Post-quantum Resettably-Sound Zero Knowledge

Nir Bitansky, Michael Kellner, and Omri Shmueli(✉)

Tel-Aviv University, Tel Aviv, Israel
nirbitan@tau.ac.il,{kellner,omrishmueli}@mail.tau.ac.il

Abstract. We study post-quantum zero-knowledge (classical) protocols that are sound against *quantum resetting attacks*. Our model is inspired by the classical model of resetting provers (Barak-Goldreich-Goldwasser-Lindell, FOCS '01), providing a malicious efficient prover with oracle access to the verifier's *next-message-function*, fixed to some initial random tape; thereby allowing it to effectively reset (or equivalently, rewind) the verifier. In our model, the prover has *quantum access* to the verifier's function, and in particular can query it in superposition.

The motivation behind quantum resettable soundness is twofold: First, ensuring a strong security guarantee in scenarios where quantum resetting may be possible (e.g., smart cards, or virtual machines). Second, drawing intuition from the classical setting, we hope to improve our understanding of basic questions regarding post-quantum zero knowledge.

We prove the following results:

- **Black-Box Barriers.** Quantum resetting exactly captures the power of black-box zero knowledge quantum simulators. Accordingly, resettable soundness cannot be achieved in conjunction with black-box zero knowledge, except for languages in **BQP**. Leveraging this, we prove that constant-round public-coin, or three message, protocols cannot be black-box post-quantum zero-knowledge. For this, we show how to transform such protocols into quantumly resettably sound ones. The transformations are similar to classical ones, but their analysis is very different due to the essential difference between classical and quantum resetting.
- **A Resettably-Sound Non-Black-Box Zero-Knowledge Protocol.** Under the (quantum) Learning with Errors assumption and quantum fully-homomorphic encryption, we construct a post-quantum resettably-sound zero knowledge protocol for **NP**. We rely on non-black-box simulation techniques, thus overcoming the black-box barrier for such protocols.
- **From Resettable Soundness to The Impossibility of Quantum Obfuscation.** Assuming one-way functions, we prove that any quantumly-resettably-sound zero-knowledge protocol for **NP** implies the impossibility of quantum obfuscation. Combined with the above

A full version of this paper is available at https://eprint.iacr.org/2021/349.pdf.
M. Kellner—Member of the Check Point Institute of Information Security.

K. Nissim and B. Waters (Eds.): TCC 2021, LNCS 13042, pp. 62–89, 2021.
https://doi.org/10.1007/978-3-030-90459-3_3

result, this gives an alternative proof to several recent results on quantum unobfuscatability.

1 Introduction

Zero-knowledge protocols, introduced by Goldwasser, Micali, and Rackoff [GMR89], are a cornerstone of cryptography. They allow proving the validity of any statement in **NP** without revealing anything but its validity [GMW91]. After over three and a half decades of research, zero knowledge protocols are well understood in terms of their expressiveness and round complexity, and various enhancements of zero knowledge have been considered.

In this work, we consider zero knowledge protocols with *post-quantum security*, namely, protocols that can be executed by classical parties, but where both soundness and zero knowledge are guaranteed against efficient quantum adversaries. Starting from the seminal work of Watrous [Wat09], our understanding of post-quantum zero knowledge has been gradually improving, and yet it is still far behind our understanding of classical zero knowledge. Beyond the obvious need for post-quantum computational assumptions, the design and analysis of post-quantum zero knowledge protocols is challenged by quantum phenomena such as the no-cloning theorem [WZ82] and state disturbance [FP96], which often deem classical techniques insufficient.

Resettable Soundness. We focus on the notion of *resettable soundness*, introduced by Barak, Goldreich, Goldwasser, and Lindell [BGGL01] and by Micali and Reyzin [MR01]. In the classical setting, resettably-sound protocols remain sound even against a prover that has the ability to reset the honest verifier to its initial state and random tape, and repeat the interaction in any way it chooses (equivalent to the ability to rewind the verifier to any previous message). The threat of reset attacks arises in various settings, when fresh randomness cannot be generated on the fly and parties are subject to physical resets. Examples include verifiers that run on smart cards or virtual machines. Accordingly security against resetting attacks has received much attention [CGGM00, KP01, MR01] [DGS09, GS09, COSV12, OV12, COPV13, COP+14, BP15, CPS16].

Beyond the protection it provides in the above settings, resettable soundness has played an important role in understanding a foundational question regarding (classical) zero knowledge protocols—the gap between black box zero knowledge and non black box zero knowledge. In the first, the zero knowledge simulator can only access the verifier as a black box, whereas in the second, it can make explicit use of the verifier's code. Indeed, resettably-sound protocols cannot have a black-box zero knowledge simulator [BGGL01]; roughly speaking, this is because a resetting prover effectively has the same rewinding power as a zero knowledge simulator, and can accordingly use any black box simulation strategy in order to cheat. In fact, several other black-box zero knowledge impossibilities can be derived by a reduction to the impossibility of resettably sound black-box zero knowledge [GK96b, BGGL01, PTW11].

This Work: Quantum Resettable Soundness. We investigate resettable soundness in the quantum setting. That is, we consider classical protocols that are sound against *quantum resetting attacks* and (plain) zero knowledge against quantum malicious verifiers. Our goal is twofold: First, constructing such protocols to deal with resetting scenarios in a quantum world. Second, in light of the role that resettable soundness plays in the classical setting, we expect that in the quantum setting too, understanding resettable soundness would shed light on basic questions regarding post-quantum zero knowledge.

1.1 Contributions

We first model resetting attacks in a quantum world and define the corresponding notion of resettable soundness. We consider a strong definition that provides the resetting prover *quantum access* to the honest verifier's *next message function*, for some fixed verifier randomness. In particular, the resetting prover may not only rewind the verifier, but also do it in superposition. This model aims to capture the worst possible behavior of an efficient quantum attacker in a setting where resetting is possible. Furthermore, the model captures the capabilities of a black box zero knowledge simulator in the quantum setting (the model is further discussed in the technical overview). Throughout, we restrict attention to efficient resetting provers and accordingly to arguments [BCC88] (offering computational soundness) rather than proofs (offering statistical soundness).

We next describe our results regarding the construction and implications of the above notion of resettable soundness (further discussion of the model and definition can be found in the technical overview below).

Quantum Black Box Barriers. As intended our definition provides a quantum resetting prover with the power of a quantum black-box zero knowledge simulator. This yields a black box barrier analogous to the one in the classical setting.

Observation 1 (Informal). *Post-quantum resettably-sound* **black-box** *zero knowledge is impossible, except for languages in* **BQP**.

Building on this fact, we then prove that the Goldreich-Krawczyk black box zero knowledge barriers from the classical setting [GK96b] translate to the quantum setting. More generally, we show that under minimal assumptions, any *three-message* or *constant-round public-coin* zero-knowledge protocol can be converted into a quantum resettably-sound argument, while preserving black-box zero knowledge.

Theorem 2 (Informal). *Assuming post-quantum one-way functions, post-quantum zero knowledge protocols that are* **three message or constant-round public-coin**, *with a negligible soundness error, can be made resettably sound. Such protocols cannot be black-box zero knowledge, except for languages in* **BQP**.

We note that the classical barriers proven by [GK96b] do not apply here, as they only consider classical zero-knowledge simulators, rather than the quantum ones in our setting. The transformation behind the above theorem is in fact the same as the corresponding classical transformation [BGGL01]. However, the analysis of the transformation is different and more challenging due to the essential difference between classical resetting and quantum resetting, which is *superposition resetting attacks* (see technical overview).

The resulting black-box barrier holds for general zero knowledge protocols, in particular, for arguments. In the case of *proofs* (with statistical rather than computational soundness), there is evidence that three-message or constant-round public-coin zero knowledge (for non-trivial languages) is impossible altogether (even non-black-box) [BLV06,KRR17,FGJ18]. In the case of black-box zero knowledge, this barrier for proofs was proven (unconditionally) by Jain, Kolla, Midrijanis, and Reichardt [JKMR09]. Finally, we note that like in the classical setting, the resulting barriers, in fact, hold also in a semi-black-box model where the simulator is allowed to depend on the circuit size of the simulated verifier. In the fully black-box model, the barriers can be proven without relying on one way functions.

A Resettably-Sound Protocol via Quantum Non-Black-Box Techniques. Aiming to constructing post-quantum resettably-sound zero knowledge, we are faced with the above mentioned black-box impossibility. In the classical setting, the corresponding black box impossibility of resettably-sound can be circumvented relying on *non-black-box simulation.* Indeed, the pioneering work of Barak shows how to construct constant-round public-coin zero knowledge arguments from collision-resistant hashing [Bar01], to which one can apply the [BGGL01] transformation to obtain resettable soundness. In the quantum setting, however, constant-round public-coin zero knowledge arguments for now remain out of reach.

Nevertheless, under standard assumptions (Quantum Learning with Errors [Reg05] and Quantum Fully-Homomorphic Encryption [Bra18,Mah18]) we construct a post-quantum resettably-sound zero knowledge protocol relying on (quantum) non-black-box simulation.

Theorem 3 (Informal). *Assuming the hardness of* QLWE *and the existence of* QFHE *there exists a post-quantum resettably sound zero-knowledge argument for* **NP**.

Our construction starts from the recent construction of post-quantum constant-round (non-black-box) zero-knowledge [BS20] and modifies it. While non-black-box techniques do not seem inherent for constant round zero knowledge with plain soundness (see [CCY20] in related work), in our setting they become essential. While the non-black-box technique we use is similar to that of [BS20], resettable soundness, requires a new proof, which encounters several technical challenges emerging from quantum resetting.

From Resettable Soundness to Quantumly Unobfuscatable Functions. In the classical setting, resettably-sound zero knowledge is known to be intimately related to the impossibility of virtual black box obfuscation [BGI+12]. In particular, assuming one-way functions any resettably-sound zero knowledge protocol for **NP** implies *a family of unobfuscatable functions* [BP15]. We show that this result translates also to the quantum setting; specifically there exists classical function families that cannot be obfuscated as quantum states according to the quantum virtual black box notion of Alagic and Fefferman [AF16].

