\frac{\alpha_1(\alpha_3-\alpha_2)^2}{\beta_3-\beta_2} + \frac{\alpha_2(\alpha_1-\alpha_3)^2}{\beta_1-\beta_3} + \frac{\alpha_3(\alpha_2-\alpha_1)^2}{\beta_2-\beta_1}\right),\\
t_1 &= -(A/B)(b_2/b_1),\\
t_2 &= \frac{1}{b_1}\left(\frac{\beta_1(\beta_3-\beta_2)^2}{\alpha_3-\alpha_2} + \frac{\beta_2(\beta_1-\beta_3)^2}{\alpha_1-\alpha_3} + \frac{\beta_3(\beta_2-\beta_1)^2}{\alpha_2-\alpha_1}\right).
\end{aligned}$$

The morphisms φ_i extend to the Jacobian J_H by mapping

$$\Big[\sum_j P_j\Big] \to \sum_j \varphi(P_j)$$

and they combine into a $(2,2)$-isogeny $\Phi : J_H \to C \times E$, the dual of which is our isogeny of interest. To compute the image of a point $(P_c, P) \in C \times E$ under this dual isogeny, it suffices to compute some $[D] \in \Phi^{-1}\{(P_c, P)\} \subset J_H$ and then double it. Indeed, then we have

$$2[D] = \hat{\Phi}\Phi([D]) = \hat{\Phi}(P_c, P)$$

as wanted.

Let $D = P_H + Q_H - \infty_1 - \infty_2$ represent a point on J_H. As mentioned, we assume that its Mumford representation is of the form $[x^2 + u_1 x + u_0, v_1 x + v_0]$. To avoid the need for field extensions, let us express $\varphi_i(P_H + Q_H)$ for $i = 1, 2$ directly in terms of u_0, u_1, v_0, v_1. Note that the divisor $\infty_1 + \infty_2$ maps to ∞, both under φ_1 and under φ_2, so it suffices to concentrate on $P_H + Q_H$.

The calculation is easiest for φ_2, where the line connecting $\varphi_2(P_H)$ and $\varphi_2(Q_H)$ has slope

$$\lambda_2 = -\frac{(\Delta_\alpha / B^3) v_1}{t_1 u_1}$$

and then $\varphi_2(P_H + Q_H)$ is

$$\left(\lambda_2^2 + \sum_{i=1}^{3} \beta_i - t_1(u_1^2 - 2u_0) - 2t_2 \, , \, -\lambda_2 \left(\cdots - t_2 + (u_0 v_1 - u_1 v_0) \frac{t_1}{v_1} \right) \right) \tag{11}$$

with $\cdots$ denoting a copy of the first coordinate. To derive formulae for φ_1, note that this map is of a very similar kind, except for the transformation

$$\tilde{\ } : (x, y) \mapsto (1/x, y/x^3)$$

by which it is preceded. Let $\tilde{u_0}, \tilde{u}_1, \tilde{v_0}, \tilde{v}_1$ be the Mumford coordinates of $\tilde{P}_H + \tilde{Q}_H$, then an easy calculation shows:

$$\tilde{u}_0 = \frac{1}{u_0}, \quad \tilde{u}_1 = \frac{u_1}{u_0}, \quad \tilde{v}_0 = \frac{u_1 v_0 - u_0 v_1}{u_0^2}, \quad \tilde{v}_1 = \frac{u_1^2 v_0 - u_0 v_0 - u_0 u_1 v_1}{u_0^2}.$$

Thus the formulae for the coordinates of $\varphi_1(P_H + Q_H)$ are the same as in (11), except for swapping the α_i's and the β_i's and for substituting $\tilde{u_0}, \tilde{u}_1, \tilde{v_0}, \tilde{v}_1$ for u_0, u_1, v_0, v_1.

