SIDH Proof of Knowledge

Luca De Feo[1], Samuel Dobson[2], Steven D. Galbraith[2], and Lukas Zobernig[2]

[1]IBM Research Europe. luca@defeo.lu
[2]Mathematics Department, University of Auckland, New Zealand.
samuel.dobson.nz@gmail.com, s.galbraith@auckland.ac.nz,
lukas.zobernig@auckland.ac.nz

August 5, 2021

Abstract

We demonstrate the soundness proof for the De Feo, Jao and Plût identification scheme (the basis for SIDH signatures) contains an invalid assumption and provide a counterexample for this assumption - thus showing the proof of soundness is invalid. As this proof was repeated in a number of works by various authors, multiple pieces of literature are affected by this result. Due to the importance of being able to prove knowledge of an SIDH key (for example, to prevent adaptive attacks), soundness is a vital property. We propose a modified identification scheme fixing the issue with the De Feo, Jao and Plût scheme, and provide a proof of security of this new scheme. We also prove that a modification of this scheme allows the torsion points in the public key to be verified too, which previous schemes did not cover. This results in a secure proof of knowledge for SIDH keys and a secure SIDH-based signature scheme. In particular, these schemes provide a non-interactive way of verifying that SIDH public keys are well formed as protection against adaptive attacks, more efficient than generic NIZKs.

1 Introduction

De Feo, Jao and Plût [DJP14] introduced an identification protocol based on SIDH. This was used to build an undeniable signature by Jao and Soukharev [JS14], and was also used to build a signature scheme by Yoo, Azarderakhsh, Jalali, Jao and Soukharev [YAJ+17] and also by Galbraith, Petit and Silva [GPS20]. All these papers claim that the scheme has the soundness property based on the SIDH assumption. In this paper we show that these claims are not rigorously proved, and provide a counterexample to all such previous proofs. The issue may also affect some subsequent work, such as that by Urbanik and Jao [UJ20].

Soundness of the identification scheme is especially important considered in the light of adaptive attacks on the SIDH scheme, given by Galbraith et al. [GPST16] and extended in Dobson et al. [DGL+20]. Being able to successfully complete an identification protocol for a given public key should prove knowledge of the corresponding secret which should allow the verifier to check that the key is well-formed and not maliciously modified. Unfortunately, without a proof of soundness, we cannot rely on this assumption. Additionally, the original SIDH identification and signature schemes did not provide any guarantees about the torsion points in the public key - only verifying the elliptic curve itself - and it is precisely these points that allow adaptive attacks.

Without an efficient method of proving that SIDH public keys are honestly formed, inefficient protocols such as k-SIDH have been proposed as the best known solution to allow static keys to be used in a key exchange without risk of adaptive attack. The k-SIDH protocol runs many instances of SIDH in parallel and combines the parallel secrets into the final shared secret. The authors suggest $k = 92$ is required for a secure key exchange. Urbanik and Jao [UJ20]'s proposal attempted to improve the efficiency of this protocol by making use of the special automorphisms on curves with j-invariant 0 or 1728, but it was shown by Basso et al. [BKM$^+$20] that Urbanik and Jao's proposal is vulnerable to adaptive attack and actually scales worse than k-SIDH itself. Thus, while key exchange with a non-static key is possible á la SIKE [ACC$^+$17], we do not have an efficient way of obtaining static-static key exchange from isogenies to date.

In this work we examine the issue with the existing soundness proofs and propose a new SIDH-based identification scheme which we prove does satisfy special soundness. This gives the first sound proof of knowledge protocol of a secret isogeny for a given public key. We then propose a modification to the scheme which allows the two torsion points in the public key to be proved correct, thus giving a secure method for proving well-formedness of SIDH public keys. This scheme is much more efficient than using generic NIZKs for the task, and has important applications in all areas where SIDH key exchanges could be used with static keys. By providing an accompanying proof of well-formedness for static public keys, they can safely be used in SIDH exchanges without the need for inefficient protocols like k-SIDH (while the size of our NIZK proof is larger than a k-SIDH public key of the same security level, it is much more efficient to verify than computing a k-SIDH exchange, which requires k^2 isogeny computations).

1.1 Outline

This work begins in Section 2 with revision of some preliminary background material. We then recall the De Feo-Jao-Plût identification scheme in Section 3.1 and outline the issue with its proof of soundness (given in multiple previous works) in Section 3.2. Subsequently, we present a new SIDH identification scheme in Section 4 which modifies the De Feo-Jao-Plût scheme and allows us to prove soundness (and thus security). We then show how the points in the SIDH public key can also be verified under this identification scheme in Section 5. From this, we construct a secure signature scheme which is a Proof of Knowledge (PoK) of an SIDH secret key, and is the first such scheme to prove correctness of the points in the public key (a protection mechanism against adaptive attacks [GPST16, DGL$^+$20]) in Section 6.

1.2 Acknowledgements

Thank you to David Jao, Jason LeGrow, and Yi-Fu Lai for useful discussion about this work.

2 Preliminaries

Notation. We begin with some notation and conventions that we will use throughout this paper. We will use K_ϕ to denote a point which generates the kernel of an isogeny ϕ. Let $[t]$ denote the set $\{1, \ldots, t\}$.

2.1 SIDH

We now provide a brief refresher on the Supersingular Isogeny Diffie-Hellman (SIDH) key exchange protocol [JD11, DJP14] by De Feo, Jao, and Plût.

As public parameters, we have a prime $p = \ell_1^{e_1} \cdot \ell_2^{e_2} \cdot f \pm 1$, where ℓ_1, ℓ_2 are small primes, f is an integer cofactor, and $\ell_1^{e_1} \approx \ell_2^{e_2}$. We work over the finite field $\mathbb{F}_{p^2}$. Additionally we fix a base supersingular elliptic curve E_0 and bases $\{P_1, Q_1\}$, $\{P_2, Q_2\}$ for both the $\ell_1^{e_1}$ and $\ell_2^{e_2}$ torsion subgroups of $E_0(\mathbb{F}_{p^2})$ respectively (such that $E_0[\ell_i^{e_i}] = \langle P_i, Q_i \rangle$). Typically $\ell_1 = 2$ and $\ell_2 = 3$.

It is well known that knowledge of an isogeny and knowledge of its kernel are equivalent, and we can convert between them at will, via Vélu's formulae [Vél71]. In SIDH, the secret key of Alice (respectively Bob) is an isogeny $\phi : E(\mathbb{F}_{p^2}) \to E_A(\mathbb{F}_{p^2})$ of degree $\ell_1^{e_1}$ (respectively $\ell_2^{e_2}$). These isogenies are generated by randomly choosing secret integers $a_i, b_i \in \mathbb{Z}/\ell_i^{e_i}\mathbb{Z}$ (not both divisible by ℓ_i) and computing the isogeny with kernel $K_i = \langle [a_i]P_i + [b_i]Q_i \rangle$. We thus unambiguously refer to the isogeny, its kernel, and such integers a, b, as "the secret key." Figure 1 depicts the commutative diagram making up the key exchange.

$$\begin{tikzcd} E \arrow[r, "\phi_A"] \arrow[d, "\phi_B"'] & E_A \arrow[d, "\phi_{AB}"] \\ E_B \arrow[r, "\phi_{BA}"'] & E_{AB} \end{tikzcd}$$

Figure 1: Commutative diagram of SIDH, where $\ker(\phi_{BA}) = \phi_B(\ker(\phi_A))$ and $\ker(\phi_{AB}) = \phi_A(\ker(\phi_B))$.