Theorem 4 (Informal). *If there exists a post-quantum resettably-sound zero-knowledge argument for* **NP** *and post-quantum one-way functions, then quantum virtual black-box obfuscation is impossible.*

Such an impossibility was recently shown by Ananth and La Placa [AP20b] and by Alagic, Brakerski, Dulek, and Schaffner [ABDS20]. The combination of Theorems 3, 4 yields an alternative, albeit more complicated, proof of this result (under similar assumptions). We note that differently from the classical setting where the impossibility of black box obfuscation is unconditional, in the quantum setting it relies on QLWE and strongly relies on quantum homomorphic encryption. Following the above theorem, any advancement in the construction of quantumly resettably sound protocols, and in particular the construction of constant-round public-coin or three-message protocols, is likely to also advance our understanding of quantum unobfuscatability.

2 Technical Overview

In this section, we provide a technical overview of the paper.

2.1 Defining Post-quantum Resettable Soundness

In the classical setting [BGGL01], a resetting attack by a malicious prover rP is modeled by providing the prover oracle access to the *next-message function* of honest verifier $\mathsf{V}(x, \cdot\ ; r)$ for the common input x and randomness r that is sampled uniformly and fixed once and for all. The prover then has the ability to query a partial transcript ts, including prover messages up to some round i, and obtain back the verifier message in round $i+1$. In a successful attack, after polynomially many queries, the prover manages to output a full transcript ts for some false statement x, which yet convinces the verifier $\mathsf{V}(x, \mathsf{ts}; r)$.

Aiming to generalize this to the quantum setting, there are two conceivable definitions. The first considers quantum provers, which are only given *classical access* to $\mathsf{V}(x, \cdot\ ; r)$. The second, which we consider in this work, provides the prover with *quantum access* to $\mathsf{V}(x, \cdot\ ; r)$; namely, access to the unitary map $|\mathsf{ts}\rangle|y\rangle \mapsto |\mathsf{ts}\rangle|y \oplus \mathsf{V}(x, \mathsf{ts}; r)\rangle$; in particular, it may now query $\mathsf{V}(x, \cdot\ ; r)$ in superposition. While the first may still provide meaningful security in settings where classical access can be enforced, the second resists stronger resetting scenarios in which the attacker can perform quantum resetting and remain secure even in

settings where classical access could be hard to enforce (similar considerations arise when considering CCA and signatures against quantum adversaries, see for instance [BZ13]). Finally, our definition captures the abilities of a black-box zero-knowledge simulator, and will thus be useful for proving black-box barriers on post-quantum zero knowledge.

Proving that resettably-sound protocols cannot be *black box* zero knowledge, except for languages in **BQP**, now follows a standard argument similar to the classical one [BGGL01]. Roughly, speaking this is because a quantum resetting prover has the ability to run a quantum black-box simulator for the verifier $\mathsf{V}(x, \cdot\ ; r)$, in order to produce a cheating transcript. Indeed, by zero knowledge and completeness, for any true statement x, the simulator almost always generates an accepting transcript, and unless it can decide the underlying language (meaning that it is in **BQP**), it must also be able to do so for some false statements.

Variants. A natural strengthening of the above definition allows the prover to also choose the statements x that it provides the oracle with; namely get access to $\mathsf{V}(\cdot\ , \cdot\ ; r)$. In the body, we prove that this stronger notion can be obtained from the simpler notion assuming subexponentially-secure (post-quantum) pseudorandom functions. We note that all the implications of resettable soundness shown in this work, already follow from the simpler notion of resettable soundness.

Also, as already noted we restrict attention to efficient resetting provers, namely arguments. We note that classically, resettably-sound zero knowledge proofs, namely against unbounded provers, are only possible for trivial languages [BGGL01], and this carries over to the quantum setting. Again, all implications shown in this work already follow from resettably-sound zero knowledge arguments.

2.2 3-Message and Constant-Round-Public-Coin Protocols Can Be Made Resettably Sound

We now explain how 3-message protocols and constant-round public-coin protocols are made resettably sound. The transformation does not change the honest prover, and thus preserves black box zero knowledge, and any other privacy guarantee, such as witness indistinguishability (which we will use later on). This in turn yields quantum black-box zero-knowledge barriers on 3-message or constant-round public-coin protocols (with a negligible soundness error).

3-Message Protocols. The transformation for three-message protocols is essentially identical to the classical one [BGGL01]. Given the original verifier V for the protocol, we consider a new verifier $\tilde{\mathsf{V}}$ whose randomness consists of a random seed k for a pseudorandom function secure under quantum access [Zha12]. Given a statement x and first prover message α, the verifier $\tilde{\mathsf{V}}$ derives

randomness r by applying the PRF and derives the second message β, by applying the original verifier with corresponding randomness:

$$r = \mathsf{PRF}_k(\alpha), \qquad \beta = \mathsf{V}(x, \alpha; r) \ .$$

As expected $\tilde{\mathsf{V}}(x, \alpha, \beta, \gamma; k)$ accepts if the original verifier $\mathsf{V}(x, \alpha, \beta, \gamma; r)$ accepts.

In the classical setting, resettable soundness is proven by a relatively simple reduction to the soundness of the original protocol. In the quantum setting, however, proving security is significantly more challenging. Before we address these challenges let us start by recalling the classical reduction to develop basic intuition. We are given a resetting prover rP, which without loss of generality, never makes the same query twice, and always queries the oracle $\tilde{\mathsf{V}}$ on the cheating transcript it eventually outputs. Roughly speaking, the reduction, which aims to cheat V in a single interaction, will aim to embed this interaction in a random position in an execution of the resetting $\mathsf{rP}^{\tilde{\mathsf{V}}(x,\cdot\,;k)}$ and forward that execution to the external verifier V. All other executions are internally simulated by the reduction. By pseudorandomness, the view of the simulated rP is indistinguishable from its view in a resetting attack and will include some cheating execution. With noticeable probability (inverse proportional to the number of queries that rP makes), the reduction hits the cheating execution and wins.

In the quantum setting, however, it is not a-priori clear how such a reduction would work. In particular, any query made by rP to $\tilde{\mathsf{V}}$ may now include a superposition of super-polynomially many transcripts. Furthermore, merely observing the prover queries disrupts its state and could affect the probability it produces a cheating transcript. Embedding an execution at a random position is also tricky. When we forward some message α to the external verifier, and obtain back a message β, we have to answer consistently with β all oracle queries to α. However, whereas in the classical case, we could assume that no α is queried more than once (because queries can be stored), now it may be that α takes part in all superposition queries that the prover makes.

Similar difficulties arise when trying to prove the soundness of the Fiat-Shamir transformation [FS86] in the quantum random oracle model [BDF+10], and were, in fact, successfully circumvented in recent works [LZ19,DFMS19, DFM20]. Indeed, both in the Fiat-Shamir setting and in our setting, we can still hope to obtain an analog of the classical reduction. Specifically, by measuring a random query made by rP, forwarding the result α to the external verifier, and consistently answering with β any *future query* α by *reprogramming* the classical function $\tilde{\mathsf{V}}$.

The intuition is that for the prover to succeed in outputting a convincing transcript (α, β, γ), the message α has to appear in one of his superposition queries with noticeable weight; otherwise, it gains almost no information on the corresponding verifier message β, and will fail to break soundness. Furthermore, when measuring such a query we are likely to obtain α, without disturbing the prover's state too much (in the extreme case that α occurs with probability one, the state is not disturbed at all). If the reduction hits the first such query (where

α is significant), then it suffices that it is consistent with α in future queries and does not have to worry about past queries.

This intuition is elegantly captured and made rigorous by Don, Fehr, Majenz, and Schaffner [DFMS19,DFM20]. They prove reprogramming and simulation lemmas that establish the validity of (a slight variant of) the described reduction in the case of Fiat Shamir, where the message β is chosen uniformly at random. In our setting, β is an arbitrary message derived by the verifier. Nevertheless, relying on their reprogramming lemma, we can prove an appropriate simulation lemma for our setting.

A Useful Generalization: Many-Round Almost Resettable Protocols. We also show a generalization of the three-message transformation that allows to take any *single-prefix resettably-sound protocol* and make it (fully) resettably sound. Single-prefix resettably sound protocols are almost resettably sound. They allow the resetting prover to use a *single classical first message* and accordingly obtain a single response to this message from the verifier. Only starting from the prover's next message it is allowed to quantumly reset; namely all interactions (even if in superposition) start with the same classical prover message and verifier response. A three message protocol is indeed the simplest example of a single-prefix resettably-sound protocol, since the verifier has a single message, and if this message is not reset, then there is no resetting whatsoever, and resettable soundness is synonymous to plain soundness.

This generalization turns out to be useful, and is used later on in our construction of a resettably sound (non-black-box) zero knowledge protocol for **NP**. To obtain this generalization, we first extend the reprogramming lemma from [DFM20] to the case of reprogramming an entire oracle, specified by some prefix. This allows us to extend the previously described reduction, which given a fully resetting prover can turn it into a single prefix resetting prover. The difference is that now rather than obtaining from the external verifier a response β to the measured α, it obtains oracle access to an oracle $\tilde{\mathsf{V}}(x, \alpha, \cdot\,; r)$ specified by the prefix α (and implicitly a response β). This oracle effectively allows to perform resetting attacks, but only starting from the next prover message.

Constant-Round Public-Coin Protocols. Another example where classical resettable soundness can be achieved is that of constant round public-coin protocols. Also here we obtain an analogous transformation in the quantum setting, now based on *multi-value reprogramming lemmas* from [DFM20], used there to deal with multi-message Fiat Shamir.

Beyond 3-Message or Constant-Round Public-Coin? We note that we should not hope to transform arbitrary protocols into resettably-sound ones; indeed, multi-message post-quantum zero knowledge protocols for **NP** do exist, and are even public coin [Wat09]. But what does it take for a protocol to be (transformable to) resettably sound? Here one bottleneck is the (in)ability of

the reduction to simulate internally the interactions that are not forwarded to the external verifier. More specifically, the question is whether the reduction could simulate *continuations* that start consistently with the external verifier and then diverge. In general private-coin protocols, this may not be possible as the private coins of the external verifier are not known to the reduction. In contrast, in three-message protocols this is not a problem, as there is nothing to continue (the verifier has a single message). Similarly, also in public coin protocols, simulating continuations is easy—the reduction samples the random messages on its own.

This is, however, not the only bottleneck. A second bottleneck is that the reduction has to *hit the cheating execution* with noticeable probability, and since the reduction has to guess on the fly which messages to forward to the external verifier, this probability may decrease exponentially in the number of rounds. Hence, even for public coin protocols, the transformation only works for a constant number of rounds. In fact, this is tight—the round complexity of Watrous' zero knowledge public-coin proofs [Wat09] can be reduced to any super constant function $\omega(1)$. (For instance, by starting from Blum's Hamiltonicity protocol [Blu86] that has constant soundness, repeating it in parallel logarithmically many times, and then sequentially $\omega(1)$ times.)