This gives us 4 equations in the unknowns u_0, u_1, v_0, v_1:

$$\begin{cases} x(\varphi_1(P_H + Q_H)) = x(P_c), \\ y(\varphi_1(P_H + Q_H)) = y(P_c), \\ x(\varphi_2(P_H + Q_H)) = x(P), \\ y(\varphi_2(P_H + Q_H)) = y(P). \end{cases} \tag{12}$$

Together with the equation

$$\begin{aligned}2v_0^2 - 2v_0v_1u_1 + v_1^2(u_1^2 - 2u_0) = 2c_0 &+ (-u_1)c_1 + (u_1^2 - 2u_0)c_2 \\ &+ (-u_1^3 + 3u_0u_1)c_3 + (u_1^4 - 4u_1^2u_0 + 2u_0^2)c_4 \\ &+ (-u_1^5 + 5u_1^3u_0 - 5u_1u_0^2)c_5 \\ &+ (u_1^6 - 6u_1^4u_0 + 9u_1^2u_0^2 - 2u_0^3)c_6,\end{aligned}$$

expressing that $[D] \in J_H$, this system is expected to have 4 solutions, all of which are defined over $\mathbb{F}_{p^2}$. (In practice, we found these solutions by clearing denominators in (12), running a Gröbner basis computation, and discarding solutions having zeroes among their coordinates, because they are most likely parasite solutions that were created when clearing denominators.) Taking any of these solutions and doubling the corresponding point on J_H produces the desired image of (P_c, P).

8.2 Richelot Isogenies

By assumption, the next $a - 2$ steps are $(2, 2)$-isogenies between Jacobians of genus 2 curves. Such maps are called Richelot isogenies and they are classical; for a contemporary exposition, including explicit formulae, we refer to Smith's thesis [42, Chapter 8]. Starting from a hyperelliptic curve $H : y^2 = h(x)$ and a $(2, 2)$-subgroup

$$\langle [g_1(x), 0], [g_2(x), 0] \rangle, \quad g_1(x) = x^2 + g_{11}x + g_{10}, \quad g_2(x) = x^2 + g_{21}x + g_{20}$$

of its Jacobian, one lets $g_3(x) = h(x)/(g_1(x)g_2(x)) = g_{32}x^2 + g_{31}x + g_{30}$. One then computes

$$\delta = \det \begin{pmatrix} g_{10} & g_{11} & 1 \\ g_{20} & g_{21} & 1 \\ g_{30} & g_{31} & g_{32} \end{pmatrix}$$

and $h'(x) = g_1'(x)g_2'(x)g_3'(x)$ where

$$g_i'(x) = \delta^{-1} \left(\frac{dg_j}{dx} g_k - g_j \frac{dg_k}{dx} \right) \text{ for } (i, j, k) = (1, 2, 3)\ ,\ (2, 3, 1)\ ,\ (3, 1, 2).$$

Then the codomain of our Richelot isogeny is the Jacobian of $H' : \mathbf{y}^2 = h'(\mathbf{x})$. We use different notation for the coordinates because pushing a point through this isogeny is done via the "Richelot correspondence", which is the curve $X \subset H \times H'$ defined by

$$X : g_1(x)g_1'(\mathbf{x}) + g_2(x)g_2'(\mathbf{x}) = y\mathbf{y} - g_1(x)g_1'(\mathbf{x})(x - \mathbf{x}) = 0.$$

It naturally comes equipped with two projection maps $\pi : X \to H$, $\pi' : X \to H'$. The isogeny is then

$$J_H \to J_{H'} : [D] \mapsto [\pi'_*\pi^*D] \quad \text{(pullback along } \pi \text{ and pushforward along } \pi' \text{).}$$

This means that in order to compute the image of a point $[x^2 + u_1x + u_0, v_1x + v_0] \in J_H$, one should eliminate the variables x, y from the system

$$\begin{cases} x^2 + u_1x + u_0 = 0, \\ y = v_1x + v_0, \\ y^2 = h(x), \\ g_1(x)g_1'(\mathbf{x}) + g_2(x)g_2'(\mathbf{x}) = 0, \\ y\mathbf{y} = g_1(x)g_1'(\mathbf{x})(x - \mathbf{x}). \end{cases}$$

We expect the last two equations of its reduced Gröbner basis (with respect to the lexicographic order with $\mathbf{x} \prec \mathbf{y} \prec y \prec x$) to be of the form

$$\mathbf{y} = v_3'\mathbf{x}^3 + v_2'\mathbf{x}^2 + v_1'\mathbf{x} + v_0', \quad \mathbf{x}^4 + u_3'\mathbf{x}^3 + u_2'\mathbf{x}^2 + u_1'\mathbf{x} + u_0' = 0$$

and then $[\mathbf{x}^4 + u_3'\mathbf{x}^3 + u_2'\mathbf{x}^2 + u_1'\mathbf{x} + u_0', v_3'\mathbf{x}^3 + v_2'\mathbf{x}^2 + v_1'\mathbf{x} + v_0']$ are non-reduced Mumford coordinates for the image on $J_{H'}$.