In order to make the diagram commute, Alice and Bob are required to not just give their image curves E_A and E_B in their respective public keys, but also the images of the basis points of the other participant's kernel on E. That is, Alice provides E_A, $P_2' = \phi_A(P_2), Q_2' = \phi_A(Q_2)$ as her public key. This allows Bob to "transport" his secret isogeny to E_A and compute ϕ_{AB} whose kernel is $\langle [a_2]P_2' + [b_2]Q_2' \rangle$. Both Alice and Bob will arrive along these transported isogenies at isomorphic image curves E_{AB}, E_{BA} (using Vélu's formulae, they will actually arrive at exactly the same curve). Two elliptic curves are isomorphic over $\overline{\mathbb{F}}_{p^2}$ if and only if their j-invariants $j(E_{AB}) = j(E_{BA})$, hence this j-invariant may be used as the shared secret of the SIDH key exchange.

Remark 1. Galbraith et al. [GPST16, Lemma 2.1] formally presented the idea of "equivalent keys" (which were previously implicit in some previous work including Costello et al. [CLN16]). Two secret keys (a, b) and (a', b') are equivalent if they generate the same subgroup for any basis of the $\ell_i^{e_i}$ torsion subgroup. This is true when $(a', b') = (\theta a, \theta b)$ for $\theta \in \mathbb{Z}_{\ell_i}^*$. Because we have the condition that at least one of a, b is not divisible by ℓ_i (assume for now this is a), a is invertible modulo $\ell_i^{e_i}$. Thus we can choose $\theta \equiv a^{-1} \pmod{\ell_i^{e_i}}$. This gives an equivalent key $(1, b')$. Similarly, if b was not divisible by ℓ_i, we can invert it and obtain equivalent key $(a', 1)$. Hence we obtain a shorter representation of secret keys without loss of generality, to a single element and one extra bit.

2.2 SIDH assumptions

We shall recall the standard isogeny-based hardness assumptions of relevance to this work.

Definition 1 (General isogeny problem)**.** *Given j-invariants $j, j' \in \mathbb{F}_q$, find an isogeny $\phi : E \to E'$ if one exists, where $j(E) = j$ and $j(E') = j'$.*

This is the foundational hardness of assumption of isogeny-based cryptography, that it is hard to find an isogeny between two given curves. Note the decisional version, determining whether an isogeny exists, is easy - an isogeny exists if and only if the cardinality $\#E(\mathbb{F}_q) = \#E'(\mathbb{F}_q)$.

Definition 2 (Computational Supersingular Isogeny (CSSI) problem)**.** *For fixed SIDH public parameters $(p, E_0, P_1, Q_1, P_2, Q_2)$, let $\phi : E_0 \to E_A$ be an isogeny of degree $\ell_1^{e_1}$. Given the SIDH public key $(E_A, P = \phi(P_2), Q = \phi(Q_2))$, find an isogeny $\phi' : E_0 \to E_A$ of degree $\ell_1^{e_1}$ such that $P, Q = \phi'(P_2), \phi'(Q_2)$.*

This is problem 5.2 of [DJP14], and essentially states that it is hard to find the secret key corresponding to a given public key.

Definition 3 (Decisional Supersingular Product (DSSP) problem)**.** *Let E_0, E_1 be supersingular elliptic curves such that there exists an isogeny $\phi : E_0 \to E_1$ of degree $\ell_1^{e_1}$ between them. Let $P_2, Q_2 \in E_0[\ell_2^{e_2}]$ be a fixed basis of the $\ell_2^{e_2}$ torsion subgroup. Suppose we have the following two distributions:*

- *(E_2, E_3, ϕ') such that there exists a cyclic subgroup $G \subseteq E[\ell_2^{e_2}]$ of order $\ell_2^{e_2}$ and $E_2 \cong E_0/G$ and $E_3 \cong E_1/\phi(G)$, and $\phi' : E_2 \to E_3$ is a degree $\ell_1^{e_1}$ isogeny.*
- *(E_2, E_3, ϕ') such that E_2 is a random supersingular curve with the same cardinality as E_0, and E_3 is the codomain of a random isogeny $\phi' : E_2 \to E_3$ of degree $\ell_1^{e_1}$.*

The Decisional Supersingular Product problem is, given E_0, E_1 as well as the points $P_2, Q_2, \phi(P_2), \phi(Q_2)$, and given a tuple (E_2, E_3, ϕ') drawn randomly with probability 1/2 from the above two distributions, to determine which of the two distributions it was drawn from.

This is problem 5.5 of [DJP14] and intuitively states that it is hard to determine whether there exists valid "vertical sides" to an SIDH square given the corners and the bottom horizontal side.

2.3 Sigma protocols

A sigma protocol Π_Σ for a relation $\mathcal{R} = \{(X, W)\}$ is a public-coin three-move interactive proof system consisting of two parties - a verifier V and a prover P.

Definition 4 (Sigma protocol)**.** *A sigma protocol Π_Σ for a family of relations $\{\mathcal{R}\}_\kappa$ parametrized by security parameter κ consists of PPT algorithms $((P_1, P_2), (V_1, V_2))$ where V_2 is deterministic and we assume P_1, P_2 share states. The protocol proceeds as follows:*

1. *Round 1: The prover, on input $(X, W) \in \mathcal{R}$, returns a commitment $\mathsf{com} \leftarrow P_1(X, W)$ and sends com to the verifier.*
2. *Round 2: The verifier, on receipt of com, runs $\mathsf{chall} \leftarrow V_1(1^\kappa)$ to obtain a random challenge, and sends this to the prover.*
3. *Round 3: The prover then runs $\mathsf{resp} \leftarrow P_2(X, W, \mathsf{chall})$ and returns resp to the verifier.*
4. *Verification: The verifier runs $V_2(X, \mathsf{com}, \mathsf{chall}, \mathsf{resp})$ and outputs either $\top$ (accept) or $\perp$ (reject).*

A transcript $(\mathsf{com}, \mathsf{chall}, \mathsf{resp})$ is said to be valid if $V_2(X, \mathsf{com}, \mathsf{chall}, \mathsf{resp})$ outputs $\top$. Let $\langle P, V\rangle$ denote the transcript for interaction between prover P and verifier V. Relevant properties of a sigma protocol are:

Correctness: If the prover P knows $(X, W) \in \mathcal{R}$ and behaves honestly, then the verifier V accepts.