2.3 Constructing a Resettably Sound Non-Black-Box Zero-Knowledge Protocol

We now outline the main ideas and techniques behind our construction of a resettably-sound non-black-box zero-knowledge protocol for **NP**. Our starting point is the post-quantum zero knowledge protocol of Bitansky and Shmueli [BS20]. We next describe the main challenges in turning this protocol into a quantumly resettably sound protocol.

A Bird's Eye View of the BS Protocol. At a high level (and oversimplifying), the BS protocol consists of two phases. First, the verifier provides a quantum extractable commitment to a challenge message. Then the parties execute a standard zero knowledge sigma protocol to prove the statement x, where the verifier opens the commitment from the first phase. The extractor for the first-phase commitment is non-black-box, using the code of a sender (the verifier in this case), it can extract the underlying message while faithfully simulating the quantum state of the sender. This gives rise to a corresponding non-black-box simulation strategy, which first extracts the verifier challenge and can then cheat in the sigma protocol.

Already at this level, one can see that the protocol is not resettably sound, even classically, let alone quantumly. A resetting prover can first run the verifier until the opening phase, obtain the challenge, then reset the verifier, and like the simulator use the obtained challenge to cheat in the sigma protocol. Indeed, the reason that the actual simulator in the BS protocol does not follow this black-box strategy is that it does not work for malicious quantum verifiers, whereas a resetting prover only has to cheat a classical verifier.

Following the above observation, we change the above high level blueprint. We rely on the Feige-Lapidot-Shamir [FLS99] *trapdoor paradigm.* In the first-phase, the BS extractable commitment is used to set up a trapdoor statement t. In the second phase, the prover provides a witness-indistinguishable proof that either x is a true statement or t is a true statement. To guarantee soundness, the trapdoor statement is set up so that it is indistinguishable from a false statement, and thus relying on the soundness of the second-phase proof, a convincing proof must mean that x is a true statement. In contrast, a simulator given the code of the verifier should be able to efficiently extract a witness for the trapdoor statement t, and can then use it in the second phase proof indistinguishably from the prover (who uses the witness for x).

Given that we are interested in quantum resettable soundness, we have to guarantee that the indistinguishability of the trapdoor statement t from a false statement, holds even against quantum resetting attacks. Furthermore, we have to guarantee that the second-phase proof is resettably sound. For the latter, we can use standard constant-round public-coin witness-indistinguishable proofs; indeed, we have already shown that such proofs can be made quantumly-resettably sound, while preserving witness indistinguishability. The more involved part is establishing indistinguishability of the trapdoor statement from a false one under resetting.

A Resettably-Secure Trapdoor Phase. We now dive deeper into the construction of a resettably-secure trapdoor phase. In terms of extractability (of a trapdoor witness), we first present a trapdoor phase that is only extractable against a restricted class of verifiers that are *non-aborting and explainable.* The notion of non-aborting explainable verifiers considers verifiers whose messages can always be *explained* as a behavior of the honest (classical) verifier with respect to *some* randomness (finding this explanation may be inefficient); in particular, they never abort. This simpler setting will already capture the main challenges we need to deal with. We will later discuss how this restriction is removed.

Similarly to the BS extractable commitment, we rely on three basic tools:

- *Quantum fully-homomorphic encryption* (QFHE)—an encryption scheme that allows to homomorphically apply any polynomial-size quantum circuit C to an encryption of x to obtain a new encryption of $C(x)$, proportional in size to the result $|C(x)|$ (the size requirement is known as *compactness*).
- *Compute-and-compare program obfuscation* (CCO). A compute-and-compare program $\mathbf{CC}[f, v, z]$ is given by a function f (represented as a classical circuit) and a target string v in its range; it accepts every input x such that $f(x) = v$, and rejects all other inputs. A corresponding obfuscator compiles any such program into a program $\widetilde{\mathbf{CC}}$ with the same functionality. In terms of security, provided that the target v has high entropy conditioned on f, the obfuscated program is computationally indistinguishable from a simulated dummy program that is independent of f, v, z, and rejects *all* inputs.

– *Secure function evaluation* (SFE) that can be thought of as homomorphic encryption with an additional *circuit privacy* guarantee, which says that the result of homomorphic evaluation of a circuit, reveals nothing about the evaluated circuit to the decryptor, except of course from the result of evaluation.

We now describe a (still simplified) trapdoor phase, which is essentially the same as the BS extractable commitment, except for how the randomness of the verifier is handled. In the trapdoor phase the verifier has two randomized steps; we denote the randomness used in these rounds by r_1 and r_2, respectively.

1. The prover P samples a secret key sk for SFE, and sends a commitment cmt to sk.
2. The verifier V uses randomness r_1 to sample:
   – two random strings u and v,
   – a secret key sk' for an FHE scheme,
   – an FHE encryption $\mathsf{ct}'_u = \mathsf{QFHE.Enc}_{\mathsf{sk}'}(u)$ of u,
   – an obfuscation $\widetilde{\mathbf{CC}}$ of $\mathbf{CC}[f, v, \mathsf{sk}']$, where $f = \mathsf{QFHE.Dec}_{\mathsf{sk}'}$ is the FHE decryption circuit.

   It then sends $(\mathsf{ct}'_u, \widetilde{\mathbf{CC}})$ to the prover P.
3. The prover P:
   – sends $\mathsf{ct}_{u'}$, a string u' encrypted using SFE (the honest prover sets u' arbitrarily).
   – proves using a resettably-sound witness-indistinguishable argument that $\mathsf{ct}_{u'}$ is a valid SFE encryption corresponding to the secret key sk underlying the commitment cmt.
4. The verifier V:
   – uses the SFE homomorphic evaluation to compute the function $C_{u \to v}$ that given input u, returns v (and otherwise $\perp$).
   – To derive the randomness for this evaluation, V interprets its randomness r_2 as a seed for a pseudorandom function and applies it to the prover messages $(\mathsf{cmt}, \mathsf{ct}_{u'})$.
   – V then returns the resulting ciphertext to P.
5. The trapdoor statement t is set to be:

"There exists a ciphertext ct^* *that the program* $\widetilde{\mathbf{CC}}$ *does not reject."*

Basic Intuition. We start by building basic intuition on how the above protocol achieves the goal of a trapdoor phase. For starters we will ignore the resetting attacks, and recall the intuition from BS. Then we will address the main challenges in proving resettable security, and how they are met. (A reader familiar with BS may want to skip directly to the resettable security paragraph.)

Let us start by explaining how a non-black-box simulator can use the circuit of an explainable verifier in order to obtain a witness proving the trapdoor statement. The simulator acts honestly in the first step, and then obtains the CC obfuscation $\widetilde{\mathbf{CC}}$ and FHE encryption ct'_u of the string u. The main point is that now the simulator can *homomorphically continue the protocol under the FHE*

encryption. That is, it will evaluate the (quantum) verifier under the encryption, where it has the secret u *in the clear* and can use it in the SFE protocol to obtain back the secret target value v (the hiding of SFE encryption is used to argue that such an execution is indistinguishable from a real one where a dummy encryption is sent). Going back out of the encryption, the simulator now actually holds an encryption ct^* of v, and in particular $\widetilde{\mathbf{CC}}$ does not reject ct^*, but rather outputs the FHE secret key sk'. Thus, the ciphertext ct^* obtained by the simulator is a valid trapdoor witness. The reason we require $\widetilde{\mathbf{CC}}$ to output sk', rather than an arbitrary accept value, is for the simulator to be able to decrypt the internal verifier quantum state and faithfully continue the simulation.

We now turn to explain why to a malicious (but for now, non-resetting) prover, who does not obtain the code of the verifier, the trapdoor statement is indistinguishable from a false statement. Specifically, we would like to argue that we can replace the obfuscation $\widetilde{\mathbf{CC}}$ with a simulated one that rejects all inputs. To see this, we first argue that the prover cannot send an SFE encryption $\mathsf{ct}_{u'}$ such that $u' = u$, except with negligible probability. Indeed, given only the first sender message $(\mathsf{ct}'_u, \widetilde{\mathbf{CC}})$, the receiver obtains no information about u. Hence, we can invoke the CCO security and replace the obfuscation $\widetilde{\mathbf{CC}}$ with a simulated one, which is independent of the secret FHE key sk. This, in turn, allows us to invoke the security of encryption to argue that the first message $(\mathsf{ct}'_u, \widetilde{\mathbf{CC}})$ hides u. While this means that the prover does not obtain u in the clear, we still need to argue that it cannot send an encryption of u. This is done using a non-uniform reduction and is exactly the purpose of the prover commitment cmt to the SFE secret key sk, which allows us to provide the reduction with sk as non-uniform advice. Having established that no SFE encryption of u is sent we can invoke the circuit privacy guarantee to completely remove the value v from the prover's view and now we also replace $\widetilde{\mathbf{CC}}$ with a simulated one that rejects all inputs.

Resettable Security. The above argument establishing indistinguishability of the trapdoor statement from a false statement, does not consider resettable attackers. We now discuss the difficulties arising from resetting attacks and how they are dealt with.

Recall that a resetting quantum attacker may perform superposition queries. Accordingly, now when arguing that it cannot produce an SFE encryption of u, we would like to argue that SFE encryptions of u have negligible weight in any query made by rP; in other words, projecting the queries on the space of non-u queries has little effect on the experiment. Indeed, we can prove this if the resetting prover is guaranteed to always use the same SFE encryption key, in which case we can non-uniformly hardwire this key into our reduction like before. The problem is that a resetting prover may start many executions, each with a different SFE key; in fact it can run exponentially many such executions in superposition. This is where we use our reduction to *single-prefix resetting provers* (discussed in the previous section). The reduction allows us to obtain new prover that in all executions sends the same commitment cmt and uses

the same secret key; any resetting attempt is done from the next message and onward.

