Modeling Quantum-Safe Authenticated Key Establishment, and an Isogeny-Based Protocol

Jason LeGrow[1,2,3], David Jao[1,2,4], and Reza Azarderakhsh[5]

[1] Department of Combinatorics and Optimization,
University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada
[2] Centre for Applied Cryptographic Research,
University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada
[3] Institute for Quantum Computing,
University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada
[4] evolutionQ, Inc., Waterloo, Ontario, Canada
[5] Department of Computer and Electrical Engineering and Computer Science,
Florida Atlantic University, Boca Raton, Florida

Abstract. We propose a security model for authenticated key establishment in the quantum setting. Our model is the first for authenticated key establishment that allows for quantum superpositions of queries. The model builds on the classical Canetti-Krawczyk model but allows quantum interactions between the adversary and quantum oracles that emulate classical parties. We demonstrate that this new security definition is satisfiable by giving a generic construction from simpler cryptographic primitives and a specific protocol which is secure in the quantum random oracle model, under the supersingular isogeny decisional Diffie-Hellman assumption (SIDH).

1 Introduction

Key establishment is a fundamental cryptographic primitive which allows parties communicating over an insecure but authenticated channel (that is, a channel whose messages can be eavesdropped on, but whose origin and authenticity are guaranteed) to establish a secure common cryptographic key, and hence establish a secure communication channel. Of course, modern Internet communications are not inherently authenticated, and so more sophisticated *authenticated* key establishment protocols are required to establish secure keys in this regime; in particular, ubiquitous Internet communication protocol suites such as SSL/TLS and IPsec include a wide array of authenticated key establishment protocols.

Though public-key key establishment protocols date back at least to the Diffie-Hellman protocol [9], the first formal security analysis of authenticated key establishment was due to Bellare and Rogaway [4]; since this seminal work, many extensions and modifications have been made to the Bellare-Rogaway security model in the classical setting [3,7,16], and some work has been done to model quantum protocols in a similar regime [18]. Despite the breadth of work that has been done on modelling the security of authenticated key establishment

protocols, there has been little explicit consideration of quantum-safe security of authenticated key establishment protocols, beyond using quantum- safe computational assumptions. In particular—and in stark contrast with the situation of encryption [1,6,11] and signature schemes [6]—to the best of our knowledge no security model for authenticated key establishment has been proposed which allow for quantum *interactions* between the adversary and key-establishing parties. The first major result of this work is to establish a formal framework for truly quantum-safe security of authenticated key establishment protocols. It is important that we understand how to analyze the security of our protocols in the quantum setting to face the looming threat of quantum computers.

Of course, ideally we would like to have a protocol for authenticated key establishment which is secure in our quantum-safe model. To this end, we develop a generic construction which combines the simpler primitives of key establishment and digital signatures to obtain a secure authenticated key establishment protocol. We then apply our generic construction to isogeny-based key establishment and signatures to obtain a truly quantum-safe authenticated key establishment protocol, which can in principle be used to create secure classical channels, even when faced with an adversary with significant quantum capabilities.

To reiterate, in this work we:

1. Give a novel quantum-safe security model and definition for authenticated key establishment;
2. Give a generic construction for secure protocols in this model, and;
3. Apply our generic construction to build an isogeny-based authenticated key establishment protocol.

Related work. There are a number of quantum-safe authenticated key establishment protocols based on standard security assumptions. In the realm of isogenies, [27] uses a standard signature-based construction to add authentication to the isogeny-based protocol due to Jao *et al.* Our protocol is reminiscent of this one, but uses a different signature scheme with security amplifications to achieve a much stronger security property than is shown in [27]. There is also a code-based construction [2] which uses a key encapsulation mechanism authenticated by a signature scheme. Again, this method is not known to achieve the strong security properties achieved by the protocol in this work. Arguably the most novel and satisfying previously-known solution, presented in [30], is based on the Ring Learning with Errors (Ring-LWE) problem in ideal lattices. This protocol does not follow the framework of the previous two protocols; parties do not encrypt or sign the messages of an unauthenticated key establishment protocol to produce an authenticated variant. The protocol has many desirable properties (in particular, because it does not use signatures its security is based solely on the hardness of the Ring-LWE problem; it has perfect forward secrecy, and it can be formulated as a one-pass scheme) and is extremely efficient; however, this protocol is not known to be quantum-safe in the sense we propose.

Structure of this work. In Sections 2 and 3 we detail the necessary cryptographic and algebrogeometric preliminaries for our security model, construction, and final protocol. Section 4 lists the constructions from previous works that we use to construct our protocol. Section 5 contains an explanation of the security model and Section 6 contains our generic construction using signature schemes. Section 7 contains an explanation of our specific scheme. Finally, we conclude and consider future avenues of study in Section 8.

2 Aspects of Post-Quantum Cryptography

2.1 Quantum Oracles and Post-Quantum Security Definitions

In classical cryptography, often an adversary is provided *oracle access* to a function. For instance, in the well-known EUF-CMA security definition for signature schemes, the adversary is given oracle access to the signing function $\mathrm{Sign}_{\mathrm{sk}}$; that is, the adversary is given a "black box" which takes in messages m and outputs a signature $\sigma = \mathrm{Sign}_{\mathrm{sk}}(m)$. In the quantum-safe setting, the adversary may instead be provided with a *quantum oracle* for a function. If $f\colon \{0,1\}^m \to \{0,1\}^n$ is a function, then a quantum oracle U_f for f acts on basis states as

$$U_f\colon \ |x\rangle\,|y\rangle \mapsto |x\rangle\,|y \oplus f(x)\rangle\,.$$

In the EUF-qCMA security definition for signature schemes [6], the adversary is given quantum oracle access to the signing function:

Definition 1 (Strong EUF-qCMA Security).
A signature scheme $(\mathrm{Sign}, \mathrm{Ver})$ *is* strongly existentially unforgeable against a quantum adaptive chosen message attack (strongly EUF-qCMA secure) *if the advantage that any polynomial-time adversary has at winning the following game is negligible:*

1. *The challenger* $\mathcal{C}$ *runs the key generation algorithm on input* 1^λ *to obtain the key pair* $(\mathrm{sk}, \mathrm{pk})$*, and publishes* pk*.*
2. *For* $i = 1, 2, \ldots, t$*:*
   a) *The adversary* $\mathcal{A}$ *sends* $|\psi_i\rangle = \sum_{m,y} \alpha_{m,y}\,|m\rangle\,|y\rangle$ *to* $\mathcal{C}$*.*
   b) $\mathcal{C}$ *returns* $U_{\mathrm{Sign}_{\mathrm{sk}}}\,|\psi_i\rangle = \sum_{m,y} \alpha_{m,y}\,|m\rangle\,|y + \mathrm{Sign}_{\mathrm{sk}}(m)\rangle$
3. $\mathcal{A}$ *produces* $(m_1^*, \sigma_1^*), (m_2^*, \sigma_2^*), \ldots, (m_{t+1}^*, \sigma_{t+1}^*)$*.* $\mathcal{A}$ *wins the game if the pairs* (m_i^*, σ_i^*) *are distinct and* $\mathrm{Ver}_{\mathrm{pk}}(m_i^*, \sigma_i^*) = 1$ *for* $1 \le i \le t+1$*.*

Quantum oracle access to encryption functions (for cryptosystems) and signing functions (for signature schemes) are standard in the study of quantum cryptanalysis (see, for instance, [6,11,14]). In Section 5 we present a security model for authenticated key establishment that allows the analogous interaction with key establishing parties.

2.2 The Quantum Random Oracle Model

In the quantum setting, any actual implementation of a protocol with a random oracle will use a concrete hash function. Hence the most natural approach to security analysis is to allow quantum calls to any random oracle, a construction known as the quantum random oracle model (QRO). It is easy to see how this capability might cause trouble in security proofs—it is not clear how an entity can generate new random values for queries to new inputs while at the same time maintaining consistency with previously-returned hash values (this, of course, is not a problem in the classical setting, where the entity can simply keep a table of values of all previously-requested hash values; this is infeasible in the quantum setting because the entity would have to keep a table which contains hash values for *all* classical input values, since the adversary could query the random oracle on a superposition of all those values). For this reason, standard proof techniques in the classical random oracle model do not necessarily translate well to the quantum setting. Some progress has been made in developing proof techniques and constructions that work in the quantum random oracle model (see, for instance, [24,29]). In Section 7 we employ a construction due to Eaton and Song [10, Theorem 4] which can be used to obtain a signature scheme which is secure in the quantum random oracle model from a signature scheme which is secure in the classical random oracle model against an adversary who can run quantum computations.

Essentially all security models for quantum-safe cryptography can be framed independently of the quantum random oracle model, and so we end up with different "versions" of security definitions which differ only in whether random oracles are quantumly-accessible. If a definition allows for random oracles but does *not* allow them to be accessed quantumly, we may append -RO to the name (*e.g.*, EUF-qCMA-RO), while if a definition allows for random oracles and allows them to be accessed quantumly then we append -QRO (*e.g.*, EUF-qCMA-QRO). We may append no suffix if it is clear from context which is meant.

3 Mathematical Background

3.1 Elliptic Curves, Isogenies, and the j-Invariant

Definition 2 (Elliptic Curve [20, Section III.3]). *An* elliptic curve E *is a nonsingular curve (*i.e.*, a projective variety of dimension one) of genus one, with a distinguished point $\mathcal{O}$, called the* point at infinity.

We say that an elliptic curve E is *defined over* a field $\mathbb{K}$ if it is defined over $\mathbb{K}$ as an algebraic set and $\mathcal{O}$ has coordinates in $\mathbb{K}$.

It can be shown ([20, Section III.1]) that an elliptic curve defined over a field $\mathbb{K}$ with $\operatorname{char}\mathbb{K} \neq 2, 3$ can be written as

$$E = \{(x, y) \in \mathbb{A}^2(\overline{\mathbb{K}}) : y^2 = x^3 + ax + b\} \sqcup \{\mathcal{O}\}$$

We can introduce a group structure on an elliptic curve E in a natural way [20, Section III.2, Proposition 2.2], and this group structure allows us to define a class of functions, called *isogenies*, which preserve some of this algebraic structure, as well as the geometric structure of the curves.

Definition 3 (Isogeny [20, Section III.4]). *Let E and E' be elliptic curves defined over a field $\mathbb{K}$. An* isogeny *from E to E' is an algebraic morphism $\phi\colon E \to E'$ such that $\phi(\mathcal{O}_E) = \mathcal{O}_{E'}$.*

We say that an isogeny $\phi\colon E \to E'$ is *defined over* a field $\mathbb{K}$ if it is defined over $\mathbb{K}$ as a rational map. The *degree* of an isogeny is its degree when considered as a rational map. If ϕ is a separable isogeny, then $\deg \phi = |\ker \phi|$ [8, Section 2].

For elliptic curves E and E', we say that E' is *isogenous* to E over $\mathbb{K}$ if and only if there is an isogeny ϕ from E to E' defined over $\mathbb{K}$ such that $\phi(E) \neq \{\mathcal{O}_{E'}\}$. It can be shown (*q.v.* [20, Section III.6, Theorem 6.1]) that E' is isogenous to E over $\mathbb{K}$ if and only if E is isogenous to E' over $\mathbb{K}$; that is, the property of "being isogenous" is an equivalence relation, and we define the *isogeny class* of a curve E defined over $\mathbb{K}$ to be the set of all curves E' which are isogenous to E, up to $\overline{\mathbb{K}}$-isomorphism as algebraic sets. Since any algebraic morphism of curves is either constant or surjective [12, Chapter II, Section 6, Proposition 6.8], if $\phi\colon E \to E'$ is a nontrivial isogeny, then $\phi(E) = E'$.

A theorem of Tate [23, Section 3, Theorem 1] says that if E and E' are defined over a *finite* field $\mathbb{K} = GF(q)$, then E and E' are isogenous over $\mathbb{K}$ if and only if $|E(\mathbb{K}')| = |E'(\mathbb{K}')|$ for every finite extension $\mathbb{K}'$ of $\mathbb{K}$.

Let E be defined over a field of characteristic $p > 0$, and for each $\ell \in \mathbb{Z}$, let $E[\ell]$ denote the set of ℓ-torsion points of E. If $p \nmid \ell$, the map

$$[\ell]\colon E \to E$$
$$[\ell]\colon P \mapsto \ell P$$

is separable and has degree ℓ^2; hence, since $E[\ell] = \ker[\ell]$ is a finite Abelian group, it must be that $E[\ell] \cong \mathbb{Z}/\ell\mathbb{Z} \oplus \mathbb{Z}/\ell\mathbb{Z}$, since any other 2-generated Abelian group of order ℓ^2 has elements of order strictly greater than ℓ, by the Fundamental Theorem of Finitely Generated Abelian Groups [19, Chapter 10, Theorem 10.20]. Additionally, either $E[p^r] = \{\mathcal{O}\}$ for all $r \in \mathbb{Z}$, or $E[p^r] \cong \mathbb{Z}/p^r\mathbb{Z}$ for all $r \in \mathbb{Z}$ [20, Chapter V, Section 3, Theorem 3.1]; in the first case we say that E is supersingular, while in the second case we say that E is ordinary. Any two isogenous elliptic curves are either both ordinary or both supersingular. We will be concerned only with supersingular elliptic curves for our applications.