2-special soundness: There exists a polynomial time extraction algorithm $\mathsf{Extract}$, which given a statement X and access to any prover P^*, outputs a witness W such that $(X, W) \in \mathcal{R}$ with probability at least $\Pr[\langle P^*, V\rangle = 1] - \varepsilon$ for soundness error ε.

Zero Knowledge (ZK): There exists a polynomial time simulator Sim, which given a statement X, and for any (cheating) verifier V^*, outputs transcripts $(\mathsf{com}, \mathsf{chall}, \mathsf{resp})$ that are indistinguishable from valid interactions between a prover P and V^*.

3 Previous SIDH identification scheme and soundness issue

3.1 De Feo-Jao-Plût scheme

Let p be a large prime of the form $\ell_1^{e_1} \cdot \ell_2^{e_2} \cdot f \pm 1$, where ℓ_1, ℓ_2 are small primes. We start with a supersingular elliptic curve E_0 defined over $\mathbb{F}_{p^2}$ with $\#E_0(\mathbb{F}_{p^2}) = (\ell_1^{e_1}\ell_2^{e_2} f)^2$. The private key is a random point $K_\phi \in E_0(\mathbb{F}_{p^2})$ of exact order $\ell_1^{e_1}$. Define $E_1 = E_0/\langle K_\phi\rangle$ and denote the corresponding $\ell_1^{e_1}$-isogeny by $\phi : E_0 \to E_1$.

Let P_0, Q_0 be a basis of the torsion subgroup $E_0[\ell_2^{e_2}] = \langle P_0, Q_0\rangle$. The fixed public parameters are $pp = (p, E_0, P_0, Q_0)$. The public key is $(E_1, \phi(P_0), \phi(Q_0))$. The private key is the kernel generator K_ϕ (equivalently, the isogeny ϕ). The interaction goes as follows:

1. The prover chooses a random primitive $\ell_2^{e_2}$-torsion point K_ψ as $K_\psi = [a]P_0 + [b]Q_0$ for some integers $0 \le a, b < \ell_2^{e_2}$ not both divisible by ℓ_2. Note that $\phi(K_\psi) = [a]\phi(P_0) + [b]\phi(Q_0)$. The prover defines the curves $E_2 = E_0/\langle K_\psi\rangle$ and $E_3 = E_1/\langle\phi(K_\psi)\rangle = E_0/\langle K_\psi, K_\phi\rangle$, and uses Vélu's formulae to compute the following diagram.

$$\begin{tikzcd}
E_0 \arrow[r, "\phi"] \arrow[d, "\psi"'] & E_1 \arrow[d, "\psi'"] \\
E_2 \arrow[r, "\phi'"] & E_3
\end{tikzcd}$$

The prover sends commitment $\mathsf{com} = (E_2, E_3)$ to the verifier.

2. The verifier challenges the prover with a random bit $\mathsf{chall} \leftarrow \{0, 1\}$.
3. If $\mathsf{chall} = 0$, the prover reveals $\mathsf{resp} = (K_\psi, \phi(K_\psi) = K_{\psi'})$.

   If $\mathsf{chall} = 1$, the prover reveals $\mathsf{resp} = (\psi(K_\phi) = K_{\phi'})$.

In both cases, the verifier accepts the proof if the points revealed have the correct order and generate kernels of isogenies between the correct curves. We iterate this process t times to reduce the cheating probability (where t is chosen based on the security parameter κ).

Note that in an honest execution of the proof, we have

$$\widehat{\psi'} \circ \phi' \circ \psi = [\ell_2^{e_2}]\phi.$$

3.2 Issue with soundness proofs for the De Feo-Jao-Plût scheme

A core component of the security proof of the De Feo-Jao-Plût identification scheme is the soundness proof. A proof of soundness was given by multiple previous works [DJP14, YAJ+17, GPS20] based on the CSSI problem in Definition 2. A sketch of this soundness proof is as follows:

Suppose $\mathcal{A}$ is an adversary that takes as input the public key and succeeds in the identification protocol (all t iterations) with noticeable probability ϵ. Given a challenge instance $(E_0, E_1, R_2, S_2, \phi(R_2), \phi(S_2))$ for the CSSI problem, we run $\mathcal{A}$ on the tuple $(E_1, \phi(R_2), \phi(S_2))$ as the public key. In the first round, $\mathcal{A}$ outputs commitments $(E_{i,2}, E_{i,3})$ for $1 \le i \le t$. We then send a challenge $b \in \{0, 1\}^t$ to $\mathcal{A}$ and, with probability ϵ, $\mathcal{A}$ outputs a response that satisfies the verification algorithm. Now, we use the standard replay technique: Rewind $\mathcal{A}$ to the point where it had output its commitments and then respond with a different challenge $b' \in \{0, 1\}^t$. With probability ϵ, $\mathcal{A}$ outputs a valid response. This gives exactly the 2-special soundness requirement of two valid transcripts with the same commitment but different challenges.

Now, choose some index i such that $b_i \neq b'_i$. We now restrict our focus to the components (E_2, E_3) for that index, and the two responses. It means $\mathcal{A}$ sent E_2, E_3 and can answer both challenges $b = 0$ and $b = 1$ successfully. Hence $\mathcal{A}$ has provided the maps ψ, ϕ', ψ' in the following diagram.

$$\begin{tikzcd}
E_0 \arrow[r, dashed, "\phi"] \arrow[r, bend right, "\tilde{\phi}"'] \arrow[d, "\psi"'] & E_1 \arrow[d, "\psi'"] \\
E_2 \arrow[r, "\phi'"] & E_3
\end{tikzcd}$$

The argument proceeds as follows: We have an explicit description of an isogeny $\tilde{\phi} = \widehat{\psi'} \circ \phi' \circ \psi$ from E_0 to E_1. The degree of $\tilde{\phi}$ is $\ell_1^{e_1}\ell_2^{2e_2}$. One can determine $\ker(\tilde{\phi}) \cap E_0[\ell_1^{e_1}]$ by iteratively testing points in $E_0[\ell_1^j]$ for $j = 1, 2, \ldots$. Hence, one determines the kernel of ϕ, as desired.

However, the important issue with this argument which has so far gone unnoticed, is that it assumes $\ker(\phi) = \ker(\tilde{\phi}) \cap E_0[\ell_1^{e_1}]$. This assumption has no basis, and we will provide a simple counterexample to this argument in the following section. While we always recover an isogeny, it may not be ϕ at all - it is entirely possible the isogeny we recover does not even have codomain E_1 so this proof of 2-special soundness is not valid.