Having established that the prover queries do not include encryptions of the secret u (or rather have a small projection on this space), we would like to invoke as before the circuit privacy guarantee. However, this should be done with care. The problem is the prover still has the ability to send many ciphertexts and receive evaluations on each one of them. This is the reason we invoke a pseudorandom function to derive randomness in this step, which ensures that each evaluation uses (pseudo)independent randomness. Proving security, however, is not straightforward. In the classical setting, this is not an issue—the overall number of queries is polynomial and thus we can use a standard hybrid argument, invoking circuit privacy polynomially many times. In the quantum setting, however, where queries include a superposition over exponentially many ciphertexts, this is unclear. In fact, there is a basic problem here, which we find interesting on its own. Assume that for two efficient samplers $S_0(x)$ is computationally indistinguishable from $S_1(x)$ for any input x; are the two oracles $F_i(x) := S_i(x; R(x))$ indistinguishable (quantumly), when R is a random function? Zhandry [Zha12] shows that this is the case if $S_i(x) = S_i(y)$ for any x, y, but the general case is unclear.

Fortunately, in our case, we can take a straightforward approach to solve it, by guaranteeing that circuit privacy is statistical, and ensuring that the statistical error is smaller than the total number of ciphertexts in the support, and thus a naive hybrid argument still works. Doing so again requires care, as the size of SFE ciphertexts and the statistical security guaranteed may be related. We show how to deal with this by forcing the prover to also commit to the randomness used in SFE encryptions so that the number of hybrids only depends (exponentially) on the fixed length of the encrypted plaintext.

General Verifiers. In the described trapdoor protocol, we have made two simplifying assumptions regarding the verifier—that it is explainable and that it is non-aborting. We deal with the first restriction using a common approach based on witness indistinguishable proofs by the verifier [BKP19,BS20]. This time however, we need to rely on *resettable* statistical witness indistinguishability. Statistically-witness-indistinguishable ZAPs are known under super-polynomial hardness of QLWE [GJJM20,BFJ+20] and are resettable as they only include one round. We also give a solution using only polynomial hardness of QLWE, based on Unruh's notion of collapse binding statistically-hiding hash functions, which leads to statistical witness-indistinguishable protocols [Unr16b,Unr16a], while these protocols are not resettably-witness-indistinguishable as is, we show how to make them resettably secure.

As for dealing with verifier aborts, we rely on a general approach from [BS20], which roughly asserts that it is sufficient to be able to construct two separate zero knowledge simulators, one for verifiers that do not abort and one for verifiers that do, and which do not affect the probability of aborting (more than negligibly).

They show that two such simulators can always be combined to one full-fledged simulator using Watrous' rewinding lemma [Wat09].

2.4 From Resettable Soundness to Quantum Unobfuscatability

Finally, we outline the construction of quantumly unobfuscatable functions from resettably-sound zero-knowledge protocols for **NP** and one-way functions. Informally, an unobfuscatable function family is a family of classical functions $\{f_k\}$ indexed by a secret k. Given quantum oracle access to a random f_k in the family, no efficient quantum learner should be able to learn some secret function $s(k)$ of the key. In contrast, given any quantum state ρ and quantum circuit C such that for some k and and all inputs x, $C(\rho, x)$ computes the classical value $f_k(x)$, one could efficiently extract from C and ρ the corresponding secret $s(k)$.

Our construction closely follows the construction of classically unobfuscatable functions from classical resettably sound zero knowledge protocols [BP15], while making some adaptations to the analysis stemming from the difference between the classical and quantum settings. Roughly speaking, our family of functions $\{f_{r,\varphi,s}\}$ is indexed by randomness r and statement φ for the (honest) verifier given by our resettably-sound protocol, and some secret s. The statement φ is taken from some **NP** language $\mathcal{L}$ where random statements $\varphi \in \mathcal{L}$ are indistinguishable from statement not in $\mathcal{L}$ (for instance pseudorandom strings vs random strings for a sufficiently stretching pseudorandom generator). The function generally computes the verifier next message function $\mathsf{V}(\varphi, \cdot; r)$ with two exceptions. For some fixed public input `statement`, the function will output the statement φ. Also, given any accepting transcript ts, the function outputs its secret s.

To argue unlearnability, we show that any efficient quantum learner L that given oracle access to a random $f_{r,\varphi,s}$ finds s can be transformed into a prover that violates quantum resettable soundness. For this, we first show that any learner that manages find s with noticeable probability, can be translated into a learner that that given access to $\mathsf{V}(\varphi, \cdot; r)$ finds an accepting transcript ts, still with noticeable probability. For this we rely on a *quantum one-way to hiding lemma* by Ambainis, Hamburg, and Unruh [AHU19]. We then rely on the fact that φ is indistinguishable from a false statement to deduce that the prover will also succeed for no statements and thus break resettable soundness.

Finally, we show that we can use the non-black-box zero knowledge simulator to extract an accepting transcript with overwhelming probability. Given a quantum circuit C and state ρ implementing the function $f_{r,\varphi,s}$, say perfectly (although almost perfectly would still do). We can realize a quantum circuit along with quantum auxiliary input ρ that implement the verifier $\mathsf{V}(\varphi, \cdot; r)$. Here perfect correctness guarantees that when the constructed verifier computes its next messages, the state ρ is not disturbed, and thus we can repeatedly compute next messages. We can now run our non-black-box simulator (which also works relative to quantum auxiliary input), and by zero knowledge and completeness obtain an accepting transcript.

2.5 Related Work

We now mention additional related work, elaborate on some of the related works mentioned earlier, and address concurrent work.

Classical Resettable Security. The notion of resetting attacks was first considered by Canetti, Goldreich, Goldwasser, and Micali [CGGM00]. They defined and constructed protocols that are zero knowledge against resetting attacks. Resettable soundness was then introduced and achieved by Barak, Goldreich, Goldwasser, and Lindell [BGGL01]. Deng, Sahai, and Goyal showed how to construct a simultaneously resettable zero knowledge protocol [DGS09], this result was later followed by Goyal [Goy13] who gave a public coin protocol, by Chung, Ostrovsky, Pass and Visconti [COP+14] who gave a protocol based on one-way functions, and by Chongchitmate, Ostrovsky, and Visconti [COV17] who gave a constant round protocol, based on various standard assumptions. Goyal and Sahai [GS09] and Goyal and Maji [GM11] defined and constructed varioues forms of resettable secure computation. Bitansky and Paneth [BP12,BP13,BP15] constructed resettably-sound protocols with various improved features based on unobfuscatability. Chung, Pass, and Seth [CPS13] constructed resettably-sound zero knowledge based on one-way functions. Finally, Chung, Ostrovsky, Pass, and Venkitasubramaniam [COP+14] presented a 4-round resettably sound zero-knowledge based on one-way functions.

Post-Quantum Zero-Knowledge for NP. The study of post-quantum zero-knowledge (QZK) protocols was initiated by Van De Graaf [VDGC97], who first observed that traditional zero-knowledge simulation techniques, based on rewinding, fail against quantum verifiers. Subsequent work has further explored different flavors of zero knowledge and their limitations [Wat02], and also demonstrated that relaxed notions such as zero-knowledge with a trusted common reference string can be achieved [Kob03,DFS04]. Watrous [Wat09] was the first to show that the barriers of quantum information theory can be crossed, demonstrating a post-quantum zero-knowledge protocol for **NP** (in a polynomial number of rounds). A constant round non-black-box zero knowledge protocol was constructed by Bitansky and Shmueli [BS20] based on QLWE and quantum fully homomorphic encryption. Similar techniques for non black-box extraction were also developed by [AP20a]. Subsequently, Agarwal, Bartusek, Goyal, Khurana, and Malavolta [ABG+20] extended the BS construction to obtain parallel zero knowledge based on spooky encryptions for relations computable by quantum circuits.

Very recently Chia, Chung and Yamakawa [CCY20] showed that the Goldreich-Kahan protocol [GK96a] satisfies a relaxed notion called (post-quantum) ε-zero knowledge; the protocol is based on collapse binding hash functions in the case of proofs, and on one-way functions in the case of arguments.

Barriers for 3-Message and Constant-Round Public-Coin Proofs. Classically, 3-message and constant-round public-coin zero knowledge arguments are

subject to black-box barriers [GK96b], but can in fact be classically achieved using non-black-box simulation (under appropriate computational assumptions) [Bar01,BKP18]. In the case of proofs, there is evidence that they are unlikely to exists altogether (including non-black-box zero knowledge). Specifically, constant-round public-coin proofs do not exist assuming appropriate Fiat-Shamir hash functions [FS86,DNRS03,BLV06]. Kalai and the Rothblums [KRR17] gave such an instantiation of a Fiat Shamir hash assuming subepxoenential indistinguishability obfuscation, and strong forms of point obfuscation. Jain, Fleischhacker, and Goyal [FGJ18] extended their impossibility to also rule out three-message proofs. The mentioned implications also hold in the quantum setting, assuming post-quantum analogs of the corresponding assumptions. Jain, Kolla, Midrijanis and Reichardt [JKMR09] showed that for black-box zero knowledge, proofs can be ruled out unconditionally.

Simulating Quantum Oracles. Quantum oracles have been a fundamental aspect of quantum computation from the start. Querying the oracle in superposition created the need to develop new proof techniques. Specifically when proving security of quantum protocols in the Quantum Random Oracle Model ([BDF+10]). The main issue is the lack of ability to record the queries asked by the adversaries and to easily reprogram the answers. Nevertheless, many results were achieved even without these abilities [Zha12,Unr14,Zha15,ES15,Unr15, TU16,ABB+17,KLS18]. Following Zhandry's work [Zha18] on recording random oracles, many other results were proven such as the Fiat-Shamir transform [LZ19,DFMS19,DFM20], the Micali CS Proofs [CMS19], 4-round Luby-Rackoff construction [HI19] and more.

Quantum Obfuscation. Quantum obfuscation was first proposed by [AF16]. It's impossibility is not implied by the impossibility proved in [BGI+12]. In recent work, [ABDS20] showed the impossibility of such schemes based on the hardness of QLWE. A related stronger notion called Secure Software Leasing was dealt in [AP20b] and [KNY20], showing the impossibility of such generic scheme (based on QLWE and the existence of QFHE), and the possibility of such schemes for restricted classes of functions (pseudo-random functions and evasive functions) under sub-exponential QLWE.

Concurrent Work. In a concurrent and independent work, Chia, Chung, Liu and Yamakawa [CCLY21], prove new black-box barriers on post-quantum zero knowledge. They show that black-box ε-zero-knowledge is impossible for three-message and constant-round public-coin protocols, and that black-box zero knowledge is impossible for general constant round protocols (also private coin). The barriers on ε-zero-knowledge for public-coin and three-message also follow directly from our resettable-soundness transformations, but the barrier for general constant-round protocols does not. The other results in this paper (the construction of a resettably-sound protocol and the connection to unobfuscatability) do not overlap with their work.