Any supersingular elliptic curve E is defined over $GF(p^2)$ for some prime p, and for each prime $\ell \neq p$ there are $\ell + 1$ isogenies of degree ℓ with domain E (though not all of them are defined over $GF(p^2)$, in general) [8]. These isogenies of degree ℓ are in one-to-one correspondence with the subgroups Φ of E of order ℓ; moreover, each such subgroup is the kernel of a unique isogeny ϕ, and we write $\phi(E) = E/\Phi$ [20, Chapter III, Section 4, Proposition 4.12]. Hence to specify an isogeny it suffices to specify its kernel, and conversely given a subgroup Φ of E we can construct the isogeny ϕ whose kernel is Φ, using Vélu's formulae [26]. In

particular, if Φ is generated by a point $R \in E(GF(p^2))$, then we have a compact representation of ϕ, and we can compute ϕ efficiently from R [8]. We will use such isogenies in an authenticated key establishment protocol in Section 7.

Associated to every elliptic curve E defined over $\mathbb{K}$ is a field $j(E) \in \mathbb{K}$, called the *j-invariant* of the curve. As the name suggests, the j-invariant is invariant under $\overline{\mathbb{K}}$-isomorphisms of algebraic sets, and so a j-invariant uniquely identifies a $\overline{\mathbb{K}}$-isomorphism class of elliptic curves over $\mathbb{K}$. Given an elliptic curve E, its j-invariant can be found in polynomial-time; moreover, given a j-invariant $j^* \in \mathbb{K}$, one can find in polynomial time the curve E with $j(E) = j^*$. Knowing this, we have a compact description of an elliptic curve for the purposes of communication during a key establishment protocol.

3.2 Supersingular Elliptic Curve Isogeny-based Cryptography

In the following definition, let $p = \ell_A^{e_A} \ell_B^{e_B} f \pm 1$ be a prime, where ℓ_A and ℓ_B are distinct small primes, and f is a small cofactor used to ensure that p is prime—we will not quantify what is meant by "small." As well, let E be a supersingular elliptic curve defined over $\mathbb{K} = GF(p^2)$ with $E(GF(p^2)) \cong \mathbb{Z}/(p \mp 1)\mathbb{Z} \oplus \mathbb{Z}/(p \mp 1)\mathbb{Z}$, and let $\{P_A, Q_B\}$ and $\{P_B, Q_B\}$ be bases for $E[\ell_A^{e_A}]$ and $E[\ell_B^{e_B}]$, respectively.

Definition 4 (Supersingular Isogeny Decisional Diffie-Hellman Problem (SIDH)). *Let $\phi_A \colon E \to E_A$ be an isogeny with kernel $\langle m_A P_A + n_A Q_A \rangle$ where m_A, n_A are chosen uniformly at random from $\mathbb{Z}/\ell_A^{e_A}\mathbb{Z}$, not both divisible by ℓ_A. Let $\phi_B \colon E \to E_B$ be an isogeny with kernel $\langle m_B P_B + n_B Q_B \rangle$ where m_B, n_B are chosen uniformly at random from $\mathbb{Z}/\ell_B^{e_B}\mathbb{Z}$, not both divisible by ℓ_B. Given a tuple $(E, E_A, E_B, \phi_A(P_B), \phi_A(Q_B), \phi_B(P_A), \phi_B(Q_A), E_C)$ where either $E_C = E_{AB} = E/\langle m_A P_A + n_A Q_A, m_B P_B + n_B Q_B \rangle$ or E_C is sampled uniformly at random from the set of all curves of the form $E/\langle x_A P_A + y_A Q_A, x_B P_B + y_B Q_B \rangle$ where x_A, y_A and x_B, y_B are chosen with the same conditions as m_A, n_A and m_B, n_B, each with probability $\frac{1}{2}$, the* supersingular isogeny decisional Diffie-Hellman problem (SIDH) *is to determine which is the case.*

The corresponding security assumption is that arbitrary instances of SIDH cannot be solved in polynomial-time on a quantum computer with non-negligible advantage. Presently the best known quantum algorithm for this problem runs in fully-exponential time $O(\sqrt[6]{p})$, [5] while the best known classical attack runs in time $O(\sqrt[4]{p})$ [8, Section 5.2].

In particular, we will use a strong designated verifier signature scheme due to Sun *et al.* [22] and a key establishment protocol due to De Feo, Jao, and Plût [8], both based on supersingular elliptic curve isogenies to make an authenticated key establishment protocol. We note that it is not necessary to use a strong designated verifier signature scheme for the construction—any sufficiently secure signature scheme will do. We chose this signature scheme primarily because it is conceptually simple, stateless, and is based on more conventional computational assumptions than other isogeny-based signature schemes.

4 Tools for Constructing a Secure AKE Protocol

In this section we briefly mention three tools we require to construct our key establishment protocol: chameleon hash functions, and two generic security-strengthening transformations for signature schemes.

4.1 Chameleon Hash Functions

Intuitively, chameleon hash functions—introduced by Krawczyk and Rabin [15]—are a type of hash function which are collision resistant for anybody who does not know an associated piece of secret information, but for which collisions can easily be found for any input given that piece of secret information. We will use them to construct signature schemes which are secure in the quantum random oracle model. A precise definition of chameleon hash function is given in Definition 5.

Definition 5 (Chameleon Hash Function (Adapted from [6, Definition 3.9])).
A chameleon hash function $\mathcal{H}$ *is a tuple* $(\mathrm{Gen}, H, \mathrm{Inv}, \mathrm{Sample})$ *of algorithms such that:*

1. $\mathrm{Gen}(\lambda)$ *generates a key pair* $(\mathrm{sk}, \mathrm{pk})$ *with security parameter* λ*;*
2. $H_{\mathrm{pk}}(m, r)$ *maps messages* m *to some target space* $\mathcal{Y}$*;*
3. $\mathrm{Sample}(\lambda)$ *samples* r *such that* $H_{\mathrm{pk}}(m, r)$ *is distributed computationally indistinguishably from uniform over the image of* $H_{\mathrm{pk}}(m, \cdot)$ *for every pair* (pk, m)[6]*;*
4. $\mathrm{Inv}_{\mathrm{sk}}(h, m)$ *produces* r *such that* $H_{\mathrm{pk}}(m, r) = h$ *(where* $(\mathrm{sk}, \mathrm{pk})$ *is generated by* $\mathrm{Gen}(\lambda)$*), with distribution computationally indistinguishable from that of* $\mathrm{Sample}(\lambda)$ *conditioned on* $H_{\mathrm{pk}}(m, r) = h$*; and,*
5. *For any* pk, $H_{\mathrm{pk}}(\cdot, \cdot)$ *is collision resistant.*

A chameleon hash function is *quantum-safe* if it is collision-resistant against a quantum adversary who can query the function in superposition.

4.2 Generic Security-Strengthening Transformations

We would like to use a signature scheme for authentication in an authenticated key establishment protocol which is secure in a quantum model where the adversary can make quantum queries to oracles which emulate classical parties, and to random oracles. For this purpose, clearly a signature scheme which is EUF-qCMA secure in the quantum random oracle model (EUF-qCMA-QRO) is required. The strong designated verifier signature scheme that we use in Section 7 provides only EUF-CMA-RO security. To get from EUF-CMA-RO to EUF-qCMA-QRO, we apply two generic transformations:

[6] Here and elsewhere, we use $\cdot$ as above to indicate inputs to a function which are *not* fixed when others (pk and m, above) are fixed.

1. Eaton & Song [10, Theorem 4]: EUF-CMA-RO → EUF-CMA-QRO;
2. Boneh & Zhandry [6, Constr. 3.12]: EUF-CMA-QRO → EUF-qCMA-QRO.

The resulting protocols are described in detail in Section 7.

5 The Security Model

To model different "levels" of security, security models for authenticated key establishment do not typically have fixed assumptions on the *computational* capabilities of the adversary. In particular, they do not explicitly disallow adversaries access to a quantum computer. However, to model the *interaction* of the adversary and honest parties, these security models typically follow the lead of the extended Canetti-Krawczyk model [16] and its predecessors [4,7] and allow the adversary to interact only classically with the parties; that is, they can only deliver single, classical messages, rather than quantum superpositions. On the contrary, in this work we consider the case where the parties are modelled as quantum oracles (with memory) with which the adversary interacts. As noted in the introduction, such quantum-aware security models have been previously introduced in other settings [1,6,11], and the attacks which make particular use of the quantum nature of the oracles have been studied under the name "superposition attacks," "quantum chosen message attacks" (for MAC and signature schemes), and "quantum chosen ciphertext attacks" (for encryption schemes).

Though the model may seem to give the adversary unreasonable abilities, it has a number of desirable properties that make it very natural and useful to study. For instance:

1. It is simple enough that analyzing the security of specific protocols and the effects of natural quantum attacks is sufficiently easy;
2. It encompasses all common classical and non-interactive quantum attacks on stateless key establishment protocols, and in that sense is "stronger" than previously-considered models;
3. Because it explicitly allows for quantum interactions, protocols proved secure in this model can be used without modification on a quantum computer without worrying about novel attacks. Though in principle this could be solved by (for instance) requiring all quantum input as part of a classical protocol be measured to collapse superpositions, this requires particular engineering of the devices; moreover, our method will work even on *untrusted* devices (*q.v.* "device-independent" quantum cryptography [17,25]), and;
4. It allows for security in some unorthodox quantum-safe scenarios which allow quantum interactions; for instance, if an adversary with quantum computing capabilities is given an obfuscated classical circuit which emulates a classical key-establishing party on given sessions, in order to (for example) temporarily delegate key establishment to a server.

5.1 Definitions

Definition 6 (Party). *A party $\mathcal{P}$ is an interactive classical Turing machine*[1] *with access to a source of random bits.*

Associated to each party $\mathcal{P}$ is a (possibly empty) private key/public key pair $(\mathrm{sk}, \mathrm{pk})$. For the purposes of the model, it is assumed that each party has a genuine copy of each other party's public key, in order to allow for authentication. Moreover, to each party we associate a unique identifier `id`. It is assumed that each party has a genuine copy of each other party's identifier.

Definition 7 (Protocol). *A protocol Π is a specification of subroutines, to be run by some number of parties, to establish a session key.*

A protocol Π is said to be *correct* if, when Π is executed correctly and all messages are relayed faithfully (*i.e.*, without changes to their content or ordering), all parties involved compute the same key.

Protocols are *message-driven*; that is, upon receiving a message, a party computes the response message and sends it to the intended recipient. The party does no further computations and sends no more messages until it is activated again by an incoming message. This assumption sacrifices no generality, since any computations can be performed *before* sending out a new message.

Definition 8 (Session). *A session is an instance of a protocol at a party.*

Associated to each session is a unique session identifier Ψ—which we use to refer to the session—chosen by the session's owner. If a party $\mathcal{P}$ has a session Ψ, the parties with whom $\mathcal{P}$ believes they are attempting to establish a key are called *peers* to $\mathcal{P}$ in session Ψ, and the peers' associated sessions (if they exist) are called *matching sessions.*

If a party $\mathcal{P}$ with identifier `id` who owns a given session Ψ, with matching session Ψ' and peer $\mathcal{P}'$ with identifier $\mathtt{id}'$ has received messages $m_1, \ldots, m_{k-1}$ in this session, then we denote by $\mathcal{P}(\Psi, \Psi', \mathtt{id}, \mathtt{id}'; \mathrm{pk}, \mathrm{pk}', \mathrm{sk}; m_1, \ldots, m_k; r_\Psi)$ the message that $\mathcal{P}$ sends given that the next message it receives in this session is m_k, and it uses randomness r_Ψ for this session. We abbreviate this expression as $\mathcal{P}(m_k)$ if the other inputs are clear from context.

Invalid Messages In a key establishment security model, it is typical to model a mechanism for parties to abort a session in the event that it receives an "invalid" message. What exactly constitutes an invalid message is defined by a given protocol, but typically an invalid message is one which does not make sense in the context of the protocol or fails to validate under the public key of the party

[1] Though we will allow quantum access to messaging oracles which emulate parties, the parties themselves must be defined to be classical, because we intend to model classical cryptosystems. The quantum-safe security comes from the quantum interactions that we allow for the security experiment.

believed to have sent the message. If we allow the adversary to deliver superpositions of messages, however, it does not make sense to consider such termination since parties cannot "read off" messages from a superposition and thus cannot tell whether to terminate the session. Measuring the state would collapse it and defeat the purpose of considering quantum queries entirely. To handle such cases, we introduce a fail character $\perp$ to be used when a session would be terminated. Define $\mathcal{P}(\Psi, \Psi', \mathtt{id}, \mathtt{id}'; \mathrm{pk}, \mathrm{pk}', \mathrm{sk}; m_1, \ldots, m_k; r_\Psi) = \perp$ when m_k is invalid, and define all further messages in a session after a response has been $\perp$ to be $\perp$. This formalism allows a session to be in a superposition of terminated and active.