8.3 Split or Not?

We now want to check whether or not the a-th $(2, 2)$-isogeny takes us back to a product of elliptic curves. This is easy: we proceed as if we are dealing with a Richelot isogeny (just the codomain computation, no points need be pushed through anymore). It can be shown that the determinant δ vanishes if and only if the codomain is a product of elliptic curves instead of the Jacobian of a genus 2 curve. Therefore the final and deciding step in our computation simply amounts to verifying whether or not $\delta = 0$.

9 Magma Code

This paper comes with the following auxiliary Magma files, which are available at https://homes.esat.kuleuven.be/~wcastryc/.

- `richelot_aux.m` contains auxiliary functions, mainly for computing chains of $(2, 2)$-isogenies, where the functions `FromProdtoJac` and `FromJactoJac` are implementations of the methods described in Sect. 8,
- `uvtable.m` contains precomputed values of u and v as described in Sect. 7.2,
- runs of `SIKE_challenge1.m`, resp. `SIKE_challenge2.m`, load the first two files and break `$IKEp182`, resp. `$IKEp217`, by running the algorithm from Sect. 6, incorporating the speed-ups from Sect. 7,
- a run of `SIKEp434.m` generates random input for the `SIKEp434` parameters and runs the algorithm from Sect. 6, again incorporating the speed-ups from Sect. 7; to attack `SIKEp503`, `SIKEp610` and `SIKEp751` one simply replaces the line `a := 216; b := 137;` by

  `a := 250; b := 159;`, `a := 305; b := 192;`, `a := 372; b := 239;`,

  respectively.

The reader can execute these files in order to confirm the approximate timings mentioned in Sect. 1. We ran them in Magma V2.27-5 on an Intel Xeon CPU E5-2630v2 at 2.60 GHz.

10 Achieving (heuristic) Polynomial Runtime

As $x \to \infty$, the number of integers c in the interval $[0, x]$ that admit a decomposition of the form $c = u^2 + 4v^2$ is asymptotic to

$$\frac{0.5731...}{\sqrt{\ln x}} x,$$

by (a variation on) a theorem of Landau, see [41]. We can use this to estimate the probability that our strategy from Sect. 5 succeeds in constructing an isogeny $\gamma : E_0 \to C$ of degree $c = 2^a - 3^b$: it is about $0.5731/\sqrt{a \ln 2} \approx 0.6884/\sqrt{a}$.

Let us now revisit the first iteration of our key recovery algorithm from Sect. 6, where we choose $\beta_1 \geq 1$ such that there exists an $\alpha_1 \geq 0$ for which $c_1 = 2^{a-\alpha_1} - 3^{b-\beta_1}$ is of the form $u_1^2 + 4v_1^2$. In view of Landau's theorem, we expect that we should try in the order of $\sqrt{a}$ pairs (α_1, β_1) before we succeed. So the smallest β_1 is expected to be of magnitude $\sqrt[4]{a}$. While this is good enough for breaking the concrete parameter sets of SIKE, the asymptotic runtime is $L_p(1/4)$ rather than polynomial: indeed, there are 3^{β_1} options for κ_1 to guess from.

Remark 8. The first iteration dominates the overall runtime. Indeed, once suitable α_1, β_1 are found, the expression $2^{a-\alpha_1} - 3^{b-\beta_1}$ can be recycled in the remaining iterations by extending Bob's secret isogeny, as explained in Sect. 7.3.

To achieve a polynomial time complexity, we extend the attack from sums of squares to more general quadratic forms and hope that there is a prime number $n \leq a$ such that c_1 can be written as $u_1^2 + nv_1^2$. Heuristically, this happens with overwhelming probability. We can loosely argue this as follows. Based on a generalization of Landau's theorem, see again [41], for every n the success probability remains inversely proportional to $\sqrt{a}$. If the events of being of the form $u_1^2 + nv_1^2$ are "sufficiently independent" as n varies, and if the implicit constants do not decay too quickly, then the probability of failure overall is in the order of

$$\left(1 - \frac{1}{\sqrt{a}}\right)^{\pi(a)} \approx \left(1 - \frac{1}{\sqrt{a}}\right)^{a/\ln a},$$

which decreases as $e^{-\sqrt{a}/\ln a}$ (here π is the prime-counting function). In particular, we expect that we can simply take $\beta_1 = 1$ in this case.