Technically, while Chia et al. do not explicitly consider resettable soundness, the barriers on three-message and public-coin protocols are proven similarly (using measure-and-reprogram techniques). To achieve the result on general constant-round, they first extend a classical result by Barak and Lindell [BL04] on the impossibility of a *strict* polynomial-time black-box simulator. This is again done using similar measure-and-reprogram techniques. Then, they further extend the result to expected-time simulators. This requires novel ideas and strongly relies on quantum entanglement; in particular, in the classical setting, such a barrier does not exist.

3 Defining Post-Quantum Resettable Soundness

In this section, we present our definition of resettable soundness, and show and immediate implication of this definition, regarding the triviality of black-box zero-knowledge arguments with resettable soundness.

3.1 Post-Quantum Resettable Soundness

We present our definition for post-quantum resettable soundness. Our definition deals with giving oracle access to fixed verifier. We shall use $\mathsf{V}(x,\cdot;r)$ to denote the interaction of algorithm V on instance x fixed randomness r (where the input is a partial transcript). Also, to denote the application of V's predicate on a transcript ts we shall write $\mathsf{V}(x,\mathsf{ts};r)$. The definition of resettable soundness is as follows,

Definition 1 (Post-Quantum Resettable Soundness). *A classical interactive protocol $\langle \mathsf{P},\mathsf{V}\rangle$ for language $\mathcal{L}$ has resettable soundness against quantum provers, if for any malicious* QPT *resetting prover $\mathsf{rP}=\{\mathsf{rP}_\lambda,|\psi_\lambda\rangle\}_{\lambda\in\mathbb{N}}$ there exists a negligible function $\mu(\cdot)$ such for any security parameter $\lambda\in\mathbb{N}$ and any $x\in\{0,1\}^\lambda\setminus\mathcal{L}$ it holds that,*

$$\Pr_r\left[\mathsf{V}(x,\mathsf{ts};r)=1 \;\middle|\; \mathsf{ts}\leftarrow \mathsf{rP}_\lambda^{\mathsf{V}(x,\cdot;r)}(|\psi_\lambda\rangle)\right]\leq \mathrm{negl}(\lambda) \quad ,$$

where ts is a transcript of a possible interaction between P,V. $\mathsf{V}(x,\cdot;r)$ is the function that computes V's next message, on instance x and some fixed randomness r, given as input a transcript of a partial interaction.

4 Transforming Protocols to Achieve Quantum Resettable Soundness

In this section we show that classical three-message protocols as well as constant-round public-coin protocols can be made resettably sound assuming one-way functions. The transformation is simple and similar to the one from the classical setting [BGGL01], however, having to deal with quantum resetting attacks, the analysis is significantly different. The transformation preserves black-box zero-knowledge; accordingly, we deduce as a corollary that post-quantum black-box zero-knowledge protocols cannot be 3-message or constant-round public-coin, except for trivial languages.

4.1 Quantum Oracle Notations

We rely on a couple of lemmas proved in [DFM20]. We restate them here again, while augmenting some of the notation, to fit with our conventions. Let A^H be a quantum oracle-aided algorithm. For a q-query algorithm, without loss of generality, A can be described as having the following registers, query registers on which we apply the unitary $\mathcal{O}_H$ computing $|x\rangle|y\rangle \rightarrow |x\rangle|y \oplus H(x)\rangle$, X, Z which are output registers, and E holds any other internal qubits used by A. More so, the operation of A on its initial state can be described as,

$$\mathsf{A}^H = \mathsf{A}_q \mathcal{O}_H \dots \mathsf{A}_1 \mathcal{O}_H \ ,$$

where A_i is a sequence of unitaries. Like [DFM20] we use the following notation for $i < j \in [q]$

$$\mathsf{A}^H_{i \rightarrow j} = \mathsf{A}_j \mathcal{O}_H \dots \mathsf{A}_{i+1} \mathcal{O}_H \ .$$

We also denote $\mathsf{A}^H_{i \rightarrow j} = \mathsf{Id}$ for $i \geq j \in [q]$. Assuming A gets as initial input a pure state $|\phi_0\rangle$, we denote,

$$|\phi_i^H\rangle = \mathsf{A}^H_{0 \rightarrow i} |\phi_0\rangle \ .$$

For a function H we denote by $H_{x \rightarrow \theta}$ the same function where x is remapped to θ:

$$H_{x \rightarrow \theta}(x') = \begin{cases} H(x') & x' \neq x \\ \theta & x' = x \end{cases} \ .$$

4.2 Transforming 3 Message Private Coin Protocols

We show that any 3 message interactive protocol $\langle \mathsf{P}, \mathsf{V} \rangle$ can be transformed to a quantum resettably sound one, assuming the existence of quantum secure PRFs. More formally we show the following,

Proposition 1 (Compiler For 3 Message Protocols). *Assuming quantum-secure one-way functions, any 3 message protocol $\langle \mathsf{P}, \mathsf{V} \rangle$ with negligible soundness for a language $\mathcal{L}$, can be transformed to into a post-quantum resettably sound protocol $\langle \mathsf{P}, \tilde{\mathsf{V}} \rangle$. More so, if $\langle \mathsf{P}, \mathsf{V} \rangle$ is (black-box) zero-knowledge then so is $\langle \mathsf{P}, \tilde{\mathsf{V}} \rangle$.*

Combining proposition 1 with observation 1 immediately implies the following corollary,

Corollary 1. *If $\mathcal{L}$ has a 3 message post-quantum black-box zero-knowledge protocol, then $\mathcal{L} \in$ **BQP**.*

Single Value Reprogramming. To prove our construction presented in 4.2, we shall rely on a lemma by [DFM20].

Lemma 1 (Single Value Reprogramming Lemma ([DFM20])). *Let* A *be a* q*-query oracle quantum algorithm. Then, for any function* $H : \mathcal{X} \to \mathcal{Y}$*, any* $x \in \mathcal{X}$ *and* $\theta \in \mathcal{Y}$*, and any projection* $\Pi_{x,\theta}$ *acting on the* Z *register (which may depend on* x, θ*), it holds that*

$$\mathop{\mathbb{E}}_{i,b}\left[\left\|(|x\rangle\langle x| \otimes \Pi_{x,\theta})\left(\mathsf{A}^{H_{x\to\theta}}_{i+b\to q}\right)\left(\mathsf{A}^{H}_{i\to i+b}\right)(|x\rangle\langle x|)\,|\phi^{H}_{i}\rangle\right\|_2^2\right] \geq \frac{\left\|(|x\rangle\langle x| \otimes \Pi_{x,\theta})\,|\phi^{H_{x\to\theta}}_{q}\rangle\right\|_2^2}{(2q+1)^2} \quad ,$$

where the expectation is over uniform $(i, b) \in \{0, \ldots, q-1\} \times \{0, 1\} \cup \{(q, 0)\}$*. We emphasize that first* $|x\rangle\langle x|$ *acts on query register, while the second acts on the* X *register.*

Remark 1. We state here the technical lemma and not the existence of a simulator, as done in the multiple values reprogramming in the public-coin case, since unlike [DFM20] we use this lemma to reprogram a non-uniform output function, in our private-coin transform.

Construction. Fix some language $\mathcal{L}$ with a three-message protocol $\langle \mathsf{P}, \mathsf{V}\rangle$ whose message we denote by (α, β, γ). Assume V uses $m(\lambda)$ bits of randomness. We present the protocol $\langle \mathsf{P}, \tilde{\mathsf{V}}\rangle$. P is exactly the same, where as $\tilde{\mathsf{V}}$ is described in 1.

Algorithm 1: $\tilde{\mathsf{V}}(x; k)$

1 Use k as a key for $\mathsf{PRF}_k(\cdot)$, a pseudo-random function.
2 Given α compute $\beta = \mathsf{V}(x, \alpha; \mathsf{PRF}_k(\alpha))$.
3 Given a transcript α, β, γ compute $\mathsf{V}(x, (\alpha, \beta, \gamma); \mathsf{PRF}_k(\alpha))$ and output it.

The fact that the protocol preserves completeness and zero-knowledge follows readily, we focus on proving resettable soundness. To show resettable soundness, we show an efficient reduction from a resetting prover rP to a prover $\tilde{\mathsf{P}}$ for the original protocol, which preserves the cheating probability up to a polynomial loss.

Fix a malicious quantum resetting prover rP for a false instance x. Assume that rP makes at most q oracle queries, and has non-uniform advice $|\psi_0\rangle$. Assume rP has registers A, Z, E and query registers. The query registers are for querying a first message α and receiving the corresponding second message β. A, Z will hold the outputted first and third message, and E holds any internal qubits used. Then, $\tilde{\mathsf{P}}$ will perform as follows,

We show that,

Claim.

$$\Pr\left[\langle \tilde{\mathsf{P}}, \mathsf{V}\rangle(x) = 1\right] \geq \frac{1}{(2q+1)^2}\Pr_k\left[\langle \mathsf{rP}, \tilde{\mathsf{V}}(x, \cdot; k)\rangle(x) = 1\right] - \mathrm{negl}(\lambda) \quad .$$

Algorithm 2: $\tilde{\mathsf{P}}(x)$ - Malicious Quantum Prover for $\langle \mathsf{P}, \mathsf{V} \rangle$

1 Sample $(i, b) \leftarrow \{0, \ldots, q-1\} \times \{0,1\} \cup \{(q, 0)\}$.
2 Sample $k \leftarrow \{0,1\}^{\lambda}$.
3 Run $\mathsf{rP}_{0 \to i}^{\tilde{\mathsf{V}}(x,\cdot;k)} |\psi_0\rangle$ and denote the resulting state $|\psi_i^{\tilde{\mathsf{V}}(x,\cdot;k)}\rangle$.
4 Measure the query register to obtain a value α and send it as the first message. Denote the state after measurement by $|\phi_i^{\tilde{\mathsf{V}}(x,\cdot;k)}(\alpha)\rangle$.
5 Upon receiving the second message β, run $\left(\mathsf{rP}_{i+b \to q}^{\tilde{\mathsf{V}}(x,\cdot;k)_{\alpha \to \beta}}\right)\left(\mathsf{rP}_{i \to i+b}^{\tilde{\mathsf{V}}(x,\cdot;k)}\right)|\phi_i^{\tilde{\mathsf{V}}(x,\cdot;k)}(\alpha)\rangle$.
6 Measure A, Z to obtain (α', γ) if $\alpha' = \alpha$ output γ as the third message, otherwise abort.