5.2 Party and Adversarial Capabilities

Aside from classical computations, parties can issue a $\mathtt{Send}(\mathtt{id}, m)$ command, which requests that the message m be delivered to to the party identified by $\mathtt{id}$. Parties may store private/public value pairs (k, K) in memory, associated to sessions. That is, for a key establishment session Ψ a party may construct an ephemeral secret key sk_Ψ and corresponding ephemeral public value pk_Ψ. To be consistent with other quantum-safe security definitions, randomness is drawn classically; that is, if a function $f(\cdot; r)$ which depends on randomness r is to be applied in superposition, the value of r is chosen classically and the map $U_{f(\cdot;r)}$ is applied.

In order to make protocols meaningfully quantum-resistant, during the security experiment the challenger will provide a *quantum messaging oracle* $O_\mathcal{P}$ for each party $\mathcal{P}$, defined inductively as follows. Before $O_\mathcal{P}$ receives any messages in session Ψ with matching session Ψ' and peer $\mathcal{P}'$ with identifier $\mathtt{id}'$, we define

$$O_\mathcal{P}(\Psi) |m_1\rangle |y\rangle = |m_1\rangle \left|y \oplus \mathcal{P}(\Psi, \Psi', \mathtt{id}, \mathtt{id}'; \mathrm{pk}, \mathrm{pk}', \mathrm{sk}; m_1; r_\Psi)\right\rangle$$

for all $m_1 \in \mathcal{M}$, and extend linearly to superpositions. Notice that the session Ψ considered by $O_\mathcal{P}$ must be given as *classical* input. $O_\mathcal{P}$ then *keeps* the first register, and returns the second. It is important that $O_\mathcal{P}$ hold onto the register, since otherwise it "forgets" what the adversary has sent it, and then does not know its stage in the protocol and cannot respond to future messages appropriately.

After receiving $k-1$ messages, $O_\mathcal{P}$ will hold $k-1$ registers. When queried again, we consider its input as the first $k+1$ registers of some global (pure) state

$$|\Gamma\rangle = \sum_{\boldsymbol{m}, y} \alpha_{\boldsymbol{m}, y} \underbrace{|m_1\rangle \cdots |m_{k-1}\rangle}_{\text{Held by Challenger}} \overbrace{|m_k\rangle |y\rangle}^{\text{Provided by } \mathcal{A}} \underbrace{|\mu_{\boldsymbol{m}, y}\rangle}_{\text{Remaining registers}} .$$

It acts as $O_\mathcal{P}(\Psi) |m_1\rangle \cdots |m_k\rangle |y\rangle = |m_1\rangle \cdots |m_k\rangle |y \oplus \mathcal{P}(m_k)\rangle$.

This model differs from others in that it does not allow for parties to keep an internal "state" which specifies what has happened previously in a session; rather, it requires that parties essentially keep a (quantum) transcript of received messages. For many protocols (in particular, the one we present in Section 7) our choice makes no difference for security. However, there may be subtle differences

for stateful protocols—in particular, for protocols using hash-based signatures for authentication. We do not consider such protocols here.

If $\mathcal{P}$ does not own a session $\hat{\Psi}$, then we simply define $O_{\mathcal{P}}(\hat{\Psi})\,|m\rangle\,|y\rangle$ to be $|m\rangle\,|y \oplus \bot\rangle$. For simplicity, for the purposes of the security experiment, the adversary interacts only with these quantum messaging oracles.

The adversary interacts with the quantum messaging oracles by delivering (quantum superpositions of) messages to them. The adversary may also issue the following "information-leakage" commands:

1. `RevealEphemeralKey(id,`Ψ`)`: If the identified party owns a session Ψ, the challenger reveals any ephemeral secret key associated to Ψ.
2. `RevealPrivateKey(id)`: The identified party reveals their private key.
3. `Corrupt(id)`: The party identified by `id` reveals all classical information is knows to the adversary, turns over all its quantum memory to the adversary, and becomes adversarially-controlled.

As well, the adversary may issue a `RevealSessionKey(id,`Ψ`)` query, defined as follows: if $K_\Psi := K(\Psi, \Psi', \mathtt{id}, \mathtt{id}'; \mathrm{pk}, \mathrm{pk}', \mathrm{sk}; m_1, \ldots, m_k; r_\Psi)$ is the key that the party identified by `id` would compute in session Ψ with peer identifier `id`$'$ and matching session Ψ', and it has so far received messages $m_1, \ldots, m_k$, then if the global state is

$$|\Gamma\rangle = \sum_{\boldsymbol{m},y} \alpha_{\boldsymbol{m},y} \underbrace{|m_1\rangle \cdots |m_k\rangle}_{\text{Held by Challenger}} \overbrace{|y\rangle}^{\text{Provided by } \mathcal{A}} \underbrace{|\mu_{\boldsymbol{m},y}\rangle}_{\text{Remaining registers}} ,$$

the result of this query is defined by

$$|m_1\rangle \cdots |m_k\rangle\,|y\rangle \mapsto |m_1\rangle \cdots |m_k\rangle\,|y \oplus K_\Psi\rangle\,.$$

As a result of this query $O_{\mathcal{P}}$ returns the first $k+1$ registers of the global state (that is, the target register and the message registers it held).

In particular, if the adversary chooses not to entangle the registers of different pairs of sessions, then for each session Ψ whose session key the adversary reveals, the adversary will have the state

$$\sum_{m_1,\ldots,m_k;y} \alpha_{m_1,\ldots,m_k,y}\,|m_1\rangle \cdots |m_k\rangle\,|y \oplus K_\Psi\rangle$$

and the registers corresponding to different revealed sessions will be disentangled from one another. This is precisely analogous to qCPA security of encryption schemes and qCMA security of signatures schemes, where the adversary receives

$$\sum_{m,y} \alpha_{m,y}\,|m\rangle\,|y \oplus \mathrm{Enc}_{\mathrm{sk}}(m;r)\rangle\,, \text{ or } \sum_{m,y} \alpha_{m,y}\,|m\rangle\,|y \oplus \mathrm{Sign}_{\mathrm{sk}}(m;r)\rangle\,.$$

5.3 The Security Experiment

A session Ψ owned by party $\mathcal{P}$ with peer $\mathcal{P}'$ and partner session Ψ' is "clean" if:

1. At session completion, neither $\mathcal{P}$ nor $\mathcal{P}'$ was adversarially-controlled;
2. At session completion, $\mathcal{A}$ had issued neither $\texttt{RequestPrivateKey}(\texttt{id})$ nor $\texttt{RequestPrivateKey}(\texttt{id}')$;
3. $\mathcal{A}$ has not revealed the ephemeral secret key for Ψ or Ψ', and;
4. $\mathcal{A}$ has not revealed the session key for Ψ or Ψ'.

For the security experiment, the adversary issues a $\texttt{Test}(\texttt{id}, \Psi)$ query on a clean session Ψ owned by the party with identifier $\texttt{id}$, defined as follows: the adversary provides a target register $|y\rangle$, and the challenger selects $b \in \{0, 1\}$ uniformly at random. If $b = 1$, $\texttt{Test}(\texttt{id}, \Psi)$ acts like a $\texttt{RevealSessionKey}$ query; if $b = 0$, the result is defined by

$$|m_1\rangle \cdots |m_k\rangle\, |y\rangle \mapsto |m_1\rangle \cdots |m_k\rangle\, |y \oplus R(m_1, \ldots m_k)\rangle$$

where $R(m_1, m_2, \ldots, m_k)$ is a random string in $\mathcal{K}$ subject to

$$R(m_1, m_2, \ldots, m_k) = \perp \iff K(\Psi, \Psi', \texttt{id}, \texttt{id}'; \mathrm{pk}, \mathrm{pk}', \mathrm{sk}; \boldsymbol{m}; r_\Psi) = \perp .$$

In any case, all $k+1$ of these registers are returned. We call a key establishment protocol Π *secure* if the probability that a polynomial-time adversary $\mathcal{A}$ can correctly guess the value of b is at most negligibly greater than $\frac{1}{2}$.

6 A Construction for Secure Protocols

In this section we present a construction for secure authenticated key establishment protocols from an unauthenticated key establishment protocol and an EUF-qCMA signature scheme. This is analogous to the well-known result of Bellare, Canetti, and Krawczyk [3, Proposition 4], which states that an EUF-CMA signature scheme can be used as an authentication method for an unauthenticated key establishment protocol.

Theorem 1. *Let $\mathcal{S} = (\mathrm{Gen}, \mathrm{Sign}, \mathrm{Ver})$ be a strongly EUF-qCMA secure signature scheme, and let Π be a two-round, two-party key establishment protocol. Consider the protocol Π' with the following properties:*

1. *Each party $\mathcal{P}'_k$ has a key pair $(\mathrm{sk}_k, \mathrm{pk}_k)$ for $\mathcal{S}$ and, moreover, each party knows each other party's public key.*
2. *Whenever an initiating party $\mathcal{P}'_I$ would send a message m, it instead constructs the message $m_{\Psi^{(I)}} = (m, \texttt{id}_I, \texttt{id}_R, \Psi^{(I)})$ (where $\texttt{id}_R$ is the peer's session identifier and $\Psi^{(I)}$ is $\mathcal{P}'_I$'s session identifier) and sends $(m_{\Psi^{(I)}}, \sigma_{\Psi^{(I)}})$, where $\sigma_{\Psi^{(I)}} = \mathrm{Sign}_{\mathrm{sk}_I}(m, \texttt{id}_I, \texttt{id}_R, \Psi^{(I)})$*

3. *Whenever a party $\mathcal{P}'_R$ would respond to a message $(m, \Psi^{(I)}, \sigma)$ from an initiating party $\mathcal{P}'_I$ with a message m', it computes*

$$b = \mathrm{Ver}_{\mathrm{pk}_I}(m_{\Psi^{(I)}}, \sigma_{\Psi^{(I)}}) = \mathrm{Ver}_{\mathrm{pk}_I}((m, \mathtt{id}_I, \mathtt{id}_R, \Psi^{(I)}), \sigma).$$

If $b = 0$, P'_R responds with $(\perp, \perp, \perp)$; otherwise, $\mathcal{P}'_R$ constructs the message $m_{\Psi^{(R)}} = (m', \mathtt{id}_I, \mathtt{id}_R, \Psi^{(I)}, \Psi^{(R)})$ and sends $(m_{\Psi^{(R)}}, \sigma_{\Psi^{(R)}})$, where $\sigma_{\Psi^{(R)}} = \mathrm{Sign}_{\mathrm{sk}_R}(m_{\Psi^{(R)}}) = \mathrm{Sign}_{\mathrm{sk}_R}(m', \mathtt{id}_I, \mathtt{id}_R, \Psi^{(I)}, \Psi^{(R)})$

4. *If a party $\mathcal{P}_k$ would compute a session key for a session $\Psi^{(k)}$ with partner $\mathcal{P}_\ell$ and partner session $\Psi^{(\ell)}$, it checks whether the signature in the message it received was valid; if not it outputs session key $\perp$. If the signature is valid, it outputs the session key as usual.*

Then:

1. *If Π is correct, then Π' is correct; and,*
2. *If Π is secure against adversaries $\mathcal{A}$ which are required to deliver all messages faithfully, then Π' is secure.*

Proof Idea. The idea of the proof is intuitive, though the proof itself is involved; the proof appears in appendix A.

At its core, the proof resembles analogous classical results (for instance, [3, Proposition 4]), which use the security of the signature scheme to argue that the adversary cannot deliver any message m in session Ψ to a party $\mathcal{P}$ with identifier $\mathtt{id}$ which:

1. Is unsent (*i.e.*, no party issued $\mathrm{Send}(\mathtt{id}, m)$ in a session matching Ψ),
2. Does not come from a corrupted party, and
3. Is valid (*i.e.,* the messages pass $\mathcal{P}$'s authentication check)

The complication that arises in our setting is that messages can be delivered in superposition, and so may "simultaneously" be valid *and* invalid, in the sense that some parts of the superposition may be valid while other are not. To address this, we demonstrate that under the security assumptions of the theorem, no adversary can construct a superposition $|\Gamma\rangle = \sum_{\boldsymbol{\mu},m,\sigma} \alpha_{\boldsymbol{\mu},m,\sigma} |\boldsymbol{\mu}\rangle |m, \sigma\rangle$ of messages which:

1. Is unsent,
2. Does not come from a corrupted party, and
3. Has nonnegligible valid content (*i.e.,* $\sum_{\boldsymbol{\mu};(m,\sigma)} |\alpha_{\boldsymbol{\mu},m,\sigma}|^2$ is nonnegligible, where the sum is taken over valid pairs (m, σ).)