Once such a decomposition $u_1^2 + nv_1^2$ is found, we proceed as follows. The techniques from Love and Boneh [29] allow for the polynomial-time construction of an isogeny $\nu : E_{\text{start}} \to N_{\text{start}}$, where N_{start} is an elliptic curve possessing an endomorphism $\sqrt{n}\mathbf{i}$ satisfying $\sqrt{n}\mathbf{i} \circ \sqrt{n}\mathbf{i} = [-n]$. Thus we can consider the degree-c endomorphism $\gamma_{\text{start}} = [u_1] + \sqrt{n}\mathbf{i} \circ [v_1]$ on N_{start}. This endomorphism can be transformed into the desired degree-c isogeny $\gamma : E_0 \to C$ along $\nu \circ \tau : E_0 \to N_{\text{start}}$, as outlined in Sect. 5.

Remark 9. In general, when compared to the method from Remark 3, it becomes more cumbersome to test whether or not an integer of the form $c = 2^a - 3^b$ admits a decomposition $u^2 + nv^2$ (and find corresponding u, v). Again we need to factor

$$c = \ell_1 \ell_2 \cdots \ell_s,$$

where for simplicity we assume that c is squarefree, i.e., the ℓ_i are pairwise distinct primes. Then a necessary condition is that $-n$ is a quadratic residue modulo each ℓ_i. In this case we can decompose $\ell_i \mathbb{Z}[\sqrt{-n}] = \mathfrak{l}_i \bar{\mathfrak{l}}_i$ into a product of two prime ideals of norm ℓ_i. We then look for a relation of the form

$$1 = \prod_{i=1}^{s} [\mathfrak{l}_i]^{\sigma_i}, \qquad \sigma_i \in \{\pm 1\} \tag{13}$$

in the ideal-class group of $\mathbb{Z}[\sqrt{-n}]$. If we succeed, then the ideal

$$\prod_{i=1}^{s} \mathfrak{l}_i^{\delta_{\sigma_i,1}} \bar{\mathfrak{l}}_i^{\delta_{\sigma_i,-1}}$$

(with $\delta_{\cdot,\cdot}$ the Kronecker delta) is a principal ideal of norm c, hence generated by $u + \sqrt{-n}v$ for integers u, v of the desired form. All ideal-class group arithmetic can be done in polynomial time, see e.g. [22], because $n \leq a$. The identity (13) is of knapsack type, but we nevertheless expect being able to decide if it exists (and find it) in polynomial time, because the expected value of s is $\log \log c \approx \log a$ by the Hardy–Ramanujan theorem.

11 Generalizations

In this final section, we move away from the SIKE set-up and discuss how to attack more general instantiations of SIDH.

11.1 Arbitrary Torsion

There is no theoretical obstruction to attacking Alice's public key instead of Bob's. In this case one will end up computing a chain of $(3, 3)$-isogenies, which is more convoluted, but still doable using the machinery from [4]; see also [17]. The formulae are still practical and recovering Alice's private key can then be done bit by bit (except possibly for some offset of the kind discussed in Sect. 7.3). Altogether, we expect having to compute approximately a chains of $(3, 3)$-isogenies of length at most b in order to retrieve Alice's private key. The expression $\varDelta$ in the formulae from [4] plays a similar role as δ in the Richelot isogeny formulae, in the sense that $\varDelta = 0$ occurs if and only if the codomain of the $(3, 3)$-isogeny is the product of two elliptic curves, see [6]. Therefore, verifying whether the final $(3, 3)$-isogeny splits is just as easy.

More generally, one can attack SIDH when set up using arbitrary small primes $\ell_\mathrm{A}, \ell_\mathrm{B}$ instead of just $2, 3$, or even more general smooth torsion as in B-SIDH.