3.3 Counterexample to soundness

Fix a supersingular curve E_0 as above. Generate a random $\ell_2^{e_2}$-torsion point $K_\psi \in E_0(\mathbb{F}_{p^2})$ as $K_\psi = [a]P_2 + [b]Q_2$ for some integers $0 \le a, b < \ell_2^{e_2}$ not both divisible by ℓ_2. Let $\psi : E_0 \to E_2$ have kernel generated by K_ψ. Then choose a random isogeny $\phi' : E_2 \to E_3$ of degree $\ell_1^{e_1}$ with kernel generated by $K_{\phi'}$. Then choose a random isogeny $\psi' : E_3 \to E_1$ of degree $\ell_2^{e_2}$. Choose points $P_2', Q_2' \in E_1(\mathbb{F}_{p^2})$ such that $\ker(\widehat{\psi'}) = \langle [a]P_2' + [b]Q_2' \rangle$. Then publish

$$(E_0, E_1, P_2, Q_2, P_2', Q_2')$$

as a public key. In other words, we have

$$E_0 \xrightarrow{\psi} E_2 \xrightarrow{\phi'} E_3 \xrightarrow{\psi'} E_1$$

Now there is no reason to believe that there exists an isogeny from E_0 to E_1 of degree $\ell_1^{e_1}$, yet we can respond to both challenge bits 0 and 1 in a single round of the identification scheme. Pulling back the kernel of ϕ' via ψ to E_0 will result in the kernel of an isogeny which, in general, will not have codomain E_1 (but instead a random other curve). This is because ψ' is entirely unrelated to ψ in this case (they are not "parallel"), so we have no SIDH square.

The key observation is that a verifier could be fooled into accepting this public key by a prover who always uses the same curves (E_2, E_3) instead of randomly chosen ones. When $b = 0$ the prover responds with the pair (a, b) corresponding to the kernel of ψ and $\widehat{\psi'}$, and when $b = 1$ the prover responds with $K_{\phi'}$. The verifier will agree that all responses are correct and will accept the proof.

The reader may immediately have several thoughts:

1. This is not the correct protocol description, since the isogenies ψ and ψ' are supposed to be random. The verifier can check if the same commitments (E_2, E_3) are always being re-used.
2. This scheme would not be zero-knowledge. If the protocol is repeated many times with the same pair (E_2, E_3) then the composition $\psi' \circ \phi' \circ \psi$ will be revealed to the verifier, leaking an isogeny from E_0 to E_1 and therefore allowing the verifier to impersonate the prover in the future.
3. Proving identity (or forging signatures) still requires knowledge of *some* isogeny from E_0 to E_1. So we can rescue the security proof by basing security on the general isogeny problem (Definition 1) instead of the SIDH problem.
4. The SIDH assumption as stated claims that an isogeny from E_0 to E_1 of degree $\ell_1^{e_1}$ exists, and asks to compute it. So surely that prevents the "attack" as well.

In response we say:

1. It is true that the verifier could test if the commitments (E_2, E_3) are being re-used, but this has never been stated as a requirement in any of the protocol descriptions. To tweak the verification protocol we need to know how "random" the pairs (E_2, E_3) (or, more realistically, the pairs (a, b)) need to be.

2. It is true that repeating (E_2, E_3) means the protocol is no longer zero knowledge. But soundness and zero-knowledge are independent security properties that are proved separately (and affect different parties: one gives an assurance to the verifier and the other to the prover). Our counterexample is a counterexample to the soundness proof. The fact that the counterexample is not consistent with the proof that the protocol is zero knowledge is irrelevant.

3-4. It is true that we could instead base security of the protocol on the general isogeny problem. Interestingly, none of the previous authors chose to do it that way. But some applications may require using the identification/signature protocols to prove that an SIDH public key is well-formed. For such applications we need soundness to be rigorously proved. The issue in the security proofs in the literature is not only that it is implicitly assumed that there is an isogeny of degree $\ell_1^{e_1}$ between E_0 and E_1. The key issue is that it is implicitly assumed that the pullback under ψ of $\ker(\phi')$ is the kernel of this isogeny. Our counterexample calls these assumptions into question, and shows that the proofs are incorrect as written down.

To make this very clear, consider the soundness proof from De Feo, Jao and Plût [DJP14]. The following diagram is written within the proof. It implicitly assumes that the horizontal isogeny ϕ' has kernel given by $\psi(S)$, so that the image curve is $E/\langle S, R\rangle$.

$$\begin{tikzcd}
E \arrow[d, "\psi"'] & E/\langle S\rangle \arrow[d, "\psi'"] \\
E/\langle R\rangle \arrow[r, "\phi'"'] & E/\langle S, R\rangle
\end{tikzcd}$$

This implicit assumption seems to have been repeated in all subsequent works, such as [YAJ+17] and [GPS20].

Note: we conjecture the original scheme to be secure despite the issue with the proof, as long as the commitment E_2, E_3 is not reused every time (point 1. above). This simple check can be added into the original scheme.

Conjecture 1. The De Feo-Jao-Plût identification scheme with t rounds (for appropriately chosen t) is sound, assuming the CSSI problem is hard, when the additional requirement that commitments (E_2, E_3) are not reused between rounds.

We leave proof of this conjecture for future work.

4 New SIDH identification scheme

Let public parameters $pp = (p, E_0, P_0, Q_0)$ such that $E_0(\mathbb{F}_p)[\ell_2^{e_2}] = \langle P_0, Q_0\rangle$. As before, suppose a user has a secret isogeny $\phi : E_0 \to E_1$ with kernel $\ker\phi = K_\phi$. Without loss of generality we assume that the secret isogeny has degree $\ell_1^{e_1}$. We propose a new sigma protocol to prove knowledge of this isogeny given the public key $(E_1, P_1 = \phi(P_0), Q_1 = \phi(Q_0))$. The protocol is presented in Figure 3. $\mathsf{IsogenyFromKernel}()$ is a function taking a kernel point and outputting an isogeny and codomain curve with said kernel. $\mathsf{CanonicalBasis}()$ is a function taking a curve and outputting a canonical $\ell_2^{e_2}$ torsion basis on the given curve. Figure 2 shows the commutative diagram of the sigma protocol.

Intuitively, the identification scheme follows 3.1, with a single bit challenge - if the challenge is 0, we reveal the vertical isogenies ψ, ψ', while if the challenge is 1, we reveal the horizontal ϕ'. The difference is the introduction of additional points on E_3 to the commitment, which force ψ, ψ' to be, in some sense "compatible" or "parallel". This restriction allows the proof of 2-special soundness to work.

We then repeat the identification scheme t times in parallel (where t is chosen based on the security parameter κ) and set com to be the concatenation of all individual $[\mathsf{com}_i]_{i\in[t]}$ for each iteration i, $\mathsf{chall} = [\mathsf{chall}_i]_{i\in[t]}$ and $\mathsf{resp} = [\mathsf{resp}_i]_{i\in[t]}$.

$$\begin{tikzcd} E_0 \arrow[r, dashed, "\phi"] \arrow[d, "\psi"'] & E_1 \arrow[d, "\psi'"] \\ E_2 \arrow[r, "\phi'"'] & E_3 \end{tikzcd}$$

Figure 2: Commutative diagram of SIDH identification scheme

Note: Verification requires checking that there exists integers c, d generating the kernels of dual isogenies $\widehat{\psi}, \widehat{\psi}'$. This computation can be offloaded to the prover by requiring them to send the correct integers. In fact, these integers uniquely determine the horizontal isogenies so they could be sent as resp by the prover without needing $K_\psi, K_{\psi'}$, but this would require more computation to verify.