From this, we show that in the test session, either

1. The adversary does not interfere with the test session, and so cannot win the security game by virtue of the security of the underlying key establishment protocol, or
2. The adversary interferes with the test session, and by doing so ensures that the valid content of the output of the test query is negligible.

In the second scenario, the adversary cannot distinguish the output of either the $b = 0$ or $b = 1$ case from being given a state where all key registers are $\perp$ except with negligible probability, and hence in particular cannot distinguish $b = 0$ from $b = 1$; this means that Π' is secure.

Figure 2 in Appendix A.2 illustrates a pair of matching sessions of a generic protocol Π' constructed according to Theorem 1.

7 A Protocol using Supersingular Elliptic Curve Isogenies

Here we apply the generic construction from Section 6 to construct a secure authenticated key establishment protocol based on the Supersingular Isogeny Decisional Diffie-Hellman assumption (SIDH). The underling key establishment protocol is De Feo, Jao, and Plût's scheme [8], with authentication provided by a signature scheme constructed by applying Eaton and Song's [10] and Boneh and Zhandry's [6] transformations to Sun *et. al's* [22] strong designated verifier signature scheme, reminiscent of Soukharev, Jao, and Seshadri's work in [21].

For readability we use the "vanilla" protocol from [8] without consideration for optimizations (*e.g.* [28]); such optimizations can be adapted for the protocol our protocol in a straightforward fashion without affecting the proof of security.

7.1 Our Authenticated Key Establishment Protocol

First we list the required global parameters. Authentication requires

1. $p_A = \ell_S^{e_S} \ell_V^{e_V} f_A \pm 1$, a prime, for primes ℓ_S and ℓ_V and small cofactor f_A;
2. E_A, a supersingular elliptic curve defined over $\mathbb{K}_A = GF(p_A^2)$;
3. $\{P_S, Q_S\}$ and $\{P_V, Q_V\}$, bases for $E_A[\ell_S^{e_S}]$ and $E_A[\ell_V^{e_V}]$, respectively;
4. $\mathcal{H} = (\text{Gen}^{(H_c)}, H_c, \text{Inv}, \text{Sample})$, a quantum-safe chameleon hash; and,
5. $\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3$, random oracles (in practice, hash functions).

For key establishment we require

1. $p_K = \ell_I^{e_I} \ell_R^{e_R} f_K \pm 1$, a prime, for primes ℓ_I and ℓ_R and small cofactor f_K;
2. E_K, a supersingular elliptic curve defined over $\mathbb{K}_K = GF(p_K^2)$; and,
3. $\{P_I, Q_I\}$ and $\{P_R, Q_R\}$, bases for $E_K[\ell_I^{e_I}]$ and $E_K[\ell_R^{e_R}]$, respectively.

Each party $\mathcal{P}_k$ must establish authentication keys; associated to each will be a key pair for signing and a key pair for verification. In particular, $\mathcal{P}_k$ selects $m_S^{(k)}, n_S^{(k)} \in \mathbb{Z}/\ell_S^{e_S}\mathbb{Z}$ not both divisible by ℓ_S uniformly at random, and sets $E_S^{(k)} = E_A/\langle m_S^{(k)} P_S + n_S^{(k)} Q_S \rangle$. Let $\phi_S^{(k)}$ denote the isogeny with domain E_A and image $E_S^{(k)}$. Similarly, $\mathcal{P}_k$ selects $m_V^{(k)}, n_V^{(k)} \in \mathbb{Z}/\ell_V^{e_V}\mathbb{Z}$ not both divisible by ℓ_V uniformly at random, sets $E_V^{(k)} = E_A/\langle m_V^{(k)} P_V + n_V^{(k)} Q_V \rangle$, and further sets $\phi_V^{(k)}$ to be the isogeny with domain E_A and image $E_V^{(k)}$. $\mathcal{P}_k$ must also select a

private key/public key pair $(\mathrm{sk}_H^{(k)}, \mathrm{pk}_H^{(k)})$ for the chameleon hash function. Then $\mathcal{P}_k$'s authentication key pair is

$$\begin{aligned}(\mathrm{sk}^{(k)}, \mathrm{pk}^{(k)}) &= \left((\mathrm{sk}_S^{(k)}, \mathrm{sk}_V^{(k)}, \mathrm{sk}_H^{(k)}), (\mathrm{pk}_S^{(k)}, \mathrm{pk}_V^{(k)}, \mathrm{pk}_H^{(k)})\right) \\ &= \Big(\left((m_S^{(k)}, n_S^{(k)}), (m_V^{(k)}, n_V^{(k)}), \mathrm{sk}_H^{(k)}\right) , \left((E_S^{(k)}, \phi_S^{(k)}(P_V),\ \phi_S^{(k)}(Q_V)),\right. \\ &\qquad \left.\left.(E_V^{(k)}, \phi_V^{(k)}(P_S), \phi_V^{(k)}(Q_S)), \mathrm{pk}_H^{(k)}\right)\Big) .\end{aligned}$$

Then for each pair $(\mathcal{P}_k, \mathcal{P}_\ell)$ of parties there is a curve $E_{SV}^{(k,\ell)}$ defined by

$$E_{SV}^{(k,\ell)} = E_V^{(\ell)} / \langle m_S^{(k)} \phi_V^{(\ell)}(P_S) + n_S^{(k)} \phi_V^{(\ell)}(Q_S) \rangle;$$

notice that $E_{SV}^{(k,\ell)} = E_{SV}^{(\ell,k)}$, and so both $\mathcal{P}_k$ and $\mathcal{P}_\ell$ can compute it using their secret keys and the other's public keys. This curve will be used for $\mathcal{P}_k$ to sign a message to $\mathcal{P}_\ell$. Figure 1 in Appendix A.2 depicts the j-invariant computation for this protocol.

The protocol is as follows (with $\mathcal{P}_k$ initiating and $\mathcal{P}_\ell$ responding):

1. Upon being instructed to start a session with $\mathcal{P}_\ell$, $\mathcal{P}_k$:
   a) Selects a session identifier Ψ;
   b) Selects $x_I^{(\Psi)}, y_I^{(\Psi)} \in \mathbb{Z}/\ell_I^{e_I}\mathbb{Z}$, not both divisible by ℓ_I, uniformly at random;
   c) Constructs $R^{(\Psi)} = x_I^{(\Psi)} P_I + y_I^{(\Psi)} Q_I$, defines $\phi^{(\Psi)}$ to be the isogeny with kernel $\langle R^{(\Psi)} \rangle$, and sets

$$m^{(\Psi)} = (E^{(\Psi)} = E_K / \langle R^{(\Psi)} \rangle, \phi^{(\Psi)}(P_R), \phi^{(\Psi)}(Q_R), \mathtt{id}_k, \mathtt{id}_\ell, \Psi);$$

   d) Selects $r^{(\Psi)} \in \{0,1\}^*$ at random;
   e) Sets

$$\begin{aligned}\sigma^{(\Psi)} &= (\sigma_1, \sigma_2, \sigma_3) \\ &= (\mathrm{Sample}(\lambda), \mathcal{O}_1(r^{(\Psi)} || j(E_{SV}^{(k,\ell)})), \mathrm{Inv}_{\mathrm{sk}_H^{(k)}}(r^{(\Psi)}, \mathcal{O}_2(\mathcal{O}_3(m^{(\Psi)}, \sigma_1) || \sigma_2))),\end{aligned}$$

   and;
   f) Activates Send$(\mathtt{id}_\ell, m^{(\Psi)}, \sigma^{(\Psi)})$
2. Upon receiving (m, σ), $\mathcal{P}_\ell$:
   a) Computes

$$b^{(\Psi')} = \begin{cases} 1 \text{ if } \sigma_2 = \mathcal{O}_1(H_{\mathrm{pk}_H^{(k)}}(\mathcal{O}_2(\mathcal{O}_3(m, \sigma_1) || \sigma_2), \sigma_3) || j(E_{SV}^{(k,\ell)})) \\ 0 \text{ otherwise} \end{cases};$$

   If $b^{(\Psi')} = 0$ the delivered message is invalid and is hence rejected; then $\mathcal{P}_\ell$ activates Send$(\mathtt{id}_k; \perp, \perp, \perp)$. Otherwise, $\mathcal{P}_\ell$:
   b) Selects a session identifier Ψ';

c) Selects $x_R^{(\Psi')}, y_R^{(\Psi')} \in \mathbb{Z}/\ell_R^{e_R}\mathbb{Z}$, not both divisible by ℓ_R, uniformly at random;

d) Constructs $R^{(\Psi')} = x_R^{(\Psi')} P_R + y_R^{(\Psi')} Q_R$, defines $\phi^{(\Psi')}$ to be the isogeny with kernel $\langle R^{(\Psi')} \rangle$, and sets

$$m^{(\Psi')} = (E^{(\Psi')} = E_K/\langle R^{(\Psi')} \rangle, \phi^{(\Psi')}(P_I), \phi^{(\Psi')}(Q_I), \mathtt{id}_k, \mathtt{id}_\ell, \Psi, \Psi');$$

e) Selects $r^{(\Psi')} \in \{0,1\}^*$ at random;

f) Sets

$$\begin{aligned}\sigma^{(\Psi')} &= (\sigma_1, \sigma_2, \sigma_3) \\ &= (\mathrm{Sample}(\lambda), \mathcal{O}_1(r^{(\Psi')} || j(E_{SV}^{(\ell,k)})), \mathrm{Inv}_{\mathrm{sk}_H^{(\ell)}}(r^{(\Psi')}, \mathcal{O}_2(\mathcal{O}_3(m^{(\Psi')}, \sigma_1)||\sigma_2))),\end{aligned}$$

and;

g) Activates $\mathrm{Send}(\mathtt{id}_\ell, m^{(\Psi')}, \sigma^{(\Psi')})$

After receiving the message, $\mathcal{P}_\ell$ computes the session key

$$K^{(\Psi')} = \begin{cases} \perp \text{ if } b^{(\Psi')} = 0 \\ m_1/\langle x_R^{(\Psi')} m_2 + y_R^{(\Psi')} m_3 \rangle \text{ otherwise} \end{cases}.$$

3. Upon receiving (m, σ), $\mathcal{P}_k$ computes the session key

$$b^{(\Psi)} = \begin{cases} 1 \text{ if } \sigma_2 = \mathcal{O}_1(H_{\mathrm{pk}_H^{(\ell)}}(\mathcal{O}_2(\mathcal{O}_3(m, \sigma_1)||\sigma_2), \sigma_3)||j(E_{SV}^{(\ell,k)})) \\ 0 \text{ otherwise} \end{cases};$$

After receiving the message, $\mathcal{P}_k$ computes the session key

$$K^{(\Psi)} = \begin{cases} \perp \text{ if } b^{(\Psi)} = 0 \\ m_1/\langle x_I^{(\Psi)} m_2 + y_I^{(\Psi)} m_3 \rangle \text{ otherwise} \end{cases}.$$

Theorem 2. *The scheme described above is correct.*

Proof. This follows immediately from the results of [8, Section 3.1].

Theorem 3. *Under the Supersingular Isogeny Decisional Diffie-Hellman assumption, the scheme described above is secure in the security model described in Section 5 in the quantum random oracle model.*

Proof. This follows from Theorem 1 by the security results of [8,22,10,6].

Thus we have demonstrated that we have constructed a protocol, based on now-standard post-quantum cryptographic assumptions, which resists superposition attacks in the quantum random oracle model.

8 Conclusion

We have presented a security model for authenticated key establishment in which the adversary can deliver quantum superpositions of messages to parties who would ordinarily be participating in a classical protocol, analogous to allowing quantum signing queries in EUF-qCMA security of signature schemes or quantum encryption/decryption queries in standard quantum-safe security definitions of encryption. We demonstrated that the corresponding new security definition is achievable by constructing a specific example of a secure key establishment protocol assuming the quantum hardness of a Diffie-Hellman-type problem for isogenies of supersingular elliptic curves, and gave a generic construction for secure protocols using sufficiently secure signature schemes and unauthenticated key establishment protocols.

Limitations and Future Work. Although the security model we present is a natural for quantum-safe authenticated key establishment, it remains to establish a separation between this security definition and, for instance, the more standard Canetti-Krawczyk model with a quantum adversary—that is, we must show that there are protocols which are secure in the Canetti-Krawczyk model when the adversary has access to a quantum computer which are insecure when the adversary can deliver quantum superpositions of messages. We would like to establish this separation to demonstrate the *necessity* of our quantum-safe security model. Once this separation is established, we can work solving the following problems:

1. Are current "quantum-safe" key establishment protocols secure in this model?
2. Are there other simple generic constructions for secure protocols; in particular, does our construction generalize to the case of many-round protocols?
3. Can we introduce more security properties: *e.g.,* resilience against key compromise impersonation and malicious insiders [16]?
4. Our security definitions do not adequately model stateful protocols; this choice was made to avoid entangling the (classical) state of the protocol user with the global (quantum) state of the adversary and environment. This means that the model is not sufficient for analyzing certain well-known types of protocols, such as hash-based schemes. How best to allow stateful protocols while remaining grounded in reality is a matter for further consideration.