Inherently, this changes nothing to our attack, except that now one must compute (ℓ, ℓ)-isogenies for primes $\ell \geq 5$. For isogenies between Jacobians of genus 2 curves, we refer to the work of Cosset and Robert [8], whose formulae are a lot more involved than those to compute $(2,2)$- and $(3,3)$-isogenies, but they are polynomial in ℓ and likely practical enough to complete the attack. The gluing of elliptic curves and splitting of Jacobians is succinctly explained by Kuhn in [28]; for a more elaborate and practical exposition, see also [15, Sect. 1.4]. Away from $\ell = 2, 3$ we are not aware of a straightforward decision algorithm to verify whether an (ℓ, ℓ)-subgroup of a given Jacobian of a genus 2 curve results in a product of elliptic curves: the easiest way seems to try and compute an (ℓ, ℓ)-isogeny to a Jacobian as in [8] and see if the theta constants fail to create a genus 2 curve. Alternatively, one can write down a system of equations expressing that our Jacobian is "(ℓ, ℓ)-split" (i.e., (ℓ, ℓ)-isogenous to a product of elliptic curves) via our given subgroup, and verify whether this system is consistent, see [15].

11.2 Other Starting Curves with a Known Endomorphism Ring

Setting up SIDH with another starting curve E_0 with known endomorphism ring does not prevent the attack. Indeed, in view of [16,44], such a curve can always be assumed to come equipped with an explicit 3^{β}-isogeny $\tau : E_0 \to E_{\text{start}}$ for some $\beta \geq 0$, where E_{start} is any of the curves from (1). Therefore we fall under the set-up from Sect. 3.

11.3 Base Curves Whose Endomorphism Ring is Unknown

We now discuss the scenario of a base curve E_0 without known endomorphism ring. In particular, no path to E_{start} is known. As indicated in Sect. 5.4, if $c = 2^a - 3^b$ is smooth then it remains possible to construct the auxiliary isogeny γ. In fact, if we no longer exploit special features of E_0, then it makes more sense to let γ emanate from E rather than E_0, leading us to considering $\gamma \circ \varphi : E_0 \to C$. This isogeny has degree $c3^b$ and can again be used to decide whether or not (D) is true: this should be the case if and only if the subgroup $\langle (P_0, x\gamma(P)), (Q_0, x\gamma(Q)) \rangle \subset E_0 \times C$ is reducible, with x a multiplicative inverse of 3^b modulo 2^a.

Remark 10. Computing γ works as follows. Write c as a product of small primes $\ell_1 \ell_2 \cdots \ell_s$ and for each $i = 1, \ldots, s$ let r_i denote the multiplicative order of $-p$ modulo ℓ_i. Because p^2-Frobenius acts as $[-p]$, we can find a non-trivial point in $E_0[\ell_1] \subset E_0(\mathbb{F}_{p^{2r_1}})$ and the subgroup it generates is defined over $\mathbb{F}_{p^2}$. So this is the kernel of an $\mathbb{F}_{p^2}$-rational degree-ℓ_1 isogeny $\gamma_1 : E_0 \to C_1$ that can be computed and evaluated using formulae of Vélu type. By repeating this construction, we eventually obtain γ as a composition $\gamma_s \circ \gamma_{s-1} \circ \ldots \circ \gamma_1$ where each γ_i is an $\mathbb{F}_{p^2}$-rational ℓ_i-isogeny.

Turning this decision method into a key recovery algorithm works along the lines of Sect. 6. First, we look for the smallest $\beta \geq 1$ for which there exists an integer $\alpha \geq 0$ such that

$$c = 2^{a-\alpha} - 3^{b-\beta} \tag{14}$$

is smooth (this is an optimistic goal!). Then, for each guess for the first degree-3^β-component κ_1 of φ, we run our test to see whether or not there exists a degree-$3^{b-\beta}$-isogeny $\kappa_1(E_0) \to E$ mapping $2^\alpha \kappa_1(P_0)$ to $2^\alpha P$ and $2^\alpha \kappa_1(Q_0)$ to $2^\alpha Q$. There are 3^β possible guesses, so clearly β should be small enough for this to be feasible.

Once κ_1 is found, we can proceed by steps of degree 3 as in Sect. 7.3. Since smoothness is such a rare event, it actually makes sense to recycle the expression (14) all along. Then we can also recycle our auxiliary isogeny γ, i.e., it only has to be computed once, including pushing through torsion points. Concretely: when guessing κ_2, we extend γ with an extra degree-3 isogeny $\varphi' : C \to E'$ and we test if we took the right direction by checking whether or not there is a degree $c3^{b-\beta}$-isogeny mapping $2^\alpha \kappa_2 \kappa_1(P_0)$ to $2^\alpha \varphi' \gamma(P)$ and $2^\alpha \kappa_2 \kappa_1(Q_0)$ to $2^\alpha \varphi' \gamma(Q)$. Iterating this process will recreate the entire isogeny chain.