Remark 2. There are certainly improvements that can be made to improve efficiency and compress the size of signatures, but these are standard and we will not explore them here. For example, in practice the commitment information (E_3, P_3, Q_3) would be replaced with a triplet of x-coordinates, as in SIKE [ACC$^+$17].

Theorem 1. *The sigma protocol in Figure 3 is complete, 2-special sound, and zero knowledge, assuming the CSSI and DSSP problems are hard.*

Proof. We prove the three properties of Theorem 1 separately below.

Completeness: It is clear that following the protocol honestly will result in an accepting transcript.

2-special soundness: Suppose we have an adversarial prover P^* which, given a public key, succeeds in the identification protocol (with t parallel rounds) with noticeable probability ϵ. By rewinding P^* after obtaining an output, and providing a different commitment, we can obtain two accepting transcripts $(\mathsf{com}, \mathsf{chall}, \mathsf{resp})$ and $(\mathsf{com}, \mathsf{chall}', \mathsf{resp}')$ with $\mathsf{chall} \neq \mathsf{chall}'$. Note that the probability of two random challenges of length t-bits being equal is $1/2^t$ which is the soundness error. Consider one of the t rounds i where the challenge bit chall_i differs. The secret isogeny corresponding to the public key (E_1, P_1, Q_1) can be recovered as follows, hence $\mathsf{Extract}$, given access to P^*, can extract a valid witness (and break the CSSI problem).

Without loss of generality, suppose $\mathsf{chall}_i = 0$ and $\mathsf{chall}'_i = 1$. Then recover (a, b) and thus $(K_\psi, K_{\psi'})$ from resp_i, and $K_{\phi'}$ from resp'_i. Compute the dual isogeny $\widehat{\psi}$ and use this to pull the kernel $K_{\phi'}$ back to E_0 (this works because the degrees of $K_{\phi'}$ and $\widehat{\psi}$ are coprime). Let φ be the isogeny with kernel $\langle K_\varphi = \widehat{\psi}(K_{\phi'})\rangle$, so that $\varphi : E_0 \to E_0/\langle K_\varphi\rangle$.

We first demonstrate that $E_0/\langle K_\varphi\rangle \cong E_1$. This follows by considering the diagram of Figure 2 as an SIDH square starting from base curve E_2. We have that $E_1 \cong E_2/\langle K_{\phi'}, G\rangle$ for subgroup G of order $\ell_2^{e_2}$ such that $\phi'(G) = \ker \widehat{\psi}'$. However, note that the restriction on the kernels of $\widehat{\psi}, \widehat{\psi}'$ force $\ker \widehat{\psi}' = \phi'(\ker \widehat{\psi})$ so $G = K_{\widehat{\psi}}$. Thus, $E_0 \cong E_2/\langle G = K_{\widehat{\psi}}\rangle$ and commutativity implies φ exists and has the correct degree, and $E_1 \cong E_0/\langle K_\varphi\rangle$ as required. A perhaps simpler argument is that $\hat{\psi}' \circ \phi' \circ \psi$ is an isogeny from E_0 to E_1 that kills the entire $\ell_2^{e_2}$ torsion $E_0[\ell_2^{e_2}]$ so must factor through $[\ell_2^{e_2}]$. Hence there is a degree $\ell_1^{e_1}$ isogeny from E_0 to E_1.

The correctness of the points P_1, Q_1 is, however, much more subtle. Unfortunately we know of no efficient way to prove the points are exactly the image of the public parameter points. What we can show, though,

round 1

$a, b \leftarrow \mathbb{Z}/\ell_2^{e_2}\mathbb{Z}$ $\triangleright$ N.B. we can use equivalent keys (Rem. 1) for compactness WLOG
$K_\psi = [a]P_0 + [b]Q_0 \in E_0$
$K_{\psi'} = \phi(K_\psi) = [a]\phi(P_0) + [b]\phi(Q_0) \in E_1$
$\psi, E_2 \leftarrow \mathsf{IsogenyFromKernel}(K_\psi)$
$P_2, Q_2 \leftarrow \mathsf{CanonicalBasis}(E_2)$
$K_{\phi'} \leftarrow \psi(K_\phi) \in E_2$
$\phi', E_3 \leftarrow \mathsf{IsogenyFromKernel}(K_{\phi'})$
$P_3, Q_3 \leftarrow \phi'(P_2), \phi'(Q_2) \in E_3$
Prover sends $\mathsf{com} = (E_2, E_3, P_3, Q_3)$ to Verifier.

round 2

$c \leftarrow \{0, 1\}$
Verifier sends $\mathsf{chall} \leftarrow c$ to Prover.

round 3

$c \leftarrow \mathsf{chall}$
if $c = 1$ **then**
    $\mathsf{resp} \leftarrow K_{\phi'}$
else
    $\mathsf{resp} \leftarrow (a, b)$
Prover sends resp to Verifier.

Verification

$(E_2, E_3, P_3, Q_3, c) \leftarrow (\mathsf{com}, \mathsf{chall})$
if $c = 1$ **then**
    $K_{\phi'} \leftarrow \mathsf{resp}$
    Check $K_{\phi'}$ has order $\ell_1^{e_1}$ and lies on E_2, otherwise output reject
    $P_2, Q_2 \leftarrow \mathsf{CanonicalBasis}(E_2)$
    $\phi', E_3' \leftarrow \mathsf{IsogenyFromKernel}(K_{\phi'})$
    Verify $E_3 = E_3'$ and $P_3, Q_3 = \phi'(P_2), \phi'(Q_2)$, otherwise output reject
    Verifier outputs accept
else
    $(a, b) \leftarrow \mathsf{resp}$
    Check that $P_1, Q_1 \in E_1$
    $K_\psi, K_{\psi'} = [a]P_i + [b]Q_i$ for $i = 0, 1$ resp.
    Check K_ψ and $K_{\psi'}$ have order $\ell_2^{e_2}$, otherwise output reject
    $\psi, E_2' \leftarrow \mathsf{IsogenyFromKernel}(K_\psi)$
    $\psi', E_3' \leftarrow \mathsf{IsogenyFromKernel}(K_{\psi'})$
    Check $E_2 = E_2'$ and $E_3 = E_3'$
    $P_2, Q_2 \leftarrow \mathsf{CanonicalBasis}(E_2)$
    Check there exists $c, d \in \mathbb{Z}/\ell_2^{e_2}\mathbb{Z}$ such that, simultaneously,
    - $\ker \widehat{\psi} = [c]P_2 + [d]Q_2$
    - $\ker \widehat{\psi}' = [c]P_3 + [d]Q_3$
    Verifier outputs accept if the preceding conditions hold, otherwise reject

Figure 3: One iteration of the sigma protocol for our new SIDH identification scheme. The public parameters are $pp = (p, \ell_1, \ell_2, e_1, e_2, E_0, P_0, Q_0)$. The public key is (E_1, P_1, Q_1), and the corresponding secret isogeny is ϕ.

is that the choice of points P_1, Q_1 by the adversary is severely restricted if they must keep them consistent with "random enough" values of a, b (i.e., random choices of ψ). So while the choice of points is certainly not unique, we show in Section 5 that the adversary is still prevented from choosing the points in the malicious way required for known adaptive attack methods.