References

1. Alagic, Gorjan and Broadbent, Anne and Fefferman, Bill and Gagliardoni, Tommaso and Schaffner, Christian and St. Jules, Michael. Computational Security of Quantum Encryption. In Nascimento, Anderson C.A. and Barreto, Paulo, editor, *Information Theoretic Security*, pages 47–71. Springer International Publishing, 2016.

2. Paulo S. L. M. Barreto, Shay Gueron, Tim Gueneysu, Rafael Misoczki, Edoardo Persichetti, Nicolas Sendrier, and Jean-Pierre Tillich. Cake: Code-based algorithm for key encapsulation. Cryptology ePrint Archive, Report 2017/757, 2017.
3. Mihir Bellare, Ran Canetti, and Hugo Krawczyk. A modular approach to the design and analysis of authentication and key exchange protocols. In *Proceedings of STOC '98*, pages 419–428. ACM Press, 1998.
4. Mihir Bellare and Phillip Rogaway. Entity Authentication and Key Distribution. In Stinson, Douglas R., editor, *Advances in Cryptology — CRYPTO '93*, pages 232–249, Berlin, Heidelberg, 1994. Springer Berlin Heidelberg.
5. Jean-François Biasse, David Jao, and Anirudh Sankar. A Quantum Algorithm for Computing Isogenies between Supersingular Elliptic Curves. In Meier, Willi and Mukhopadhyay, Debdeep, editor, *Progress in Cryptology — INDOCRYPT 2014*, pages 428–442, Cham, 2014. Springer International Publishing.
6. Dan Boneh and Mark Zhandry. Secure Signatures and Chosen Ciphertext Security in a Quantum Computing World. In Canetti, Ran and Garay, Juan A., editor, *Advances in Cryptology — CRYPTO 2013*, pages 361–379, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
7. Ran Canetti and Hugo Krawczyk. Analysis of Key-Exchange Protocols and Their Use for Building Secure Channels. In Pfitzmann, Birgit, editor, *Advances in Cryptology — EUROCRYPT 2001*, pages 453–474, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.
8. Luca De Feo, David Jao, and Jérôme Plût. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. *J. Math. Cryptol.*, 8:209–247, 2014.
9. Whitfield Diffie and Martin Hellman. New Directions in Cryptography. *IEEE Trans. Inf. Theor.*, 22(6):644–654, September 2006.
10. Edward Eaton and Fang Song. Making Existential-Unforgeable Signatures Strongly Unforgeable in the Quantum Random-Oracle Model. In *Proceedings of TQC*, pages 1–16, Germany, 2015. Dagstuhl Publishing.
11. Tommaso Gagliardoni, Andreas Hülsing, and Christian Schaffner. Semantic security and indistinguishability in the quantum world. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology — CRYPTO 2016*, pages 60–89, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.
12. Robin Hartshorne. *Algebraic Geometry*. Number 52 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1977.
13. A.S. Holevo. An analog of the theory of statistical decisions in noncommutative probability theory. *Trans. Mosc. Math. Soc.*, 26:133–149, 1972.
14. Marc Kaplan, Gaëtan Leurent, Anthony Leverrier, and María Naya-Plasencia. Breaking symmetric cryptosystems using quantum period finding. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology — CRYPTO 2016*, pages 207–237, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.
15. Hugo Krawczyk and Tal Rabin. Chameleon Hashing and Signatures, 1997. Manuscript.
16. Brian LaMacchia, Kristin Lauter, and Anton Mityagin. Stronger Security of Authenticated Key Exchange. In Susilo, Willy and Liu, Joseph K. and Mu, Yi, editor, *Provable Security: First International Conference*, pages 1–16, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
17. Dominic Mayers and Andrew Yao. Quantum cryptography with imperfect apparatus. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, FOCS '98, Washington, DC, USA, 1998. IEEE Computer Society.

18. Michele Mosca, Douglas Stebila, and Berkant Ustaoğlu. Quantum Key Distribution in the Classical Authenticated Key Exchange Framework. In Gaborit, Philippe, editor, *Post-Quantum Cryptography: 5th International Workshop*, pages 136–154, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
19. Joseph J. Rotman. *An Introduction to the Theory of Groups*. Number 148 in Graduate Texts in Mathematics. Springer, New York, 1995.
20. Joseph H. Silverman. *The Arithmetic of Elliptic Curves*. Number 106 in Graduate Texts in Mathematics. Springer, New York, 1986.
21. Vladimir Soukharev, David Jao, and Srinath Seshadri. Post-Quantum Security Models for Authenticated Encryption. In Takagi, Tsuyoshi, editor, *Post-Quantum Cryptography: 7th International Workshop*, pages 64–78, Cham, 2016. Springer International Publishing.
22. Xi Sun, Haibo Tian, and Yumin Wang. Toward Quantum-resistant Strong Designated Verifier Signature. *Int. J. Grid Util. Comput.*, 5(2):80–86, 2014.
23. John Tate. Endomorphisms of Abelian Varieties Over Finite Fields. *Invent. Math.*, 2:134 – 144, 1966.
24. Dominique Unruh. Quantum Position Verification in the Random Oracle Model. In Garay, Juan A. and Gennaro, Rosario, editor, *Advances in Cryptology — CRYPTO 2014*, pages 1–18, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.
25. Umesh Vazirani and Thomas Vidick. Fully device-independent quantum key distribution. *Phys. Rev. Lett.*, 113:140501, Sep 2014.
26. Jacques Vélu. Isogénies Entre Courbes Elliptiques. *C. R. Acad. Sci. Paris Sér. A-B*, 273:A238 – A241, 1971.
27. Han Weiwei and He Debiao. An authenticated key agreement protocol using isogenies between elliptic curves. *2010 Second International Workshop on Education Technology and Computer Science*, 1:366–369, 2010.
28. Gustavo H. M. Zanon, Marcos A. Simplicio Jr., Geovandro C. C. F. Pereira, Javad Doliskani, and Paulo S. L. M. Barreto. Faster isogeny-based compressed key agreement. Cryptology ePrint Archive, Report 2017/1143, 2017. https://eprint.iacr.org/2017/1143.
29. Mark Zhandry. Secure Identity-Based Encryption in the Quantum Random Oracle Model. In *Advances in Cryptology — CRYPTO 2012*, 2012.
30. Jiang Zhang, Zhenfeng Zhang, Jintai Ding, Michael Snook, and Özgür Dagdelen. Authenticated key exchange from ideal lattices. In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015*, pages 719–751. Springer Berlin Heidelberg, 2015.

A Figures

A.1 A Depiction of the j-Invariant Computation for our Protocol

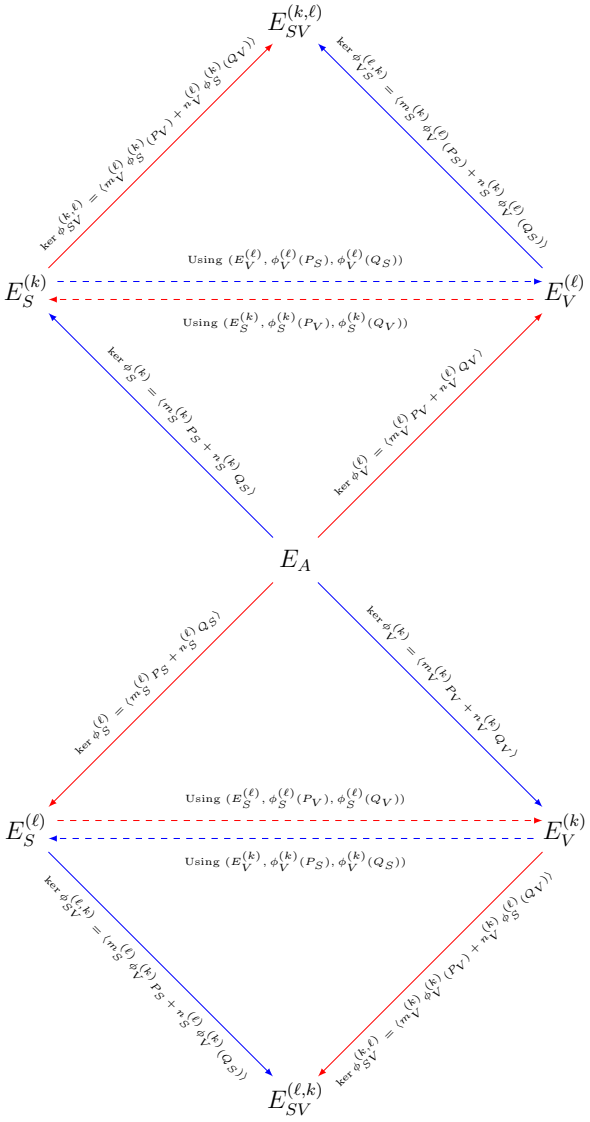

Fig. 1. A diagram explaining the j-invariant computation for the authentication portion. If $\mathcal{P}_k$ wishes to send a message to $\mathcal{P}_\ell$, they follow the blue and red arrows, respectively, in the upper half of the diagram to sign and verify. If $\mathcal{P}_\ell$ wishes to send a message to $\mathcal{P}_k$, they use the bottom half of the diagram.

A.2 A Depiction of a Pair of Matching Sessions for a Protocol Generated from our Construction (Theorem 1)

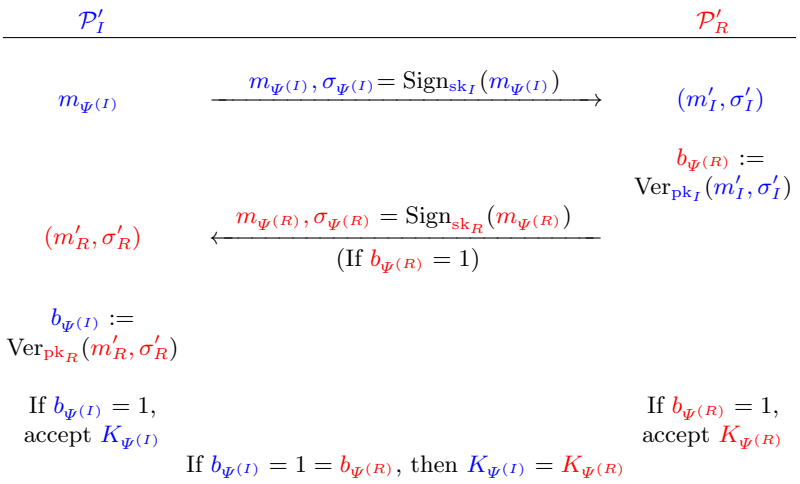

Fig. 2. An execution of a secure AKE protocol constructed as in Theorem 1. Values in blue are computed by the initiating party, while those in red are computed by the responding party.

B Proof of Theorem 1

To prove this result, we must first show how, given an instance (pk) of the strongly EUF-qCMA game, we can emulate a quantum messaging oracle $O_{\mathcal{P}_k}$ for a party $\mathcal{P}_k$ with public key pk for $\mathcal{S}$.

For an unauthenticated key exchange protocol Π, let the parties be denoted by $\mathcal{P}_k$ for some values of k, and for each such party let $\mathcal{P}'_k$ denote the corresponding party for protocol Π' defined as in Theorem 1. Notice that

$$\mathcal{P}'_k(m,\sigma) = \begin{cases} (\bot,\bot) \text{ if } \mathrm{Ver}_{\mathrm{pk}_{\mathcal{P}'}}(m,\sigma) = 0 \\ (\mathcal{P}_k(m), \mathrm{Sign}_{\mathrm{sk}_{\mathcal{P}_k}}(\mathcal{P}_k(m)) \text{ otherwise} \end{cases}.$$

Then, to emulate the quantum messaging oracle, first write $\mathcal{P}(m)$ to an auxiliary register to obtain

$$\sum_{m,\sigma,y} \alpha_{m,\sigma,y} \left|m\right\rangle \left|\sigma\right\rangle \left|\mathcal{P}(m)\right\rangle \left|y\right\rangle .$$

Use the strongly EUF-qCMA signing oracle on the third register to obtain

$$\sum_{m,\sigma,y} \alpha_{m,\sigma,y} \left|m\right\rangle \left|\sigma\right\rangle \left|\mathcal{P}(m)\right\rangle \left|\mathrm{Sign}_{\mathrm{sk}}(\mathcal{P}(m))\right\rangle \left|y\right\rangle .$$

Then apply U_{Ver} which maps $|m'\rangle\,|\sigma'\rangle\,|y\rangle\,|z\rangle$ to

$$\begin{cases} |m'\rangle\,|\sigma'\rangle\,|y \oplus m'\rangle\,|z \oplus \sigma'\rangle & \text{if } \mathrm{Ver}(m,\sigma) = 1 \\ |m'\rangle\,|\sigma'\rangle\,|y \oplus \perp\rangle\,|z \oplus \perp\rangle & \text{otherwise} \end{cases}$$

to the last registers to obtain

$$\sum_{m,\sigma,y} \alpha_{m,\sigma,y}\,|m\rangle\,|\sigma\rangle\,|\mathcal{P}(m)\rangle\,|\mathrm{Sign}_{\mathrm{sk}}(\mathcal{P}(m))\rangle\,|y \oplus \mathcal{P}(m,\sigma)\rangle\,;$$

the last register is the one we give to $\mathcal{A}$. Note in particular that we are holding onto the registers that contain valid message/signature pairs; in fact, we hold one such pair of registers for each query we make to the signing oracle. It follows that if we can persuade $\mathcal{A}$ to send us a pair of registers which, when measured, yield a valid message-signature pair different from those that we will obtain by measuring the registers we already hold, then with non-negligible probability we can win the strongly EUF-qCMA game.