In summary: as soon as we can find a small $\beta \geq 1$ with a corresponding $\alpha \geq 0$ such that (14) is smooth, then our attack applies. The likelihood of finding a smooth c of this form is very small, but there are at least two methods for creating more leeway for an attacker:

- We can extend Bob's secret isogeny $\varphi : E_0 \to E$ by an arbitrary isogeny $\varepsilon : E \to F$ of some smooth degree e and work with $\varepsilon \circ \varphi$ instead of φ. This allows us to look for a smooth integer of the form $c = 2^{a-\alpha} - e3^{b-\beta}$ and construct a corresponding degree-c isogeny $\gamma : F \to C$.
- A second tweak can be obtained by any algorithm that can efficiently solve the following problem for a fixed d:
  - Let $H/\mathbb{F}_{p^2}$ be a genus 2 curve with superspecial Jacobian J, and $d > 1$ an integer. Is there a (d, d)-isogeny $\Psi : J \to A$ such that A is a product of elliptic curves?

  Indeed, this allows us to work with expressions of the form $c = d2^{a-\alpha} - e3^{b-\beta}$. Each test then amounts to computing a $(2^{a-\alpha}, 2^{a-\alpha})$-isogeny, using the torsion point data as before, and then checking if the resulting Jacobian is (d, d)-split. Verifying whether a given Jacobian is (d, d)-split is likely to be most efficient by means of a computation similar to those in [15], [28]. Alternatively, one can exhaust over all $O(d^3)$ outgoing (d, d)-isogenies.

E.g., consider $a = 110$ and $b = 67$ as in `$IKEp217`, along with the identity

$$59 \cdot 67 \cdot 107 \cdot 443^2 \cdot 487 \cdot 1049 \cdot 2711 \cdot 8297 = 109 \cdot 2^{110-35} - 119 \cdot 3^{67-20}.$$

Assuming that we do not know a path from E_0 to E_{start}, we could still try to recover Bob's key by computing

- one-time isogenies $E \xrightarrow{\varepsilon} F \xrightarrow{\gamma} C$, dominated in cost by a 2711-isogeny and a 8297-isogeny over extension fields of respective degrees 2710 and 2074,
- computing all 3^{20}-isogenous neighbours of the base curve, gluing them together by means of a $(2^{75}, 2^{75})$-isogeny and checking which one of the resulting Jacobians is $(109, 109)$-split.

The second step immediately reveals the first 20 ternary digits of Bob's secret key and we can then easily find the remaining digits as explained above.

Remark 11. It was pointed out to us by De Feo and Wesolowski that the above considerations lead to an algorithm which, heuristically, runs in time $L_p(1/2+\epsilon)$. To see this, it suffices to pick α, β in the order of $\sqrt{a}$. Then, by letting d, e range over random integers in $[1, 2^{\sqrt{a}}]$, we can think of c as a random integer of size roughly 2^a. Following well-known heuristics [5], after about

$$\sqrt{a}^{\sqrt{a}} = L_p(1/2+\epsilon)$$

tries we expect to find an instance of c that is $2^{\sqrt{a}}$-smooth. Using these values of c, d, e, the remainder of the attack is expected to run in time $L_p(1/2)$.

Acknowledgements. We thank Craig Costello, Luca De Feo, Luciano Maino, Frederik Vercauteren, Benjamin Wesolowski, Yifan Zheng and the anonymous reviewers for helpful discussions, questions and suggestions.