Proof. We denote by $\tilde{\mathsf{V}}^R$ a version of $\tilde{\mathsf{V}}$ such that $\tilde{\mathsf{V}}$ uses a truly random function R to derive its randomness (i.e. it runs $\mathsf{V}(x, \cdot, R(\alpha))$ for a first message α). From the pseudo-randomness of the PRF it holds that,

$$\Pr_k\left[\langle \mathsf{rP}, \tilde{\mathsf{V}}(x, \cdot; k)\rangle(x) = 1\right] - \mathrm{negl}(\lambda) \leq \mathop{\mathbb{E}}_R\left[\Pr\left[\langle \mathsf{rP}, \tilde{\mathsf{V}}^R\rangle(x) = 1\right]\right] \tag{1}$$

We also denote $\tilde{\mathsf{P}}^R$ to be the malicious prover that uses $\tilde{V}^R$ (where R is a truly random function) instead of $\mathsf{V}(x, \cdot; k)$ as the oracle for rP. Again by pseudo-randomness of the PRF it holds that,

$$\Pr\left[\langle \tilde{\mathsf{P}}, \mathsf{V}\rangle(x) = 1\right] \geq \mathop{\mathbb{E}}_R\left[\Pr\left[\langle \tilde{\mathsf{P}}^R, \mathsf{V}\rangle(x) = 1\right]\right] - \mathrm{negl}(\lambda) \tag{2}$$

We define the event $W(i, b, \alpha, r, R)$ to be the event where after sampling an external verifier's randomness r, sampling i, b by $\tilde{\mathsf{P}}^R$ and measuring α as the first message in stage 4, $\tilde{\mathsf{P}}^R$ succeeds in convincing the external verifier. Then it holds that,

$$\begin{aligned}\mathop{\mathbb{E}}_R\left[\Pr\left[\langle \tilde{\mathsf{P}}^R, \mathsf{V}\rangle(x) = 1\right]\right] &= \mathop{\mathbb{E}}_{r,R}\left[\Pr\left[\langle \tilde{\mathsf{P}}^R, \mathsf{V}(x; r)\rangle(x) = 1\right]\right] \\ &= \sum_{\alpha} \mathop{\mathbb{E}}_{r,R}\left[\mathop{\mathbb{E}}_{i,b}\left[\Pr\left[W(i, b, \alpha, r, R)\right]\right]\right] .\end{aligned}$$

Also, we note that,

$$\Pr\left[W(i, b, \alpha, r, R)\right] = \left\| |\alpha\rangle\langle\alpha| \otimes \Pi^{\alpha}_{\mathsf{V}(x,\cdot;r)}\left(\mathsf{rP}_{i+b \to q}^{\tilde{\mathsf{V}}^R_{\alpha \to \mathsf{V}(x,\alpha;r)}}\right)\left(\mathsf{rP}_{i \to i+b}^{\tilde{\mathsf{V}}^R}\right)|\alpha\rangle\langle\alpha||\psi_i^{\tilde{\mathsf{V}}^R}\rangle \right\|^2 ,$$

where

$$\Pi^{\alpha}_f = \sum_{c : f(\alpha, f(\alpha), c) = 1} |c\rangle\langle c| ,$$

the first $|\alpha\rangle\langle\alpha|$ is applied to the query register, the second $|\alpha\rangle\langle\alpha|$ is applied to the A register, and $\Pi^{\alpha}_{V(x,\cdot;r)}$ is applied to the Z register. Hence, it holds,

$$\mathop{\mathbb{E}}_{R}\left[\Pr\left[\langle\tilde{\mathsf{P}}^R,\mathsf{V}\rangle(x)=1\right]\right]=$$
$$\sum_{\alpha}\mathop{\mathbb{E}}_{r,R}\left[\mathop{\mathbb{E}}_{i,b}\left[\left\|\left|\alpha\rangle\langle\alpha\right|\otimes\Pi^{\alpha}_{\mathsf{V}(x,\cdot;r)}\left(\mathsf{rP}^{\tilde{\mathsf{V}}^R_{\alpha\to\mathsf{V}(x,\alpha;r)}}_{i+b\to q}\right)\left(\mathsf{rP}^{\tilde{\mathsf{V}}^R}_{i\to i+b}\right)|\alpha\rangle\langle\alpha||\psi^{\tilde{\mathsf{V}}^R}_{i}\rangle\right\|^2\right]\right] .$$

For any fixed α, r, R by the single value reprogramming lemma (1), it holds that,

$$\mathop{\mathbb{E}}_{i,b}\left[\left\|\left|\alpha\rangle\langle\alpha\right|\otimes\Pi^{\alpha}_{\mathsf{V}(x,\cdot;r)}\left(\mathsf{rP}^{\tilde{\mathsf{V}}^R_{\alpha\to\mathsf{V}(x,\alpha;r)}}_{i+b\to q}\right)\left(\mathsf{rP}^{\tilde{\mathsf{V}}^R}_{i\to i+b}\right)|\alpha\rangle\langle\alpha||\psi^{\tilde{\mathsf{V}}^R}_{i}\rangle\right\|^2\right]\geq$$
$$\frac{\left\|(|\alpha\rangle\langle\alpha|)\otimes\Pi^{\alpha}_{\mathsf{V}(x,\cdot;r)}|\psi^{\tilde{\mathsf{V}}^R_{\alpha\to\mathsf{V}(x,\alpha;r)}}_{q}\rangle\right\|^2}{(2q+1)^2} .$$

Above, $|\psi^{\tilde{\mathsf{V}}^R_{\alpha\to\mathsf{V}(x,\alpha;r)}}_{q}\rangle=\mathsf{rP}^{\tilde{V}^R_{\alpha\to\mathsf{V}(x,\alpha;r)}}|\psi_0\rangle$. Hence it holds that,

$$\begin{aligned}\mathop{\mathbb{E}}_{R}\left[\Pr\left[\langle\tilde{\mathsf{P}}^R,\mathsf{V}\rangle(x)=1\right]\right]&\geq\sum_{\alpha}\mathop{\mathbb{E}}_{r,R}\left[\frac{\left\|(|\alpha\rangle\langle\alpha|)\otimes\Pi^{\alpha}_{\mathsf{V}(x,\cdot;r)}|\psi^{\tilde{\mathsf{V}}^R_{\alpha\to\mathsf{V}(x,\alpha;r)}}_{q}\rangle\right\|^2}{(2q+1)^2}\right]\\&=\sum_{\alpha}\mathop{\mathbb{E}}_{r,R}\left[\frac{\left\|(|\alpha\rangle\langle\alpha|)\otimes\Pi^{\alpha}_{\tilde{\mathsf{V}}^R_{\alpha\to\mathsf{V}(x,\alpha;r)}}|\psi^{\tilde{\mathsf{V}}^R_{\alpha\to\mathsf{V}(x,\alpha;r)}}_{q}\rangle\right\|^2}{(2q+1)^2}\right]\\&\underset{(*)}{=}\sum_{\alpha}\mathop{\mathbb{E}}_{r,R}\left[\frac{\left\|(|\alpha\rangle\langle\alpha|)\otimes\Pi^{\alpha}_{\tilde{\mathsf{V}}^R}|\psi^{\tilde{\mathsf{V}}^R}_{q}\rangle\right\|^2}{(2q+1)^2}\right]\\&=\mathop{\mathbb{E}}_{R}\left[\frac{\Pr\left[\langle\mathsf{rP},\tilde{\mathsf{V}}^R\rangle(x)=1\right]}{(2q+1)^2}\right] ,\end{aligned}$$

where $(*)$ follows for any x,α and uniformly sampled r, R the oracles $\tilde{V}^R$ and $\tilde{V}^R_{\alpha\to(x,\alpha;r)}$ are perfectly indistinguishable. Thus, it holds

$$\mathop{\mathbb{E}}_{R}\left[\Pr\left[\langle\tilde{\mathsf{P}}^R,\mathsf{V}\rangle(x)=1\right]\right]\geq\mathop{\mathbb{E}}_{R}\left[\frac{\Pr\left[\langle\mathsf{rP},\tilde{\mathsf{V}}^R\rangle(x)=1\right]}{(2q+1)^2}\right] .$$

Hence, by combining Eqs. 1, 2 with the equation above, the claim follows.

4.3 Deterministic-Prefix Resetting Provers

5 A Post-Quantum Resettably Sound Zero Knowledge Protocol

In this section we present a post-quantum resettably-sound zero-knowledge protocol. The protocol is also constant-round.

Ingredients and Notation:

- A post-quantum pseudorandom function PRF.
- A post-quantum non-interactive commitment scheme Com.
- A post-quantum compute and compare obfuscator Obf.
- A quantum fully-homomorphic encryption scheme $(\mathsf{QFHE.Gen}, \mathsf{QFHE.Enc}, \mathsf{QFHE.QEnc}, \mathsf{QFHE.Dec}, \mathsf{QFHE.QDec}, \mathsf{QFHE.Eval})$.
- A delayed-input 3-message post-quantum WI proof $(\mathsf{WI.P}, \mathsf{WI.V})$ for **NP**.
- A delayed-input 4-message sub-exponential statistical WI argument system $(\mathsf{sWI.P}, \mathsf{sWI.V})$ for **NP**.
- A 2-message post-quantum input hiding, sub-exponentially statistically function hiding secure function evaluation scheme $(\mathsf{SFE.Gen}, \mathsf{SFE.Enc}, \mathsf{SFE.Eval}, \mathsf{SFE.Dec})$.
- Denote by $\varepsilon \in (0,1)$ a constant such that both the 4-message WI and SFE have sub-exponential statistical security with respect to (in the statistical indistinguishability guarantee in both primitives, the statistical distance is bounded by $O(2^{-\lambda^{\varepsilon}})$).

The protocol is described in Subsect. 5.1.

5.1 Protocol Construction

The protocol is as follows,

Common Input: An instance $x \in \mathcal{L}$, security parameter $\lambda := |x|$. Below we denote $\bar{\lambda} = \lambda^{2/\varepsilon}$.

P's **private input:** A classical witness $w \in \mathcal{R}_{\mathcal{L}}(x)$ for x.