Knowing that we can use the quantum signing oracle for the strongly EUF-qCMA game to emulate a quantum messaging oracle for a party, we show that the security of the signature scheme restricts the class of messages that an adversary can construct. The results are presented in the following technical lemmas.

We first demonstrate two restrictions that we can place on the behaviour of $\mathcal{A}$ without loss of generality.

Lemma 1. *Suppose there is an adversary $\mathcal{A}$ who wins the security game for Π' with advantage* Adv *who delivers two or more messages to a single party in a given session. Then there is an adversary $\mathcal{A}'$ who wins the security game for Π' with the same advantage who never delivers two or more messages to a single party in a given session.*

Proof. The response to any message delivered to a party in a given session beyond the first is $(\perp, \perp)$; this is because Π is a two-round key establishment protocol and so any message delivered beyond the first is invalid. Let $\mathcal{A}'$ be defined as $\mathcal{A}$ is, except that whenever $\mathcal{A}'$ would deliver the second message to a party in a given session, it instead simply writes $(\perp, \perp)$ to its target register. It is clear that $\mathcal{A}'$ wins the security game with the same probability as $\mathcal{A}$.

Lemma 2. *Suppose there is an adversary $\mathcal{A}$ who wins the security game for Π' with advantage* Adv*, who at some point sends the last register of the global state*

$$|\Gamma\rangle = \sum_{\boldsymbol{\mu},m,\sigma} \alpha_{\boldsymbol{\mu},m,\sigma}\,|\boldsymbol{\mu}\rangle\,|m,\sigma\rangle$$

to a responding party $\mathcal{P}_R$ in session Ψ', who believes they are participating in a session Ψ with initiating party $\mathcal{P}_I$, such that

$$\sum_{\mathrm{Ver}_{\mathrm{sk}_I}(m,\sigma)=1} \sum_{\boldsymbol{\mu}} |\alpha_{\boldsymbol{\mu},m,\sigma}|^2$$

is negligible. Then there is an adversary $\mathcal{A}'$ who wins the security game with advantage Adv' *which differs only negligibly from* Adv*, such that $\mathcal{A}'$ never sends a register in such a state.*

Proof. Let $\mathcal{A}$ be such an adversary. First we show that Ψ' cannot possibly be the session on which $\mathcal{A}$ will choose to be tested. Suppose to the contrary that Ψ' is the test session. The global state after the `Test` query will be

$$|\Gamma\rangle = \sum_{\mathrm{Ver}_{\mathrm{sk}_I}(m,\sigma)=0} \sum_{\boldsymbol{\mu}} \alpha_{\boldsymbol{\mu},m,\sigma} |\boldsymbol{\mu}\rangle |m,\sigma\rangle |\perp\rangle + \sum_{\mathrm{Ver}_{\mathrm{sk}_I}(m,\sigma)=1} \sum_{\boldsymbol{\mu}} \alpha_{\boldsymbol{\mu},m,\sigma} |\boldsymbol{\mu}\rangle |m,\sigma\rangle |\kappa_b(m)\rangle$$

where $\kappa_b(m)$ is either a correct key for session Ψ' on incoming message m, or a random string; in particular, in either case is it not $\perp$. Consider the state

$$|\Gamma'\rangle = \sum_{\mathrm{Ver}_{\mathrm{sk}_I}(m,\sigma)=0} \sum_{\boldsymbol{\mu}} \alpha_{\boldsymbol{\mu},m,\sigma} |\boldsymbol{\mu}\rangle |m,\sigma\rangle |\perp\rangle + \sum_{\mathrm{Ver}_{\mathrm{sk}_I}(m,\sigma)=1} \sum_{\boldsymbol{\mu}} \alpha_{\boldsymbol{\mu},m,\sigma} |\boldsymbol{\mu}\rangle |m,\sigma\rangle |r(m)\rangle$$

for randomly chosen strings $r(m)$. In particular, observe that the ensembles $\mathcal{D} = \{D_{\lambda,r}\}$ and $\mathcal{D}' = \{D_{\lambda,r}\}$ of measurement outcomes of $|\Gamma\rangle$ and $|\Gamma'\rangle$ (parameterized by the security parameter and random input) are computationally indistinguishable by the previous lemma, since if they were not, we could distinguish $|\Gamma\rangle$ from $|\Gamma'\rangle$ with non-negligible advantage. In particular, in this case this means that if $\mathcal{A}$ were instead given $|\Gamma'\rangle$, and then performed his measurement in order to guess the value b, the result would, except with negligible probability, be indistinguishable from the result of measuring $|\Gamma\rangle$, *regardless of the value of b*. Since $|\Gamma'\rangle$ carries no information about b, $\mathcal{A}$ cannot possibly guess b by measuring $|\Gamma'\rangle$ with probability different from $\frac{1}{2}$. Thus when measuring $|\Gamma\rangle$ and guessing, $\mathcal{A}$ guesses correctly with probability at most negligibly greater than $\frac{1}{2}$, contradicting our assumption. Hence Ψ' cannot be the test session.

By a similar argument, $\mathcal{A}$ can construct a register that is indistinguishable from the response $\mathcal{P}_R$ would give on this input register. Hence $\mathcal{A}'$ proceeds exactly as $\mathcal{A}$ would, except that whenever he would send a register as described in the statement of the lemma, he instead constructs an indistinguishable register.

Since $\mathcal{A}$ deals with at most polynomially-many registers, we can make as many substitutions of this kind as required and the probability that the resultant state is distinguishable from the correct state is negligible; hence, $\mathcal{A}'$, defined in this way, wins the security game with advantage at most negligibly different from Adv, as required.

Hence we can assume without loss of generality that our adversary $\mathcal{A}$ never delivers more than one message in a session and never delivers a superposition of messages for which the total probability amplitude of the valid content is

negligible. This will allow us to use an adversary $\mathcal{A}$ who delivers a superposition of messages in a session for which the amplitude of a valid, but unsent, message is non-negligible as a forger for a signature scheme; this tells us then that the probability of the adversary delivering such a message is negligible.

Lemma 3. *Let $\mathcal{A}$ be an adversary which wins the security game with non-negligible advantage. Let*

$$|\Gamma\rangle = \sum_{\boldsymbol{\mu}, m_I, \sigma_I} \alpha_{\boldsymbol{\mu}, m_I, \sigma_I} |\boldsymbol{\mu}\rangle |m_I, \sigma_I\rangle$$

be the global state after the last register is delivered by $\mathcal{A}$ to $\mathcal{P}_R$ in a clean session. Further, let (m^*, σ^*) be the message and signature that $\mathcal{P}_I$ would send in this session. Let

$$\mathscr{F} = \{(m_I, \sigma_I) \colon \mathrm{Ver}_{\mathrm{pk}_I}(m_I, \sigma_I) = 1 \text{ and } (m_I, \sigma_I) \neq (m_I^*, \sigma_I^*)\}$$

of potential "forgeries." If (Gen, Sign, Ver) *is strongly EUF-qCMA, then except with negligible probability, the quantity*

$$\Phi \equiv \sum_{\boldsymbol{\mu}, (m_I, \sigma_I) \in \mathscr{F}} |\alpha_{\boldsymbol{\mu}, m_I, \sigma_I}|^2$$

is negligible.

Proof. We show how to forge a signature against (Gen, Sign, Ver) in the strongly EUF-qCMA game if Φ is non-negligible.

Suppose we are given an instance (pk) of the EUF-qCMA game. We will run $\mathcal{A}$ essentially as normal, by establishing public key/private key pairs for as many parties as $\mathcal{A}$ requires; for one party $\mathcal{P}_{i^*}$ chosen at random, however, we will set their private key as pk (and the underlying secret key will remain unknown to us). With probability at least $\frac{1}{p(\lambda)}$, where $p(\lambda)$ is a bound on the number of parties $\mathcal{A}$ requires (and which is at most polynomial in λ), we have selected the initiator of this clean session. In particular this means that, at least until the session is over, $\mathcal{A}$ will not issue `RequestPrivateKey`($\mathtt{id}_{i^*}$), and so we will not have to produce it. Whenever a party needs to sign a message we use their private key, unless that party is $\mathcal{P}_{i^*}$, in which case we simply use the signing oracle from the strongly EUF-qCMA game, as described above. Notice that by our assumption that $\mathcal{A}$ never delivers two or more messages to a party in the same session, each time we query the signing oracle we are querying it for a different session; since the session identifier is included in the signed message, this means that, in particular, if we measure the results of our queries to the signing oracle we will never obtain the same message/signature pairs. Moreover, because of the construction we use to model the party from the signing oracle, each use of the signing oracle results in a register which will, with probability 1, yield a valid message/signature pair upon measurement. In particular, this means that if $\mathcal{A}$ ever sends us a superposition of messages for which the probability amplitude

of a forged message is non-negligible, then by measuring that register and the registers we hold, we will obtain $q+1$ distinct valid message/signature pairs, where q is the number of calls we have made to the signing oracle.

$\mathcal{A}$ will perform some unitary operations on the qubits he holds; thus the global state becomes

$$|\Gamma\rangle = \sum_{\boldsymbol{\mu},m_I,\sigma_I} \alpha_{\boldsymbol{\mu},m_I,\sigma_I} |\boldsymbol{\mu}\rangle |m^*,\sigma^*\rangle |m_I,\sigma_I\rangle$$

and $\mathcal{A}$ sends the last register to $\mathcal{P}_R$ (*i.e.,* to us). If we now measure the qubits we hold, then with probability Φ, we will obtain $q+1$ valid message/signature pairs. If Φ is non-negligible, we can win the strongly EUF-qCMA game for our signature scheme; since the signature scheme is strongly EUF-qCMA, this forgery can occur with at most negligible probability, and so the probability that $\frac{\Phi}{p(\lambda)}$ is non-negligible (and hence that Φ is non-negligible) is negligible, as required.

Lemma 4. *Let*

$$|\Gamma\rangle = \sum_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R} \alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R} |\boldsymbol{\mu}\rangle |m_I,\sigma_I\rangle |m_R,\sigma_R\rangle$$

be the global state after the completion of a clean session Ψ owned by $\mathcal{P}_R$, the responding party, where the second register is the message register sent by $\mathcal{A}$ to $\mathcal{P}_R$, and the third is the message register sent by $\mathcal{A}$ to $\mathcal{P}_I$, the initiating party, if it exists. Let (m_I^*,σ_I^*) be the message and signature that would actually be sent by $\mathcal{P}_I$ in step 3e of the protocol, and let (m_R^*,σ_R^*) be the message and signature that $\mathcal{P}_R$ would respond with if the messages were relayed faithfully. Let $\mathscr{F}$ denote the set

$$\{(m_I,\sigma_I,m_R,\sigma_R)\colon (\exists \iota \in \{I,R\}\colon \mathrm{Ver}_{\mathrm{pk}_\iota}(m_\iota,\sigma_\iota) = 1 \text{ and } (m_\iota,\sigma_\iota) \neq (m_\iota^*,\sigma_\iota^*))\}$$

of potential tuples containing a "forged" signature. If (Gen, Sign, Ver) *is strongly EUF-qCMA, then, except with negligible probability, the quantity*

$$\Phi \equiv \sum_{(m_I,\sigma_I,m_R,\sigma_R)\in\mathscr{F}} \sum_{\boldsymbol{\mu}} |\alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R}|^2$$

is negligible.