References

1. Azarderakhsh, R., Jao, D., Leonardi, C.: Post-quantum static-static key agreement using multiple protocol instances. In: Adams, C., Camenisch, J. (eds.) Selected Areas in Cryptography - SAC 2017, pp. 45–63. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-72565-9_3
2. Bosma, W., Cannon, J., Playoust, C.: The Magma algebra system. I. The user language. J. Symbolic Comput. **24**(3–4), 235–265 (1997). https://doi.org/10.1006/jsco.1996.0125
3. Brock, B.: Superspecial curves of genera two and three. Ph.D. thesis, Princeton University (1994)
4. Bruin, N., Flynn, E.V., Testa, D.: Descent via $(3,3)$-isogeny on Jacobians of genus 2 curves. Acta Arithmetica **165**(3), 201–223 (2014). http://eudml.org/doc/279018
5. Canfield, E.R., Erdös, P., Pomerance, C.: On a problem of Oppenheim concerning "factorisatio numerorum." J. Number Theory **17**(1), 1–28 (1983). https://doi.org/10.1016/0022-314X(83)90002-1
6. Castryck, W., Decru, T.: Multiradical isogenies. In: Anni, S., Karemaker, V., Lorenzo García, E. (eds.) 18th International Conference Arithmetic, Geometry, Cryptography, and Coding Theory, Contemporary Mathematics, vol. 779, pp. 57–89. American Mathematical Society (2022). https://doi.org/10.1090/conm/779
7. Castryck, W., Lange, T., Martindale, C., Panny, L., Renes, J.: CSIDH: an efficient post-quantum commutative group action. In: Peyrin, T., Galbraith, S. (eds.) Advances in Cryptology - ASIACRYPT 2018, vol. 3, pp. 395–427. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03332-3_15
8. Cosset, R., Robert, D.: Computing (ℓ, ℓ)–isogenies in polynomial time on Jacobians of genus 2 curves. Math. Comput. **84**(294), 1953–1975 (2015). https://www.ams.org/journals/mcom/2015-84-294/S0025-5718-2014-02899-8/
9. Costello, C.: B-SIDH: supersingular isogeny Diffie-Hellman using twisted torsion. In: Moriai, S., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2020, vol. 2, pp. 440–463. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64834-3_15
10. Couveignes, J.M.: Hard homogeneous spaces. Cryptology ePrint Archive, Paper 2006/291 (2006). https://eprint.iacr.org/2006/291

11. De Feo, L., Jao, D., Plût, J.: Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. J. Math. Cryptol. **8**(3), 209–247 (2014). https://doi.org/10.1515/jmc-2012-0015
12. De Feo, L., Kohel, D., Leroux, A., Petit, C., Wesolowski, B.: SQISign: compact post-quantum signatures from quaternions and isogenies. In: Moriai, S., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2020, vol. 1, pp. 64–93. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64837-4_3
13. De Feo, L., et al.: Séta: supersingular encryption from torsion attacks. In: Tibouchi, M., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2021, vol. 4, pp. 249–278. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-92068-5_9
14. De Feo, L., et al.: (open project): Is SIKE broken yet? (2022). https://issikebrokenyet.github.io/
15. Djukanovic, M.: Split Jacobians and lower bounds on heights. Ph.D. thesis, Université de Bordeaux (2017)
16. Eisenträger, K., Hallgren, S., Lauter, K., Morrison, T., Petit, C.: Supersingular isogeny graphs and endomorphism rings: Reductions and solutions. In: Nielsen, J.B., Rijmen, V. (eds.) Advances in Cryptology - EUROCRYPT 2018, vol. 3, pp. 329–368. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-78372-7_11
17. Flynn, E.V., Ti, Y.B.: Genus two isogeny cryptography. In: Ding, J., Steinwandt, R. (eds.) Post-quantum Cryptography, pp. 286–306. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-25510-7_16
18. Fouotsa, T.B., Moriya, T., Petit, C.: M-SIDH and MD-SIDH: countering SIDH attacks by masking information. Cryptology ePrint Archive, Paper 2023/013 (2023). https://eprint.iacr.org/2023/013
19. Fouotsa, T.B., Petit, C.: SHealS and HealS: isogeny-based PKEs from a key validation method for SIDH. In: Tibouchi, M., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2021, vol. 4, pp. 279–307. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-92068-5_10
20. Galbraith, S.D., Petit, C., Silva, J.: Identification protocols and signature schemes based on supersingular isogeny problems. In: Takagi, T., Peyrin, T. (eds.) Advances in Cryptology - ASIACRYPT 2017, vol. 1, pp. 3–33. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70694-8_1
21. Galbraith, S.D., Vercauteren, F.: Computational problems in supersingular elliptic curve isogenies. Quantum Inf. Process. **17**(10), 1–22 (2018). https://doi.org/10.1007/s11128-018-2023-6
22. Hafner, J.L., McCurley, K.S.: A rigorous subexponential algorithm for computation of class groups. J. Am. Math. Soc. **2**(4), 837–850 (1989). https://doi.org/10.1090/S0894-0347-1989-1002631-0
23. Howe, E.W., Leprévost, F., Poonen, B.: Large torsion subgroups of split Jacobians of curves of genus two or three. Forum Math. **12**(3), 315–364 (2000). https://doi.org/10.1515/form.2000.008
24. Jao, D., et al.: Supersingular Isogeny Key Encapsulation. https://csrc.nist.gov/Projects/post-quantum-cryptography/round-4-submissions
25. Jao, D., De Feo, L.: Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In: Yang, B.Y. (ed.) Post-Quantum Cryptography, pp. 19–34. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25405-5_2
26. Kani, E.: The number of curves of genus two with elliptic differentials. J. für die reine und angewandte Mathematik **1997**(485), 93–122 (1997). https://doi.org/10.1515/crll.1997.485.93