1. **Prover Commitment:** P sends the following,
   - Non-interactive commitments to the witness, and two strings of zeros of length $\bar{\lambda}$:

     $$\mathsf{cmt}_1 \leftarrow \mathsf{Com}(1^{\lambda}, w), \;\; \mathsf{cmt}_2 \leftarrow \mathsf{Com}(1^{\lambda}, 0^{\bar{\lambda}}), \;\; \mathsf{cmt}_3 \leftarrow \mathsf{Com}(1^{\lambda}, 0^{\bar{\lambda}}) \;\; .$$

   - Two independent first messages α_1, α_2 for two independent executions of 3-message, delayed-input WI proofs $(\mathsf{WI.P}, \mathsf{WI.V})$.
   - First message h of a 4-message delayed-input statistical WI argument $(\mathsf{sWI.P}, \mathsf{sWI.V})$, with security parameter $\bar{\lambda}$.
2. **Extractable Commitment to Verifier Secret:** V samples a PRF seed $s \leftarrow \{0,1\}^{\lambda}$. V's randomness for the first message is generated by applying $\mathsf{PRF}_s(\cdot)$ to the first prover message.

(a) V computes $u \leftarrow \{0,1\}^\lambda$, $v \leftarrow \{0,1\}^\lambda$, $(\mathsf{pk}, \mathsf{sk}) \leftarrow \mathsf{QFHE.Gen}(1^\lambda)$. V sends

$$\mathsf{pk},\ \ \mathsf{ct}_\mathsf{V} \leftarrow \mathsf{QFHE.Enc}_{\mathsf{pk}}(u),\ \ \widetilde{\mathbf{CC}} \leftarrow \mathsf{Obf}\Big(\mathbf{CC}\big[\mathsf{QFHE.Dec}_{\mathsf{sk}}(\cdot), v, \mathsf{sk}\big]\Big)\ .$$

V also sends β_1, β_2 following α_1, α_2, and $\alpha_\mathbf{s}$ following h.

(b) P sends,
- $\mathsf{ct}_\mathsf{P} \leftarrow \mathsf{SFE.Enc}(1^{\bar{\lambda}}; 0^\lambda)$ an encryption of 0^λ encrypted with security parameter $\bar{\lambda}$.
- $\beta_\mathbf{s}$ for $h, \alpha_\mathbf{s}$ as the last message of $\mathsf{sWI.V}$ in the 4-message WI protocol.
- A WI proof γ_1, following α_1 and β_1, that $x \in \mathcal{L}$ or, (1) the randomness used to generate ct_P is the content of cmt_2[1], and (2) the randomness for $h, \beta_\mathbf{s}$ is the content of cmt_3.

(c) V applies $\mathsf{PRF}_s(\cdot)$ to $(\mathsf{ct}_\mathsf{P}, \beta_\mathbf{s},$ Prover's first message) to generate randomness for its current message. It sends,
- $\hat{\mathsf{ct}} \leftarrow \mathsf{SFE.Eval}\Big(\mathbf{CC}\big[\mathsf{Id}(\cdot), u, v\big], \mathsf{ct}_\mathsf{P}\Big)$ executed with security parameter $\bar{\lambda}$, where $\mathsf{Id}(\cdot)$ is the identity function.
- $\gamma_\mathbf{s}$, for $h, \alpha_\mathbf{s}, \beta_\mathbf{s}$, proving that the transcript of the verifier so far is explainable or, cmt_1 is a commitment to a non-witness $z \notin \mathcal{R}_\mathcal{L}(x)$. The witness that V uses for the proof is its randomness, that proves that the transcript is explainable.

3. **Final WI by the Prover:** P sends γ_2 which proves that $x \in \mathcal{L}$ or, that cmt_1 is a valid commitment and there exists a string c such that $\widetilde{\mathbf{CC}}(c) \neq \bot$. The witness that P uses for its proofs γ_1, γ_2 is w, which proves $x \in \mathcal{L}$.
4. **Acceptance:** V accepts if the WI statements by the prover are verified.
5. **Aborts:** During the protocol, if either party does not respond, sends a message of an incorrect form or provides a non-convincing WI proof it considered as an abort, and the other party terminates the interaction.

References

[ABB+17] Alkim, E., et al.: Revisiting TESLA in the quantum random oracle model. In: Lange, T., Takagi, T. (eds.) PQCrypto 2017. LNCS, vol. 10346, pp. 143–162. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59879-6_9

[ABDS20] Alagic, G., Brakerski, Z., Dulek, Y., Schaffner, C.: Impossibility of quantum virtual black-box obfuscation of classical circuits. CoRR, abs/2005.06432 (2020)

[ABG+20] Agarwal, A., Bartusek, J., Goyal, V., Khurana, D., Malavolta, G.: Post-quantum multi-party computation. IACR Cryptol. ePrint Arch. **2020**, 1395 (2020)

[AF16] Alagic, G., Fefferman, B.: On quantum obfuscation. CoRR, abs/1602.01771 (2016)

[AHU19] Ambainis, A., Hamburg, M., Unruh, D.: Quantum security proofs using semi-classical oracles. In: Boldyreva, A., Micciancio, D. (eds.) CRYPTO 2019. LNCS, vol. 11693, pp. 269–295. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26951-7_10

[1] Formally, there are strings r_1, r_2, r_3 such that $\mathsf{ct}_\mathsf{P} = \mathsf{SFE.Enc}(r_3; r_2)$, $\mathsf{cmt}_2 = \mathsf{Com}(1^\lambda, r_2; r_1)$.

[AP20a] Ananth, P., La Placa, R.L.: Secure quantum extraction protocols. In: Pass, R., Pietrzak, K. (eds.) TCC 2020. LNCS, vol. 12552, pp. 123–152. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64381-2_5

[AP20b] Ananth, P., La Placa, R.L.: Secure software leasing. CoRR, abs/2005.05289 (2020)

[Bar01] Barak, B.: How to go beyond the black-box simulation barrier. In: Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science, FOCS '01, USA, p. 106. IEEE Computer Society (2001)

[BCC88] Brassard, G., Chaum, D., Crépeau, C.: Minimum disclosure proofs of knowledge. J. Comput. Syst. Sci. **37**(2), 156–189 (1988)

[BDF+10] Boneh, D., Dagdelen, Ö., Fischlin, M., Lehmann, A., Schaffner, C., Zhandry, M.: Random oracles in a quantum world. IACR Cryptol. ePrint Arch. **2010**, 428 (2010)

[BFJ+20] Badrinarayanan, S., Fernando, R., Jain, A., Khurana, D., Sahai, A.: Statistical ZAP arguments. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020. LNCS, vol. 12107, pp. 642–667. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-45727-3_22

[BGGL01] Barak, B., Goldreich, O., Goldwasser, S., Lindell, Y.: Resettably-sound zero-knowledge and its applications. In: 42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, Las Vegas, Nevada, USA, 14–17 October 2001, pp. 116–125. IEEE Computer Society (2001)

[BGI+12] Barak, B., et al.: On the (im)possibility of obfuscating programs. J. ACM **59**(2), 6:1-6:48 (2012)

[BKP18] Bitansky, N., Kalai, Y.T., Paneth, O.: Multi-collision resistance: a paradigm for keyless hash functions. In: Diakonikolas, I., Kempe, D., Henzinger, M. (eds.) Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, 25–29 June 2018, pp. 671–684. ACM (2018)

[BKP19] Bitansky, N., Khurana, D., Paneth, O.: Weak zero-knowledge beyond the black-box barrier. In: Charikar, M., Cohen, E. (eds.) Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, 23–26 June 2019, pp. 1091–1102. ACM (2019)

[BL04] Barak, B., Lindell, Y.: Strict polynomial-time in simulation and extraction. SIAM J. Comput. **33**(4), 738–818 (2004)

[Blu86] Blum, M.: How to prove a theorem so no one else can claim it. In: Proceedings of the International Congress of Mathematicians, vol. 1, p. 2. Citeseer (1986)

[BLV06] Barak, B., Lindell, Y., Vadhan, S.P.: Lower bounds for non-black-box zero knowledge. J. Comput. Syst. Sci. **72**(2), 321–391 (2006)

[BP12] Bitansky, N., Paneth, O.: From the impossibility of obfuscation to a new non-black-box simulation technique. In: 53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, 20–23 October 2012, pp. 223–232. IEEE Computer Society (2012)

[BP13] Bitansky, N., Paneth, O.: On the impossibility of approximate obfuscation and applications to resettable cryptography. In: Boneh, D., Roughgarden, T., Feigenbaum, J. (eds.) Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, 1–4 June 2013, pp. 241–250. ACM (2013)

[BP15] Bitansky, N., Paneth, O.: On non-black-box simulation and the impossibility of approximate obfuscation. SIAM J. Comput. **44**(5), 1325–1383 (2015)

[Bra18] Brakerski, Z.: Quantum FHE (almost) as secure as classical. In: Shacham, H., Boldyreva, A. (eds.) CRYPTO 2018. LNCS, vol. 10993, pp. 67–95. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-96878-0_3

[BS20] Bitansky, N., Shmueli, O.: Post-quantum zero knowledge in constant rounds. In: Makarychev, K., Makarychev, Y., Tulsiani, M., Kamath, G., Chuzhoy, J. (eds.) Proccedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, 22–26 June 2020, pp. 269–279. ACM (2020)

[BZ13] Boneh, D., Zhandry, M.: Secure signatures and chosen ciphertext security in a quantum computing world. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013. LNCS, vol. 8043, pp. 361–379. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40084-1_21

[CCLY21] Chia, N.-H., Chung, K.-M., Liu, Q., Yamakawa, T.: On the impossibility of post-quantum black-box zero-knowledge in constant rounds. IACR Cryptol. ePrint Arch. **2021**, 376 (2021)

[CCY20] Chia, N.-H., Chung, K.-M., Yamakawa, T.: A black-box approach to post-quantum zero-knowledge in constant rounds. IACR Cryptol. ePrint Arch. **2020**, 1384 (2020)

[CGGM00] Canetti, R., Goldreich, O., Goldwasser, S., Micali, S.: Resettable zero-knowledge (extended abstract). In: Yao, F.F., Luks, E.M. (eds.) Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, Portland, OR, USA, 21–23 May 2000, pp. 235–244. ACM (2000)

[CMS19] Chiesa, A., Manohar, P., Spooner, N.: Succinct arguments in the quantum random oracle model. In: Hofheinz, D., Rosen, A. (eds.) TCC 2019. LNCS, vol. 11892, pp. 1–29. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-36033-7_1

[COP+14] Chung, K.-M., Ostrovsky, R., Pass, R., Venkitasubramaniam, M., Visconti, I.: 4-round resettably-sound zero knowledge. In: Lindell, Y. (ed.) TCC 2014. LNCS, vol. 8349, pp. 192–216. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-54242-8_9