Proof. Define

$$\mathscr{F}_I = \{(m_I,\sigma_I,m_R,\sigma_R) \in \mathscr{F}\colon \mathrm{Ver}_{\mathrm{pk}_I}(m_I,\sigma_I) = 1 \wedge (m_I,\sigma_I) \neq (m_I^*,\sigma_I^*)\}$$

and

$$\mathscr{F}_{\neg I} = \{(m_I,\sigma_I,m_R,\sigma_R) \in \mathscr{F}\colon \mathrm{Ver}_{\mathrm{pk}_I}(m_I,\sigma_I) = 0 \vee (m_I,\sigma_I) = (m_I^*,\sigma_I^*)\}$$

and observe that $\mathscr{F} = \mathscr{F}_I \cup \mathscr{F}_{\neg I}$ and that this union is disjoint. By Lemma 3, we know that

$$\Phi_I = \sum_{(m_I,\sigma_I,m_R,\sigma_R)\in\mathscr{F}_I} \sum_{\boldsymbol{\mu}} |\alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R}|^2$$

is negligible. Then we need only prove that

$$\Phi_{\neg I} = \sum_{(m_I,\sigma_I m_R,\sigma_R)\in\mathscr{F}_{\neg I}} \sum_{\boldsymbol{\mu}} |\alpha_{\boldsymbol{\mu},m_R,\sigma_R,m_I,\sigma_I}|^2$$

is negligible, since $\Phi = \Phi_I + \Phi_{\neg I}$. As in the proof of lemma 3, we can try to win the EUF-qCMA game by choosing a random party $\mathcal{P}$ and hoping that $\mathcal{A}$ "forges" a signature against them; then we measure the registers we hold to obtain a forgery. This succeeds with probability at least $\Phi_{\neg I}$, and so this quantity must be negligible except with negligible probability, as required.

Lemma 5. *Let*

$$|\Gamma\rangle = \sum_{m\in M} \alpha_m |m\rangle + \sum_{c\in C} \alpha_c |c\rangle \text{ and } |\Gamma'\rangle = \sum_{m\in M} \alpha_m |m\rangle + \sum_{c\in C} \alpha_c |r(c)\rangle$$

be normalized quantum states, where M and C are disjoint, nonempty, finite sets, $r(c) \notin M$ for all $c \in C$, and $\sum_{c\in C} |\alpha_c|^2$ is negligible. Then the advantage that any adversary has in distinguishing $|\Gamma\rangle$ from $|\Gamma'\rangle$ is negligible.

Proof. Note that $1 - |\langle\Gamma|\,\Gamma'\rangle|^2 \leq 2\sum_{c\in C} |\alpha_c|^2$. The result then follows from the Holevo-Helstrom theorem [13, Sections 5–7].

Lemma 6. *Consider the following state distinguishing game for some quantum states states $|\Gamma_0\rangle$, $|\Gamma_1\rangle$, $|\Gamma'_0\rangle$ and $|\Gamma'_1\rangle$:*

i. $\mathcal{C}$ selects $b \in \{0,1\}$ uniformly at random, and sends $|\Gamma_b\rangle$ to $\mathcal{A}$.
ii. $\mathcal{A}$ performs some computations and outputs a guess b'.
iii. $\mathcal{A}$ wins if $b' = b$.

Let $\mathcal{A}$ be an adversary for this game, and now consider the following game:

i. $\mathcal{C}$ selects $b \in \{0,1\}$ uniformly at random, and sends $|\Gamma'_b\rangle$ to $\mathcal{A}$.
ii. $\mathcal{A}$ performs some computations and outputs a guess b'.
iii. $\mathcal{A}$ wins if $b' = b$.

The probability that $\mathcal{A}$ wins the second game differs from the probability that $\mathcal{A}$ wins the first game by at most

$$\frac{1}{2}\left(\sqrt{1 - |\langle\Gamma_0|\,\Gamma'_0\rangle|^2} + \sqrt{1 - |\langle\Gamma_1|\,\Gamma'_1\rangle|^2}\right)$$

Proof. Consider the problem of distinguishing $|\Gamma_0\rangle$ from $|\Gamma'_0\rangle$. By the Holevo-Helstrom Theorem the advantage that any procedure has in distinguishing these two states is at most $\frac{1}{2}\sqrt{1 - |\langle\Gamma_0|\,\Gamma'_0\rangle|^2}$.

Consider the following distinguishing procedure: given a state $|\Gamma''_0\rangle$ which is either in the state $|\Gamma_0\rangle$ or $|\Gamma'_0\rangle$, each with probability $\frac{1}{2}$, give the state to $\mathcal{A}$. If

$\mathcal{A}$ produces the guess $b' = 0$, guess that the state is $|\Gamma_0\rangle$, and otherwise guess that the state is $|\Gamma_0'\rangle$. Then

$$\begin{aligned}\mathbb{P}[\text{This Method is Correct}] &= \mathbb{P}[b' = 0 \wedge |\Gamma_0''\rangle = |\Gamma_0\rangle] + \mathbb{P}[b' = 1 \wedge |\Gamma_0''\rangle = |\Gamma_0'\rangle] \\ &= \mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0\rangle] \cdot \mathbb{P}[|\Gamma_0''\rangle = |\Gamma_0\rangle] \\ &\quad + \mathbb{P}[b' = 1 \,|\, |\Gamma_0''\rangle = |\Gamma_0'\rangle] \cdot \mathbb{P}[|\Gamma_0''\rangle = |\Gamma_0'\rangle] \\ &= \frac{1}{2}\mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0\rangle] + \frac{1}{2}\left(1 - \mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0'\rangle]\right)\end{aligned}$$

so that

$$\mathbb{P}[\text{This Method is Correct}] = \frac{1}{2} + \frac{1}{2}\left(\mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0\rangle] - \mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0'\rangle]\right)$$

If this is not at least $\frac{1}{2}$, we obtain a better procedure by switching our guesses; in any case, there is a procedure that can be used to distinguish $|\Gamma_0\rangle$ from $|\Gamma_0'\rangle$ with advantage $\frac{1}{2}\left(\mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0\rangle] - \mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0'\rangle]\right)$, and so

$$|\mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0\rangle] - \mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0'\rangle]| \leq \sqrt{1 - |\langle\Gamma_0|\,\Gamma_0'\rangle|^2}.$$

A similar argument gives that

$$|\mathbb{P}[b' = 1 \,|\, |\Gamma_1''\rangle = |\Gamma_1\rangle] - \mathbb{P}[b' = 1 \,|\, |\Gamma_1''\rangle = |\Gamma_1'\rangle]| \leq \sqrt{1 - |\langle\Gamma_1|\,\Gamma_1'\rangle|^2}.$$

Let $|\Gamma\rangle$ be the state given to $\mathcal{A}$ in the first game, and $|\Gamma'\rangle$ be the state given to $\mathcal{A}$ in the second game. Then

$$\begin{aligned}&|\mathbb{P}[\mathcal{A} \text{ wins the first game}] - \mathbb{P}[\mathcal{A} \text{ wins the second game}]| \\ &\quad = |\mathbb{P}[b' = 0 \wedge |\Gamma\rangle = |\Gamma_0\rangle] + \mathbb{P}[b' = 1 \wedge |\Gamma\rangle = |\Gamma_1\rangle] \\ &\qquad - \mathbb{P}[b' = 0 \wedge |\Gamma'\rangle = |\Gamma_0'\rangle] - \mathbb{P}[b' = 1 \wedge |\Gamma'\rangle = |\Gamma_1'\rangle]| \\ &\quad \leq |\mathbb{P}[b' = 0 \,|\, |\Gamma\rangle = |\Gamma_0\rangle]\mathbb{P}[|\Gamma\rangle = |\Gamma_0\rangle] \\ &\qquad\quad - \mathbb{P}[b' = 0 \,|\, |\Gamma_0''\rangle = |\Gamma_0'\rangle]\mathbb{P}[|\Gamma'\rangle = |\Gamma_0\rangle]| \\ &\qquad + |\mathbb{P}[b' = 1 \,|\, |\Gamma\rangle = |\Gamma_1\rangle]\mathbb{P}[|\Gamma\rangle = |\Gamma_1\rangle] \\ &\qquad\quad - \mathbb{P}[b' = 1 \,|\, |\Gamma_1''\rangle = |\Gamma_1'\rangle]\mathbb{P}[|\Gamma'\rangle = |\Gamma_1\rangle]| \\ &\quad \leq \frac{1}{2}\left(\sqrt{1 - |\langle\Gamma_0|\,\Gamma_0'\rangle|^2} + \sqrt{1 - |\langle\Gamma_1|\,\Gamma_1'\rangle|^2}\right)\end{aligned}$$

as required.

Corollary 1. *Suppose the global state in an instance of the security experiment for the protocol just before the* `Test` *query is issued by* $\mathcal{A}$ *be*

$$\begin{aligned}|\Gamma\rangle =& \sum_{\boldsymbol{\mu}} \alpha_{\boldsymbol{\mu},m_I^*,\sigma_I^*,m_R^*,\sigma_R^*} |\boldsymbol{\mu}\rangle |m_I^*,\sigma_I^*\rangle |m_R^*,\sigma_R^*\rangle \\ &+ \sum_{\substack{\boldsymbol{\mu} \\ \mathrm{Ver}_{\mathrm{pk}_I}(m_I,\sigma_I)=0 \\ or\ \mathrm{Ver}_{\mathrm{pk}_R}(m_R,\sigma_R)=0}} \alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R} |\boldsymbol{\mu}\rangle |m_I^*,\sigma_I^*\rangle |m_R^*,\sigma_R^*\rangle \\ &+ \sum_{\substack{\boldsymbol{\mu} \\ (m_I,\sigma_I,m_R,\sigma_R)\in\mathscr{F}}} \alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R} |\boldsymbol{\mu}\rangle |m_I^*,\sigma_I^*\rangle |m_R^*,\sigma_R^*\rangle\end{aligned}$$

where (m_I^*,σ_I^*) *is the message/signature pair that would actually have been sent by the initiating party in the test session, and* (m_R^*,σ_R^*) *is the corresponding response. After the* `Test` *query is issued, the challenger selects* b *uniformly at random from* $\{0,1\}$*, and should return the last three registers of the global state*

$$\begin{aligned}|\Gamma_b\rangle =& \sum_{\boldsymbol{\mu}} \alpha_{\boldsymbol{\mu},m_I^*,\sigma_I^*,m_R^*,\sigma_R^*} |\boldsymbol{\mu}\rangle |m_I^*,\sigma_I^*\rangle |m_R^*,\sigma_R^*\rangle |\kappa_b(m_I^*,m_R^*)\rangle \\ &+ \sum_{\substack{\boldsymbol{\mu} \\ \mathrm{Ver}_{\mathrm{pk}_I}(m_I,\sigma_I)=0 \\ or\ \mathrm{Ver}_{\mathrm{pk}_R}(m_R,\sigma_R)=0}} \alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R} |\boldsymbol{\mu}\rangle |m_I,\sigma_I\rangle |m_R,\sigma_R\rangle |\perp\rangle \\ &+ \sum_{\substack{\boldsymbol{\mu} \\ (m_I,\sigma_I,m_R,\sigma_R)\in\mathscr{F}}} \alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R} |\boldsymbol{\mu}\rangle |m_I,\sigma_I\rangle |m_R,\sigma_R\rangle |\kappa_b(m_I,m_R)\rangle\end{aligned}$$

to $\mathcal{A}$*, where as before* $\kappa_0(m_I,m_R)$ *is the session key corresponding to messages* m_I,m_R *and* $\kappa_1(m_I,m_R)$ *is simply a random function. If instead* $\mathcal{C}$ *returns the last three registers of the state*

$$\begin{aligned}|\Gamma'_b\rangle =& \sum_{\boldsymbol{\mu}} \alpha_{\boldsymbol{\mu},m_I^*,\sigma_I^*,m_R^*,\sigma_R^*} |\boldsymbol{\mu}\rangle |m_I^*,\sigma_I^*\rangle |m_R^*,\sigma_R^*\rangle |\kappa_b(m_I^*,m_R^*)\rangle \\ &+ \sum_{\substack{\boldsymbol{\mu} \\ \mathrm{Ver}_{\mathrm{pk}_I}(m_I,\sigma_I)=0 \\ or\ \mathrm{Ver}_{\mathrm{pk}_R}(m_R,\sigma_R)=0}} \alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R} |\boldsymbol{\mu}\rangle |m_I,\sigma_I\rangle |m_R,\sigma_R\rangle |\perp\rangle \\ &+ \sum_{\substack{\boldsymbol{\mu} \\ (m_I,\sigma_I,m_R,\sigma_R)\in\mathscr{F}}} \alpha_{\boldsymbol{\mu},m_I,\sigma_I,m_R,\sigma_R} |\boldsymbol{\mu}\rangle |m_I,\sigma_I\rangle |m_R,\sigma_R\rangle |\kappa_1(m_I,m_R)\rangle\end{aligned}$$

then except with negligible probability, the probability that $\mathcal{A}$ *guesses the value of* b *correctly given this state differs at most negligibly from the probability that* $\mathcal{A}$ *guesses the value of* b *correctly given* $|\Gamma_b\rangle$*, regardless of the value of* b*.*