Thus we recover an isogeny φ of correct degree $\ell_1^{e_1}$ such that the codomain is isomorphic to E_1 and the points in the public key are correct up to scalar. This shows the protocol is 2-special sound when the CSSI problem is hard, and that it is a Proof of Knowledge of an isogeny corresponding to the given public key (which also proves the public key is honestly generated).

Zero-knowledge: Proof of ZK follows as in [DJP14]. Let V^* be a cheating verifier, which shall be used as a black box by the simulator Sim. We shall show that Sim can generate a valid transcript for t iterations of the protocol. At each step, Sim makes a guess what the next challenge bit chall will be, and then proceeds as follows.

- If $\mathsf{chall} = 0$, Sim simulates correctly by choosing $a, b \leftarrow \mathbb{Z}/\ell_2^{e_2}\mathbb{Z}$ and computing the two vertical isogenies $\psi : E_0 \to E_2, \psi' : E_1 \to E_3$ from kernels $K_\psi = [a]P_0 + [b]Q_0$ and $K_{\psi'} = [a]P_1 + [b]Q_1$. Also compute the kernels of the corresponding dual isogenies, and the canonical basis $P_2, Q_2 \leftarrow \mathsf{CanonicalBasis}(E_2)$. Write $K_{\widehat{\psi}}$ in terms of this basis as $[c]P_2 + [d]Q_2$, then choose a torsion basis on E_3 as $P_3, Q_3 \in E_3$ such that $K_{\widehat{\psi'}} = [c]P_3 + [d]Q_3$. We set the commitment to $\mathsf{com} = (E_2, E_3, P_3, Q_3)$ and the response to $\mathsf{resp} = (a, b)$ as required.
- If $\mathsf{chall} = 1$ we choose a random curve E_2 and a random point point $K \in E_2$ of correct order $\ell_1^{e_1}$. Then compute the isogeny $\phi' : E_2 \to E_3$ with kernel K with $\mathsf{IsogenyFromKernel}$. Finally generate a canonical basis $P_2, Q_2 \leftarrow \mathsf{CanonicalBasis}(E_2)$ and compute $P_3, Q_3 \leftarrow \phi'(P_2), \phi'(Q_2)$ and set the commitment to (E_2, E_3, P_3, Q_3).

After providing com to V^*, if the challenge V^* outputs is not the same as Sim's guess, simply discard that iteration and run again. Sim stops whenever V^* rejects or after t successful rounds. Suppose the probability of V^* not choosing the same bit as Sim's guess is noticeably different from $1/2$. Then V^* can be used as a distinguisher for the DSSP problem. So the probability Sim guesses correctly each round is exponentially close to $1/2$ if the DSSP problem is hard. Thus Sim will run in polynomial time.

To prove indistinguishability of simulated transcripts from true interactions of a prover P with V^*, it is enough to show that one round of the sigma protocol is indistinguishable (by the hybrid technique of Goldreich et al. [GMW91]).

When $\mathsf{chall} = 0$, the outputs of the simulator are identical to those generated according to the protocol, so the distributions are perfectly indistinguishable. Note that P_3, Q_3 are uniquely determined to be the points from the true protocol up to scalar multiple, which is possible despite not knowing ϕ' because P_1, Q_1 allow us to "jump" the gap.

When $\mathsf{chall} = 1$, we consider the distribution of (E_2, E_3, ϕ'). While this distribution is not correct a priori, the computational assumption in Definition 3 implies it is computationally hard to distinguish the simulation from the real game (as in the proof in [GPS20]). Because providing the action of ϕ' on canonical basis $P_2, Q_2 \in E_2$ cannot reveal more information than providing ϕ' itself, the distribution of (E_2, E_3, P_3, Q_3) must also be indistinguishable between simulation and real transcripts. Hence the scheme has computational zero knowledge assuming the DSSP problem is hard. $\square$

Computing a compatible basis. In the proof of ZK, we require choosing a basis P_3, Q_3 for the torsion $E_3[\ell_2^{e_2}]$ such that for a fixed $K \in E_3[\ell_2^{e_2}]$ and fixed integers c, d, $K = [c]P_3 + [d]Q_3$. This can be done efficiently due to the ease of computing discrete logarithms when p is very smooth [Tes99]. For example, simply choose any basis P', Q' and write $K = [c']P' + [d']Q'$ by solving discrete logarithms with respect to P', Q'. The shift between bases P', Q' and $P_3 = [w]P' + [x]Q', Q_3 = [y]P' + [z]Q'$ must be linear, so we can write $K = [c]P_3 + [d]Q_3 = [cw + dy]P' + [cx + dz]Q'$. Then simply solve the linear system $cw + dy = c'$, $cx + dz = d'$ for w, x, y, z and compute P_3, Q_3.

5 Correctness of the points in an SIDH public key

We have shown in Section 4 that successful completion of the new sigma protocol indeed proves knowledge of a degree $\ell_1^{e_1}$ isogeny from E_0 to E_1. However, an SIDH public key also consists of the two torsion points, and these points are the cause of issues such as the adaptive attack [GPST16].

We have that $\ker\psi = \langle K_\psi\rangle = \langle [a]P_0 + [b]Q_0\rangle$ for the fixed points P_0, Q_0. This choice of ψ also fixes $\widehat{\psi}$. Now, $\ker\widehat{\psi}' = \phi'(\ker\widehat{\psi})$ as before, so $\widehat{\psi}'$ is also fixed, and by extension ψ'. Finally then, to ensure verification succeeds, the adversary must choose $P', Q' \in E_1$ such that $\langle [a]P' + [b]Q'\rangle = \langle K_{\psi'}\rangle$ for the same a, b as before. For a single choice of a, b, there are many ways to decompose $\ker\psi'$ in terms of two basis points. The key observation though, is that once these points have been fixed in the first iteration of the sigma protocol, all future iterations must use the same two points, but answer with different (a, b) values. If the verifier checks that these (a, b) values are "random enough" whenever they are revealed (challenge bit 0), the prover is restricted in their choice of points as we will see below.

So, as stated above, the prover is in a position where they have a fixed kernel $\langle K_{\psi'}\rangle$. Obviously, the "honest" behaviour will give kernel generator $K_{\psi'} = [a]\phi(P_0) + [b]\phi(Q_0)$. Two generators generate the same kernel if and only if they are (invertible) scalar multiples of each other. Hence, we consider the case where the adversary wishes to decompose any arbitrary kernel generator K' such that $[\lambda]K' = K_{\psi'}$ in terms of a, b, that is, $[a]\phi(P_0) + [b]\phi(Q_0) = [a][\lambda]P' + [b][\lambda]Q'$. For ease of notation, let $P = \phi(P_0), Q = \phi(Q_0)$.

