On The Reprogrammability of Quantum Random Oracle: Impossibility Results on Post-Quantum Zero-knowledge

Instructed by *Prabhanjan Ananth*

Zikuan Huang

October 29, 2024

Contents

1 Introduction and Importance

The classical random oracle model, introduced by [BR93], has proven to be a powerful framework for establishing the security of various cryptographic schemes when direct security proofs in the plain model are not feasible. However, with the advent of quantum computing, there has been a growing focus on 'post-quantum' security, which addresses the robustness of cryptographic schemes against quantum adversaries.

To address post-quantum security, [BDF+11] developed a model to incorporate the effects of quantum computations into the random oracle framework. In this quantum random oracle model, a quantum adversary can query the random oracle in a superposition of states. Specifically, for a random oracle $\mathcal{O}$, a quantum adversary can make a query of the form:

$$|x\rangle |y\rangle \to |x\rangle |y \oplus \mathcal{O}(x)\rangle ,$$

where the adversary's query state is transformed into a superposition that includes the output of the random oracle.

This raises an important question: To what extent can a classical reduction under the random oracle model be extended to a post-quantum reduction? This question is crucial, given recent results such as [YZ22], which suggest that quantum algorithms can be significantly more powerful than classical algorithms within the random oracle model. Understanding the limitations and capabilities of extending classical reductions to the quantum setting is essential for ensuring the security of cryptographic schemes in a post-quantum world.

2 Problems

In this work, we focus on a specific type of reduction. In the classical setting, if an algorithm $\mathcal{A}$ can solve problem A within the random oracle model, then it is possible to construct an algorithm $\mathcal{B}$ that solves problem B in the plain model. The algorithm $\mathcal{B}$ operates as follows: it runs $\mathcal{A}$ and simulates the random oracle $\mathcal{O}$ internally during the execution of $\mathcal{A}$.

More precisely, $\mathcal{B}$ is designed to emulate the behavior of the random oracle $\mathcal{O}$ in a manner that allows $\mathcal{A}$ to solve problem A. Through this simulation, $\mathcal{B}$ can leverage the solution provided by $\mathcal{A}$ to address problem B directly in the plain model.

3 Literature Review

In classical cryptography, proving the security of a scheme S under the random oracle model often involves demonstrating that any adversary $\mathcal{A}$ that breaks S in this model can be converted into an adversary $\mathcal{B}$ that breaks an underlying assumption T in the plain model. Typically, the reduction involves the following steps:

- $\mathcal{B}$ runs $\mathcal{A}$ with the appropriate parameters.
- $\mathcal{B}$ records some of $\mathcal{A}$'s queries to the random oracle using a specific strategy (e.g., recording all queries or selecting some at random) and responds with simulated answers as if it were the random oracle.
- Using the recorded queries and $\mathcal{A}$'s behavior, $\mathcal{B}$ then attempts to break the assumption T.

3.1 Security of Fiat-Shamir Transformation under the Random Oracle Model (ROM)

One well-known example of such a reduction is the security of the Fiat-Shamir transformation under the random oracle model. Consider an adversary $\mathcal{A}$ that forges a tuple (α, β, γ) such that:

- $\mathcal{O}(\alpha) = \beta$, where $\mathcal{O}$ is the random oracle used as the hash function in the Fiat-Shamir transformation.
- The verifier accepts the tuple (α, β, γ).

The reduction works as follows, assuming $\mathcal{A}$ does not query the same α more than once:

- Suppose $\mathcal{A}$ makes $p(\lambda)$ queries to the random oracle, where $p(\cdot)$ is a fixed polynomial. Then $\mathcal{B}$ samples a random index $i \leftarrow [p(\lambda)]$ and maintains a simulated random oracle $\mathcal{O}$.
- For the first $i-1$ queries, $\mathcal{B}$ responds to $\mathcal{A}$ using $\mathcal{O}$. The i-th query α_i is forwarded to an external verifier.
- Upon receiving a response β from the external verifier, $\mathcal{B}$ updates the simulated oracle $\mathcal{O}$ to ensure $\mathcal{O}(\alpha) = \beta$ and uses this updated oracle for subsequent queries.