27. Kohel, D., Lauter, K., Petit, C., Tignol, J.P.: On the quaternion ℓ-isogeny path problem. LMS J. Comput. Math. **17**(A), 418–432 (2014). https://doi.org/10.1112/S1461157014000151
28. Kuhn, R.M.: Curves of genus 2 with split Jacobian. Trans. Am. Math. Soc. **307**(1), 41–49 (1988). https://doi.org/10.2307/2000749
29. Love, J., Boneh, D.: Supersingular curves with small non-integer endomorphisms. In: Algorithmic Number Theory Symposium (ANTS-XIV), MSP Open Book Series, vol. 4, pp. 7–22 (2020). https://doi.org/10.2140/obs.2020.4.7
30. Maino, L., Martindale, C.: An attack on SIDH with arbitrary starting curve. Cryptology ePrint Archive, Paper 2022/1026 (2022). https://eprint.iacr.org/2022/1026
31. Martindale, C., Panny, L.: How to not break SIDH. Cryptology ePrint Archive, Paper 2019/558 (2019). https://eprint.iacr.org/2019/558, Presented at CFAIL 2019, Columbia University
32. Microsoft: SIKE cryptographic challenge. https://www.microsoft.com/en-us/msrc/sike-cryptographic-challenge
33. National Institute of Standards and Technology (NIST): Post-quantum cryptography standardization process. https://csrc.nist.gov/projects/post-quantum-cryptography
34. Oudompheng, R.: A note on implementing direct isogeny determination in the Castryck–Decru attack. https://www.normalesup.org/~oudomphe/textes/202208-castryck-decru-shortcut.pdf
35. Oudompheng, R., Pope, G.: A note on reimplementing the Castryck–Decru attack and lessons learned for SageMath. Cryptology ePrint Archive, Paper 2022/1283 (2022). https://eprint.iacr.org/2022/1283
36. Petit, C.: Faster algorithms for isogeny problems using torsion point images. In: Takagi, T., Peyrin, T. (eds.) Advances in Cryptology - ASIACRYPT 2017, vol. 2, pp. 330–353. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70697-9_12
37. de Quehen, V., et al.: Improved torsion-point attacks on SIDH variants. In: Malkin, T., Peikert, C. (eds.) Advances in Cryptology - CRYPTO 2021, vol. 3, pp. 432–470. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-84252-9_15
38. Robert, D.: Breaking SIDH in polynomial time. Cryptology ePrint Archive, Paper 2022/1038 (2022). https://eprint.iacr.org/2022/1038
39. Rostovtsev, A., Stolbunov, A.: Public-key cryptosystem based on isogenies. Cryptology ePrint Archive, Paper 2006/145 (2006). https://eprint.iacr.org/2006/145
40. SageMath: The Sage Mathematics Software System. https://www.sagemath.org
41. Shanks, D., Schmid, L.P.: Variations on a theorem of Landau. Part I. Math. Comput. **20**(96), 551–569 (1966). https://doi.org/10.2307/2003544
42. Smith, B.: Explicit endomorphisms and correspondences. Ph.D. thesis, University of Sydney (2006)
43. Urbanik, D., Jao, D.: SoK: the problem landscape of SIDH. In: Emura, K., Seo, J.H., Watanabe, Y. (eds.) Proceedings of the 5th ACM on ASIA Public-Key Cryptography Workshop, APKC@AsiaCCS, Incheon, Republic of Korea, 4 June 2018, pp. 53–60. ACM (2018). https://doi.org/10.1145/3197507.3197516
44. Wesolowski, B.: The supersingular isogeny path and endomorphism ring problems are equivalent. In: 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS), pp. 1100–1111 (2022). https://doi.org/10.1109/FOCS52979.2021.00109
45. Wesolowski, B.: Understanding and improving the Castryck–Decru attack on SIDH (2022). https://www.bweso.com/papers.php