[COPV13] Chung, K.M., Ostrovsky, R., Pass, R., Visconti, I.: Simultaneous resettability from one-way functions. In: 2013 IEEE 54th Annual Symposium on Foundations of Computer Science, pp. 60–69. IEEE (2013)

[COSV12] Cho, C., Ostrovsky, R., Scafuro, A., Visconti, I.: Simultaneously resettable arguments of knowledge. In: Cramer, R. (ed.) TCC 2012. LNCS, vol. 7194, pp. 530–547. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28914-9_30

[COV17] Chongchitmate, W., Ostrovsky, R., Visconti, I.: Resettably-sound resettable zero knowledge in constant rounds. In: Kalai, Y., Reyzin, L. (eds.) TCC 2017. LNCS, vol. 10678, pp. 111–138. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70503-3_4

[CPS13] Chung, K.M., Pass, R., Seth, K.: Non-black-box simulation from one-way functions and applications to resettable security. In: Boneh, D., Roughgarden, T., Feigenbaum, J. (eds.) Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, 1–4 June 2013, pp. 231–240. ACM (2013)

[CPS16] Chung, K.-M., Pass, R., Seth, K.: Non-black-box simulation from one-way functions and applications to resettable security. SIAM J. Comput. **45**(2), 415–458 (2016)

[DFM20] Don, J., Fehr, S., Majenz, C.: The measure-and-reprogram technique 2.0: multi-round fiat-shamir and more. In: Micciancio, D., Ristenpart, T. (eds.) CRYPTO 2020. LNCS, vol. 12172, pp. 602–631. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-56877-1_21

[DFMS19] Don, J., Fehr, S., Majenz, C., Schaffner, C.: Security of the fiat-shamir transformation in the quantum random-oracle model. In: Boldyreva, A., Micciancio, D. (eds.) CRYPTO 2019. LNCS, vol. 11693, pp. 356–383. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26951-7_13

[DFS04] Damgård, I., Fehr, S., Salvail, L.: Zero-knowledge proofs and string commitments withstanding quantum attacks. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 254–272. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-28628-8_16

[DGS09] Deng, Y., Goyal, V., Sahai, A.: Resolving the simultaneous resettability conjecture and a new non-black-box simulation strategy. In: 2009 50th Annual IEEE Symposium on Foundations of Computer Science, pp. 251–260. IEEE (2009)

[DNRS03] Dwork, C., Naor, M., Reingold, O., Stockmeyer, L.J.: Magic functions. J. ACM **50**(6), 852–921 (2003)

[ES15] Eaton, E., Song, F.: Making existential-unforgeable signatures strongly unforgeable in the quantum random-oracle model. In: Beigi, S., König, R. (eds.) 10th Conference on the Theory of Quantum Computation, Communication and Cryptography, TQC 2015, Brussels, Belgium, 20–22 May 2015, vol. 44 of LIPIcs, pp. 147–162. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2015)

[FGJ18] Fleischhacker, N., Goyal, V., Jain, A.: On the existence of three round zero-knowledge proofs. In: Nielsen, J.B., Rijmen, V. (eds.) EUROCRYPT 2018. LNCS, vol. 10822, pp. 3–33. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-78372-7_1

[FLS99] Feige, U., Lapidot, D., Shamir, A.: Multiple noninteractive zero knowledge proofs under general assumptions. SIAM J. Comput. **29**(1), 1–28 (1999)

[FP96] Fuchs, C.A., Peres, A.: Quantum-state disturbance versus information gain: uncertainty relations for quantum information. Phys. Rev. A **53**, 2038–2045 (1996)

[FS86] Fiat, A., Shamir, A.: How To prove yourself: practical solutions to identification and signature problems. In: Odlyzko, A.M. (ed.) CRYPTO 1986. LNCS, vol. 263, pp. 186–194. Springer, Heidelberg (1987). https://doi.org/10.1007/3-540-47721-7_12

[GJJM20] Goyal, V., Jain, A., Jin, Z., Malavolta, G.: Statistical zaps and new oblivious transfer protocols. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020. LNCS, vol. 12107, pp. 668–699. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-45727-3_23

[GK96a] Goldreich, O., Kahan, A.: How to construct constant-round zero-knowledge proof systems for NP. J. Cryptol. **9**(3), 167–189 (1996). https://doi.org/10.1007/BF00208001

[GK96b] Goldreich, O., Krawczyk, H.: On the composition of zero-knowledge proof systems. SIAM J. Comput. **25**(1), 169–192 (1996)

[GM11] Goyal, V., Maji, H.K.: Stateless cryptographic protocols. In: Ostrovsky, R. (ed.) IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, 22–25 October 2011, pp. 678–687. IEEE Computer Society (2011)

[GMR89] Goldwasser, S., Micali, S., Rackoff, C.: The knowledge complexity of interactive proof systems. SIAM J. Comput. **18**(1), 186–208 (1989)

[GMW91] Goldreich, O., Micali, S., Wigderson, A.: Proofs that yield nothing but their validity or all languages in np have zero-knowledge proof systems. J. ACM **38**(3), 690–728 (1991)

[Goy13] Goyal, V.: Non-black-box simulation in the fully concurrent setting. In: Boneh, D., Roughgarden, T., Feigenbaum, J. (eds.) Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, 1–4 June 2013, pp. 221–230. ACM (2013)

[GS09] Goyal, V., Sahai, A.: Resettably secure computation. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 54–71. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01001-9_3

[HI19] Hosoyamada, A., Iwata, T.: 4-round luby-rackoff construction is a qPRP. In: Galbraith, S.D., Moriai, S. (eds.) ASIACRYPT 2019. LNCS, vol. 11921, pp. 145–174. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-34578-5_6

[JKMR09] Jain, R., Kolla, A., Midrijanis, G., Reichardt, B.W.: On parallel composition of zero-knowledge proofs with black-box quantum simulators. Quant. Inf. Comput. **9**(5 & 6), 513–532 (2009)

[KLS18] Kiltz, E., Lyubashevsky, V., Schaffner, C.: A concrete treatment of fiat-shamir signatures in the quantum random-oracle model. In: Nielsen, J.B., Rijmen, V. (eds.) EUROCRYPT 2018. LNCS, vol. 10822, pp. 552–586. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-78372-7_18

[KNY20] Kitagawa, F., Nishimaki, R., Yamakawa, T.: Secure software leasing from standard assumptions. IACR Cryptol. ePrint Arch. **2020**, 1314 (2020)

[Kob03] Kobayashi, H.: Non-interactive quantum perfect and statistical zero-knowledge. In: Ibaraki, T., Katoh, N., Ono, H. (eds.) ISAAC 2003. LNCS, vol. 2906, pp. 178–188. Springer, Heidelberg (2003). https://doi.org/10.1007/978-3-540-24587-2_20

[KP01] Kilian, J., Petrank, E.: Concurrent and resettable zero-knowledge in poly-loalgorithm rounds. In: Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing, pp. 560–569 (2001)

[KRR17] Kalai, Y.T., Rothblum, G.N., Rothblum, R.D.: From obfuscation to the security of fiat-shamir for proofs. In: Katz, J., Shacham, H. (eds.) CRYPTO 2017. LNCS, vol. 10402, pp. 224–251. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-63715-0_8

[LZ19] Liu, Q., Zhandry, M.: Revisiting post-quantum fiat-shamir. In: Boldyreva, A., Micciancio, D. (eds.) CRYPTO 2019. LNCS, vol. 11693, pp. 326–355. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26951-7_12

[Mah18] Mahadev, U.: Classical homomorphic encryption for quantum circuits. In: 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS), pp. 332–338. IEEE (2018)

[MR01] Micali, S., Reyzin, L.: Min-round resettable zero-knowledge in the public-key model. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 373–393. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44987-6_23

[OV12] Ostrovsky, R., Visconti, I.: Simultaneous resettability from collision resistance. Electron. Colloquium Comput. Complex. **19**, 164 (2012)

[PTW11] Pass, R., Tseng, W.L.D., Wikström, D.: On the composition of public-coin zero-knowledge protocols. SIAM J. Comput. **40**(6), 1529–1553 (2011)

[Reg05] Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. In: Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing, STOC '05, New York, NY, USA, pp. 84–93. Association for Computing Machinery (2005)

[TU16] Targhi, E.E., Unruh, D.: Post-quantum security of the fujisaki-okamoto and OAEP transforms. In: Hirt, M., Smith, A. (eds.) TCC 2016. LNCS, vol. 9986, pp. 192–216. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53644-5_8

[Unr14] Unruh, D.: Quantum position verification in the random oracle model. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014. LNCS, vol. 8617, pp. 1–18. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44381-1_1

[Unr15] Unruh, D.: Non-interactive zero-knowledge proofs in the quantum random oracle model. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9057, pp. 755–784. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-46803-6_25

[Unr16a] Unruh, D.: Collapse-binding quantum commitments without random oracles. In: Cheon, J.H., Takagi, T. (eds.) ASIACRYPT 2016. LNCS, vol. 10032, pp. 166–195. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53890-6_6

[Unr16b] Unruh, D.: Computationally binding quantum commitments. In: Fischlin, M., Coron, J.-S. (eds.) EUROCRYPT 2016. LNCS, vol. 9666, pp. 497–527. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49896-5_18

[VDGC97] Van De Graaf, J., Crepeau, C.: Towards a formal definition of security for quantum protocols. Université de Montréal (1997)

[Wat02] Watrous, J.: Limits on the power of quantum statistical zero-knowledge. In: 43rd Symposium on Foundations of Computer Science (FOCS 2002), Vancouver, BC, Canada, 16–19 November 2002, Proceedings, p. 459. IEEE Computer Society (2002)

[Wat09] Watrous, J.: Zero-knowledge against quantum attacks. SIAM J. Comput. **39**(1), 25–58 (2009)

[WZ82] Wootters, W.K., Zurek, W.H.: A single quantum cannot be cloned. Nature **299**(5886), 802–803 (1982)

[Zha12] Zhandry, M.: How to construct quantum random functions. In 53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, 20–23 October 2012, pp. 679–687. IEEE Computer Society (2012)

[Zha15] Zhandry, M.: Secure identity-based encryption in the quantum random oracle model. Int. J. Quant. Inf. **13**(04), 1550014 (2015)

[Zha18] Zhandry, M.: How to record quantum queries, and applications to quantum indifferentiability. IACR Cryptol. ePrint Arch. **2018**, 276 (2018)