Corollary 2. *Let $|\Gamma\rangle$ be drawn from one of the following distributions, each with probability $\frac{1}{2}$:*

$$\Delta : \sum_{\boldsymbol{\mu}} \alpha^*_{\boldsymbol{\mu}} |\boldsymbol{\mu}\rangle |m^*\rangle + \sum_{\boldsymbol{\mu};\, m\in M} \alpha_{\boldsymbol{\mu},m} |\boldsymbol{\mu}\rangle |m\rangle + \sum_{\boldsymbol{\mu};\, c\in C} \alpha_{\boldsymbol{\mu},c} |\boldsymbol{\mu}\rangle |c\rangle \text{ for } m^* \leftarrow \mathcal{D}, \text{ and}$$

$$\hat{\Delta} : \sum_{\boldsymbol{\mu}} \alpha^*_{\boldsymbol{\mu}} |\boldsymbol{\mu}\rangle |\hat{m}^*\rangle + \sum_{\boldsymbol{\mu};\, m\in M} \alpha_{\boldsymbol{\mu},m} |\boldsymbol{\mu}\rangle |m\rangle + \sum_{\boldsymbol{\mu};\, c\in C} \alpha_{\boldsymbol{\mu},c} |\boldsymbol{\mu}\rangle |r(c)\rangle \text{ for } \hat{m}^* \leftarrow \hat{\mathcal{D}}$$

where $\mathcal{D}$ and $\mathcal{D}'$ are probability distributions on some set, M, C, and the supports of $\mathcal{D}$ and $\hat{\mathcal{D}}$ are disjoint, finite, nonempty sets, $r(c) \notin M$ for all $c \in C$, and $\sum_{\boldsymbol{\mu};\, c\in C} |\alpha_c|^2$ is negligible. Then if there is an efficient quantum adversary $\mathcal{A}$ which determines from which distribution $|\Gamma\rangle$ is drawn with non-negligible advantage Adv*, then there is an efficient procedure, using $\mathcal{A}$ as a subroutine, which distinguishes $\mathcal{D}$ from $\hat{\mathcal{D}}$ with non-negligible advantage.*

Proof. Let $\mathcal{A}$ be as described. Suppose you are given $\tilde{m}$ and wish to know from which distribution it is drawn. Construct the state

$$|\Gamma'\rangle = \sum_{\boldsymbol{\mu}} \alpha^*_{\boldsymbol{\mu}} |\boldsymbol{\mu}\rangle |\tilde{m}\rangle + \sum_{\substack{\boldsymbol{\mu} \\ m\in M}} \alpha_{\boldsymbol{\mu},m} |\boldsymbol{\mu}\rangle |m\rangle + \sum_{\substack{\boldsymbol{\mu} \\ c\in C}} \alpha_{\boldsymbol{\mu},c} |\boldsymbol{\mu}\rangle |c\rangle .$$

Notice that if $\tilde{m}$ is drawn from $\mathcal{D}$ then $|\Gamma'\rangle$ is drawn from Δ, while if $\tilde{m}$ is drawn from $\hat{\mathcal{D}}$, then $\sqrt{1 - |\langle\Gamma|\,\Gamma'\rangle|^2}$ is negligible; hence the probability that $\mathcal{A}$ wins the game given $|\Gamma'\rangle$ differs only negligibly from the probability that $\mathcal{A}$ wins the game given a true sample from Δ or $\hat{\Delta}$ by Lemma 6.

We will guess that $\tilde{m}$ is drawn from $\mathcal{D}$ if $\mathcal{A}$ guesses that $|\Gamma'\rangle$ is drawn from Δ, and we guess that $\tilde{m}$ is drawn from $\hat{\mathcal{D}}$ if $\mathcal{A}$ guesses that $|\Gamma'\rangle$ is drawn from $\hat{\Delta}$. The probability that we guess correctly is then

$$\begin{aligned}
\mathbb{P}[\text{We guess correctly}] &= \mathbb{P}[\mathcal{A} \text{ guesses } \Delta | \tilde{m} \leftarrow \mathcal{D}] + \mathbb{P}[\mathcal{A} \text{ guesses } \hat{\Delta} | \tilde{m} \leftarrow \hat{\mathcal{D}}] \\
&\geq \mathbb{P}[\mathcal{A} \text{ guesses } \Delta |\, |\Gamma\rangle \leftarrow \Delta] + \mathbb{P}[\mathcal{A} \text{ guesses } \hat{\Delta} |\, |\Gamma'\rangle \leftarrow \hat{\Delta}] \\
&\quad - |\mathbb{P}[\mathcal{A} \text{ is correct} |\, |\Gamma\rangle \leftarrow \Delta \text{ or } \hat{\Delta}] - \mathbb{P}[\mathcal{A} \text{ is correct} |\, |\Gamma\rangle = |\Gamma'\rangle]| \\
&\geq \frac{1}{2} + \text{Adv} - \epsilon
\end{aligned}$$

for a negligible function ϵ. Indeed, this procedure works with non-negligible advantage Adv $- \epsilon$, as required.

Finally we are able to prove the security of the protocol Π'.

Proof (Theorem 1). Suppose we are faced with an instance of the security game for Π; that is, given the messages sent by the initiator $\mathcal{P}_I$ in session Ψ with responder $\mathcal{P}_R$ in session Ψ' are m^*_I and m^*_R, respectively, we wish to determine whether a given string κ_b is the true session key for session Ψ if $b = 0$, or a random string if $b = 1$. Suppose to the contrary that there is an adversary $\mathcal{A}$ who wins the security game against Π' with non-negligible advantage Adv. We will use $\mathcal{A}$ as a distinguisher for our instance of the security game against Π.

Before starting an instance of $\mathcal{A}$, select two indices i^*, j^* which are less than the (polynomially-bounded) number of parties p that $\mathcal{A}$ will require. Further, choose a number s^* less than the (polynomially-bounded) number of pairs of session ψ that $\mathcal{A}$ will use. Run $\mathcal{A}$ essentially as usual, but with the following modifications.

Set the private key/public key information for $\mathcal{P}_{i^*}$ and $\mathcal{P}_{j^*}$ to that of $\mathcal{P}_I$ and $\mathcal{P}_R$, respectively. If $\mathcal{A}$ initiates fewer than s^* sessions, if its $s^{*\text{th}}$ session is not initiated by $\mathcal{P}_{i^*}$ with responder $\mathcal{P}_{j^*}$, or if its $s^{*\text{th}}$ session is not the test session, abort and select new i^*, j^* and s^*.

Given that $\mathcal{A}$ initiates the $s^{*\text{th}}$ with initiator $\mathcal{P}_{i^*}$ and responder P_{j^*}, set the session identifier as Ψ and the peer session identifier as Ψ'. Set the initiator's outgoing message as $(m_I^*, \sigma_I^* = \text{Sign}_{\text{sk}_I}(\texttt{id}_I, \texttt{id}_R; \Psi; m_I^*; r_\Psi))$ and the responder's message as (m_R^*, σ_R^*), where $\sigma_R^* = \text{Sign}_{\text{sk}_R}(\texttt{id}_I, \texttt{id}_R; \Psi, \Psi'; m_R^*; r_{\Psi'})$. If this is not eventually the test session, abort and choose i^*, j^* and s^* again; in particular, if $\mathcal{A}$ ever issues a command that would make either session no longer clean, abort.

By Lemma 4, we know that the adversary cannot construct a state for which the amplitude of a valid responding message is non-negligible; hence by Lemma 2 we know that $\mathcal{A}$ must pass some registers to the responding party $\mathcal{P}_{j^*}$, since otherwise the probability amplitude of states for which the session key obtained from the registers sent to the initiating party will be valid for Ψ is negligible, and there will be no matching session Ψ' to test. Moreover, the adversary must either deliver some response registers to the initiating party or the test session must be Ψ', since otherwise the adversary cannot win the game with non-negligible advantage. In either case, the global state before the test session is

$$\begin{aligned}|\Gamma\rangle = &\sum_{\boldsymbol{\mu}} \alpha_{\boldsymbol{\mu}, m_I^*, \sigma_I^*, m_R^*, \sigma_R^*} |\boldsymbol{\mu}\rangle |m_I^*, \sigma_I^*\rangle |m_R^*, \sigma_R^*\rangle \\ &+ \sum_{\substack{\text{Ver}_{\text{pk}_I}(m_I, \sigma_I)=0 \\ \text{or } \text{Ver}_{\text{pk}_R}(m_R, \sigma_R)=0}} \alpha_{\boldsymbol{\mu}, m_I, \sigma_I, m_R, \sigma_R} |\boldsymbol{\mu}\rangle |m_I, \sigma_I\rangle |m_R, \sigma_R\rangle \\ &+ \sum_{(m_I, \sigma_I, m_R, \sigma_R) \in \mathscr{F}} \alpha_{\boldsymbol{\mu}, m_I, \sigma_I, m_R, \sigma_R} |\boldsymbol{\mu}\rangle |m_I, \sigma_I\rangle |m_R, \sigma_R\rangle\end{aligned}$$

where the second register is the one delivered to $\mathcal{P}_{j^*}$ and the third is the one obtained from $\mathcal{P}_{j^*}$ in session Ψ', possibly after applying some unitary operator, and, except with negligible probability,

$$\Phi = \sum_{\substack{\boldsymbol{\mu} \\ (m_I, \sigma_I, m_R, \sigma_R) \in \mathscr{F}}} |\alpha_{\boldsymbol{\mu}, m_I, \sigma_I, m_R, \sigma_R}|^2$$

is negligible, again by Lemma 4.

When the `Test` query is issued, if the test session is Ψ apply the map:

$$|m_R, \sigma_R\rangle |y\rangle \mapsto \begin{cases} |m_R, \sigma_R\rangle |y \oplus \kappa_b\rangle & \text{if } (m_R \sigma_R) = (m_R^*, \sigma_R^*) \\ |m_R, \sigma_R\rangle |y \oplus \perp\rangle & \text{if } \text{Ver}_{\text{sk}_R}(m_R, \sigma_R) = 0 \\ |m_R, \sigma_R\rangle |y \oplus \rho(m_R)\rangle & \text{otherwise} \end{cases}$$

to the register received by $\mathcal{P}_{i^*}$ in session Ψ and the target register provided by $\mathcal{A}$, where ρ maps pairs of messages to random strings. If instead the test session is Ψ', apply the map

$$|m_I, \sigma_I\rangle\,|y\rangle \mapsto \begin{cases} |m_I, \sigma_I\rangle\,|y \oplus \kappa_b\rangle & \text{if } (m_I, \sigma_I) = (m_I^*, \sigma_I^*) \\ |m_I, \sigma_I\rangle\,|y \oplus \perp\rangle & \text{if } \mathrm{Ver}_{\mathrm{sk}_I}(m_I, \sigma_I) = 0 \\ |m_I, \sigma_I\rangle\,|y \oplus \rho(m_I)\rangle & \text{otherwise} \end{cases}$$

In either case, the global state after the test query is

$$\begin{aligned} |\Gamma\rangle = &\sum_{\boldsymbol{\mu}, y} \alpha_{\boldsymbol{\mu}, m_I^*, \sigma_I^*, m_R^*, \sigma_R^*}\,|\boldsymbol{\mu}\rangle\,|m_I^*, \sigma_I^*\rangle\,|m_R^*, \sigma_R^*\rangle\,|y \oplus \kappa_b\rangle \\ &+ \sum_{\substack{\boldsymbol{\mu}, y \\ \mathrm{Ver}_{\mathrm{pk}_I}(m_I, \sigma_I) = 0 \\ \text{or } \mathrm{Ver}_{\mathrm{pk}_R}(m_R, \sigma_R) = 0}} \alpha_{\boldsymbol{\mu}, m_I, \sigma_I, m_R, \sigma_R}\,|\boldsymbol{\mu}\rangle\,|m_I, \sigma_I\rangle\,|m_R, \sigma_R\rangle\,|y \oplus \perp\rangle \\ &+ \sum_{\substack{\boldsymbol{\mu}, y \\ (m_I, \sigma_I, m_R, \sigma_R) \in \mathscr{F}}} \alpha_{\boldsymbol{\mu}, m_I, \sigma_I, m_R, \sigma_R}\,|\boldsymbol{\mu}\rangle\,|m_I, \sigma_I\rangle\,|m_R, \sigma_R\rangle\,|y \oplus \rho'(m_I, m_R)\rangle. \end{aligned}$$

Notice that this is simply $|\Gamma_b'\rangle$ from Corollary 1, and so $\mathcal{A}$ will guess b correctly (in the context of its security game) with advantage $\mathrm{Adv} - \epsilon$ for some negligible function ϵ. Then, by Corollary 2, this correct guess will be the correct guess for our security game with advantage $\mathrm{Adv} - \epsilon - \epsilon'$ where ϵ' is negligible; in particular, our advantage is non-negligible provided that we have chosen the correct i^*, j^*, and s^*. Since we choose these correctly with probability at least $\frac{1}{p^2\psi}$, a polynomial fraction, we see that our probability of winning the game using $\mathcal{A}$ as a subroutine is at least $\frac{1}{2} + \frac{\mathrm{Adv} - \epsilon - \epsilon'}{p^2\psi}$ which is non-negligibly greater than a half; that is, protocol Π is insecure. This is a contradiction, and so no such adversary $\mathcal{A}$ must exist; that is, the protocol Π' is secure.