Because both P, Q and P', Q' are bases of the same torsion subgroup, we can represent P', Q' in terms of P, Q with a change-of-basis matrix. This matrix must be invertible, so $cf - de$ must be invertible modulo $\ell_2^{e_2}$.

$$\begin{pmatrix} P' \\ Q' \end{pmatrix} = \begin{pmatrix} c & d \\ e & f \end{pmatrix} \cdot \begin{pmatrix} P \\ Q \end{pmatrix} \tag{1}$$

Now because P and Q are linearly independent, we can match coefficients (modulo the order of the generators) and obtain the following two congruences:

$$\begin{aligned} a &\equiv a\lambda c + b\lambda e \pmod{\ell_2^{e_2}} \\ b &\equiv a\lambda d + b\lambda f \pmod{\ell_2^{e_2}} \end{aligned}$$

Giving:

$$0 \equiv a(\lambda c - 1) + b\lambda e \pmod{\ell_2^{e_2}} \tag{2}$$
$$0 \equiv a\lambda d + b(\lambda f - 1) \pmod{\ell_2^{e_2}} \tag{3}$$

Because P', Q' are published by the prover before beginning the protocol, c, d, e, f are all fixed. We now add the restriction that the verifier confirms the a, b's cover the following three congruency classes modulo ℓ_2 (note that at least one of a, b must not be divisible by ℓ_2 for the kernel to have the correct order):

$$\begin{aligned} a \equiv 0, b \not\equiv 0 &\pmod{\ell_2} \\ a \not\equiv 0, b \equiv 0 &\pmod{\ell_2} \\ a, b \not\equiv 0 &\pmod{\ell_2} \end{aligned}$$

For ease of notation, we will denote these three cases as $(0, \star), (\star, 0)$, and $(\star, \star)$ respectively. It is clearly implied that $\ell_2 \nmid \star$.

From here forward, for ease of notation, we will treat all values modulo $\ell_2^{e_2}$ as integers in the range $0, \ldots, \ell_2^{e_2} - 1$. If the prover convinces the verifier that with overwhelming probability (in the security parameter κ) they can answer queries using all three classes of a, b above, then it must be the case that $e = d = 0$ and $c = f$ invertible. This indeed proves that the points P', Q' are simply an invertible scalar multiple of the original points $P' = [\lambda]P, Q' = [\lambda]Q$. This is sufficient to ensure there is no possibility of an adaptive attack being performed. In fact, using the Weil pairing check from Galbraith et al. [GPST16] as well, we can force the only choices for this scalar to be $\lambda = \pm 1$ (but we don't need this extra restriction so we won't discuss this further).

To set some notation, we use $\ell^n \parallel x$ to denote that ℓ^n divides x, but ℓ^{n+1} does not divide x. That is, ℓ^n is the highest power of ℓ dividing x. In this case, we say ℓ^n *exactly divides* x.

Theorem 2. *For a fixed security parameter κ and SIDH public key (E, P, Q), if the prover is able to successfully complete 3κ iterations of the identification scheme sigma protocol in Figure 3 as follows:*

- *κ iterations where the prover uses non-repeating challenges (a, b) for $a, b \not\equiv 0 \pmod{\ell_2}$ - case $(\star, \star)$,*
- *κ iterations where the prover uses non-repeating challenges (a, b) for $a \not\equiv 0, b \equiv 0 \pmod{\ell_2}$ - case $(\star, 0)$, and*
- *κ iterations where the prover uses non-repeating challenges (a, b) for $a \equiv 0, b \not\equiv 0 \pmod{\ell_2}$ - case $(0, \star)$*

then with probability $1 - 2^{-\kappa}$ the points P, Q are of the form $[\lambda]\phi(P_0), [\lambda]\phi(Q_0)$ for some invertible scalar λ (where ϕ is a secret $\ell_1^{e_1}$-isogeny $E_0 \to E$).

Proof. We fix c, d, e, f and suppose the prover is able to commit to and successfully answer challenges for (a, b) tuples in all three of the classes above.

If $a \equiv 0, b \not\equiv 0 \pmod{\ell_2}$, then Equation 2 implies that $e \equiv 0 \pmod{\ell_2}$, while Equation 3 requires $f \not\equiv 0 \pmod{\ell_2}$. Similarly, if Equations 2 and 3 are able to be satisfied by a, b where $a \not\equiv 0, b \equiv 0 \pmod{\ell_2}$, we get that $d \equiv 0 \pmod{\ell_2}$ and $c \not\equiv 0 \pmod{\ell_2}$.

In the simplest case, $e = d = 0$. Requiring Equations 2 and 3 to have solutions of the form $(\star, \star)$ (i.e. $a, b \not\equiv 0 \pmod{\ell_2}$) immediately implies that $\lambda c - 1 \equiv \lambda f - 1 \equiv 0 \pmod{\ell_2^{e_2}}$. Hence, $c = f$. This case is the "honest prover" scenario where the points P', Q' the prover provides in the public key are the same as the correct image points $\phi(P_0), \phi(Q_0)$ under the prover's secret isogeny, up to (co-prime) scalar multiple.

It remains to show, then, that being able to satisfy Equations 2 and 3 with (a, b) pairs across all three of the equivalence classes above force $e = d = 0$ - that they cannot be non-zero multiples of ℓ_2. We therefore proceed with a proof by contradiction. Let

$$d = d'\ell_2^g$$
$$e = e'\ell_2^h$$

where g, h are the greatest powers of ℓ_2 dividing d, e (respectively infinite if d or e is 0). Without loss of generality, we can assume that $h \geq g$, because otherwise we can swap the variables $(a, c, e) \leftrightarrow (b, f, d)$. Because we assume that at least one of e, d are non-zero, then this convention implies d (and so too d') is non-zero, while e (and e') may or may not be zero. Note that by definition, $\ell_2 \nmid d'$ and if $e' \neq 0$, then $\ell_2 \nmid e'$.

If (a, b) tuples of the form $(a, b) \equiv (\star, \star) \pmod{\ell_2}$ are able to satisfy Equation 3, then

$$\ell_2^g \parallel 1 - \lambda f$$

By considering Equation 2, we also get that

$$\ell_2^h \mid 1 - \lambda c$$

(if $e \neq 0$ this divisibility is exact, while if $e = 0$, $1 - \lambda c$ must also be 0). Because $g \leq h$, clearly $\ell_2^g \mid 1 - \lambda c$. Then,

$$\begin{aligned} 1 - \lambda f &\equiv 0 \pmod{\ell_2^g} \\ 1 - \lambda c &\equiv 0 \pmod{\ell_2^g} \\ (1 - \lambda f) - (1 - \lambda c) &\equiv 0 \pmod{\ell_2^g} \\ \lambda f &\equiv \lambda c \pmod{\ell_2^g} \end{aligned}$$

so we have that $c \equiv f \pmod{\ell_2^g}$.

Now suppose Equations 2 and 3 can be satisfied by $(a, b) \equiv (\star, 0) \pmod{\ell_2}$ as well. Because $\ell_2 \nmid a\lambda d'$, Equation 3 gives:

$$\ell_2^g \parallel b(1 - \lambda' f) \tag{4}$$

We also obtain from Equation 2 that:

$$\ell_2^h \mid 1 - \lambda' c$$

From this, using the fact that $c \equiv f \pmod{\ell_2^g}$ from the $(\star, \star)$ case, and that $g \leq h$, we get

$$\begin{aligned} \ell_2^g &\mid 1 - \lambda' c \\ \ell_2^g &\mid 1 - \lambda' c - \lambda(f - c) \\ \ell_2^g &\mid 1 - \lambda' f \end{aligned}$$

However, if $\ell_2^g \mid 1 - \lambda' f$ and $\ell_2 \mid b$, then

$$\ell_2^{g+1} \mid b(1 - \lambda' f)$$

Which contradicts Equation 4 by definition of exact divisibility.

Thus, if (a, b) tuples of both forms $(\star, \star)$ and $(\star, 0)$ modulo ℓ_2 are able to satisfy Equations 2 and 3, then necessarily $d = 0$ (and by extension of our assumption $h \geq g$, $e = 0$). To remove the assumption that $h \geq g$, we simply require that tuples of the form $(0, \star)$ are also satisfiable (due to the $(a, c, e) \leftrightarrow (b, f, d)$ variable swap). This concludes the proof. The probability given in the theorem follows trivially from the fact that, as in the original SIDH identification scheme, κ iterations convinces the verifier that the prover can answer each type of case except with probability $2^{-\kappa}$ each time. Hence, we treat each of the three cases as independent proofs and require 3κ iterations overall.

□

Remark 3. While 3κ is the trivial requirement to ensure the prover can indeed answer all three forms of (a, b) with overwhelming probability, we believe κ-bit security can be achieved with a more efficient choice. However, more thorough analysis is needed. We leave this for future work.

Because 3κ iterations of the sigma protocol are used rather than κ, this protocol will result in transcripts 3 times larger than those from Figure 3, when proving the correctness of the points is important.

In terms of the protocol in Figure 3, verification only requires on extra check: in the case that $c = 0$ (the **else** clause of the verification algorithm), after extracting (a, b) from resp, the verifier simply keeps track of how many of each case $(\star, \star)$, $(\star, 0)$ and $(0, \star)$ are seen, and accepts overall only if the number of each case is roughly equal (or perhaps, for example, the first κ iterations must match $(\star, \star)$, the second $(\star, 0)$ and the third $(0, \star)$ if 3κ iterations are used).

6 SIDH signatures and Non-Interactive Proof of Knowledge

We conclude with some brief, standard remarks about the use of the new protocol proposed above.

It is standard to construct a non-interactive signature scheme from an interactive protocol using the Fiat-Shamir transformation (secure in the (quantum) random oracle model [LZ19]). This works by making the challenge chall for the t rounds of the ID scheme a random-oracle output from input the commitment com and a message M. That is, for message M,

$$V_1^{\mathcal{O}}(\mathsf{com}) = \mathcal{O}(\mathsf{com} \parallel M)$$

Thus the prover does not need to interact with a verifier and can compute a non-interactive transcript. Because the sigma protocol described in the preceding sections not only proves knowledge of the secret isogeny between two curves, but also correctness of the torsion points in the public key, we obtain a signature scheme that is also a proof of knowledge of the secret key corresponding to a given SIDH public key, and proves that the SIDH public key is well-formed. For example, simply signing the public key with its own secret key using the new scheme gives a simple NIZK proof of well-formedness for the public key which ensures against adaptive attacks.

References

[ACC+17] Reza Azarderakhsh, Matthew Campagna, Craig Costello, Luca De Feo, Basil Hess, Amir Jalali, David Jao, Brian Koziel, Brian LaMacchia, Patrick Longa, et al. Supersingular isogeny key encapsulation. *Submission to the NIST Post-Quantum Standardization project*, 2017.

[BKM+20] Andrea Basso, Péter Kutas, Simon-Philipp Merz, Christophe Petit, and Charlotte Weitkämper. On adaptive attacks against jao-urbanik's isogeny-based protocol. In Abderrahmane Nitaj and Amr Youssef, editors, *Progress in Cryptology - AFRICACRYPT 2020*, pages 195–213, Cham, 2020. Springer International Publishing.

[CLN16] Craig Costello, Patrick Longa, and Michael Naehrig. Efficient algorithms for supersingular isogeny diffie-hellman. In *Annual International Cryptology Conference*, pages 572–601. Springer, 2016.

[DGL+20] Samuel Dobson, Steven D. Galbraith, Jason LeGrow, Yan Bo Ti, and Lukas Zobernig. An adaptive attack on 2-SIDH. *International Journal of Computer Mathematics: Computer Systems Theory*, 5(4):282–299, 2020.

[DJP14] 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.

[GMW91] Oded Goldreich, Silvio Micali, and Avi Wigderson. Proofs that yield nothing but their validity or all languages in np have zero-knowledge proof systems. *Journal of the ACM (JACM)*, 38(3):690–728, 1991.

[GPS20] Steven D. Galbraith, Christophe Petit, and Javier Silva. Identification protocols and signature schemes based on supersingular isogeny problems. *Journal of Cryptology*, 33(1):130–175, 2020.

[GPST16] Steven D. Galbraith, Christophe Petit, Barak Shani, and Yan Bo Ti. On the security of supersingular isogeny cryptosystems. In *Advances in Cryptology – ASIACRYPT 2016*, pages 63–91. Springer Berlin Heidelberg, 2016.

[JD11] David Jao and Luca De Feo. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In *Post-Quantum Cryptography*, pages 19–34, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[JS14] David Jao and Vladimir Soukharev. Isogeny-based quantum-resistant undeniable signatures. In *PQCrypto 2014*, volume 8772 of *Lecture Notes in Computer Science*, pages 160–179. Springer, 2014.

[LZ19] Qipeng Liu and Mark Zhandry. Revisiting post-quantum Fiat-Shamir. In *Annual International Cryptology Conference*, pages 326–355. Springer, 2019.

[Tes99] Edlyn Teske. The Pohlig–Hellman method generalized for group structure computation. *Journal of symbolic computation*, 27(6):521–534, 1999.

[UJ20] David Urbanik and David Jao. New techniques for SIDH-based NIKE. *Journal of Mathematical Cryptology*, 14(1):120–128, 2020.

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

[YAJ+17] Youngho Yoo, Reza Azarderakhsh, Amir Jalali, David Jao, and Vladimir Soukharev. A post-quantum digital signature scheme based on supersingular isogenies. In *International Conference on Financial Cryptography and Data Security*, pages 163–181. Springer, 2017.