Intuitively, $\mathcal{A}$ cannot distinguish whether the response from the random oracle is randomly sampled or provided by the external verifier, since the verifier uses a public-coin protocol.

3.2 Post-Quantum Security of Fiat-Shamir Transformation under the Quantum Random Oracle Model (QROM)

In recent work [DFMS19], the author investigates the security of the Fiat-Shamir transformation when the adversary $\mathcal{A}$ is a quantum adversary with quantum access to the random oracle $\mathcal{O}$. In this model, the adversary can perform superposition queries of the form:

$$|x\rangle |y\rangle \to |x\rangle |y \oplus \mathcal{O}(x)\rangle .$$

A key challenge here is that recording queries made by the quantum adversary is difficult, as any measurement of the query register could disrupt the adversary's state. However, since the adversary ultimately outputs a classical tuple (α, β, γ), it may be possible to measure some queries without significantly affecting the adversary's output. In fact, [DFMS19] provides a theorem stating (informally) that:

- Let $p^{\mathcal{O}}_{\mathcal{A}}$ be the success probability of the original adversary $\mathcal{A}$ with respect to a random oracle $\mathcal{O}$.
- After measurement and reprogramming, let (α^*, β^*) be the recorded query and response from the external verifier. Let $p^{\mathcal{O}}_{\mathcal{A},\mathsf{rep}}$ be the probability that $\mathcal{A}$ outputs an accepted transcript (α, β, γ) with $\alpha = \alpha^*$ and $\beta = \beta^*$. Then, $p^{\mathcal{O}}_{\mathcal{A},\mathsf{rep}} \geq \frac{1}{\mathsf{poly}} p^{\mathcal{O}^*}_{\mathcal{A}}$, where $\mathcal{O}^*$ is the modified oracle such that $\mathcal{O}^*(\alpha) = \beta$.

This result relies on extracting only one query-response pair. Later work [DFM20] extends this approach to extracting a constant number of query-response pairs, but the multiplicative factor on the success probability becomes $\frac{1}{\mathrm{poly}^k}$, where k is the number of pairs. Additionally, [CCLY21] used a variant of this lemma to demonstrate the impossibility of post-quantum constant-round black-box zero-knowledge arguments.

3.3 Useful Tool: Compress Oracle

Another significant approach is introduced by [Zha18], which presents the Compress Oracle technique. This method offers a novel perspective on the problem. Specifically, a phase oracle can be as powerful as a standard oracle for any problem under a quantum oracle model. A quantum algorithm $\mathcal{A}$ with a standard oracle:

$$|x\rangle |y\rangle \to |x\rangle |y \oplus \mathcal{O}(x)\rangle$$

can be transformed into another algorithm $\mathcal{B}$ with a phase oracle:

$$|x\rangle |y\rangle \to (-1)^{\langle \mathcal{O}(x), y\rangle} |x\rangle |y\rangle .$$

Zhandry observed that the phase factor $(-1)^{\langle \mathcal{O}(x), y\rangle}$ can be interpreted both as the effect of the oracle query on the adversary's state and as the adversary's effect on the oracle. Therefore, if we consider the adversary's queries as impacting the oracle's database, it may be possible to record some of the queries. However, measuring the oracle database during the algorithm might disrupt the entanglement between the adversary and the oracle. [LZ19] used this technique to prove the security of the Fiat-Shamir transformation under the Quantum Random Oracle Model.

3.4 Impossibility of Public-Coin Unbounded-Parallel Black-Box Zero-Knowledge Proofs/Arguments

When a reduction requires extracting a polynomial number of query-response pairs, the techniques mentioned above (and all known techniques) are insufficient. A classic example is the impossibility of public-coin unbounded-parallel black-box zero-knowledge arguments, as demonstrated by [PTW11]. The classical reduction approach involves:

- Constructing a malicious verifier V^* for t sessions, where V^* outputs $\mathcal{O}(m)$, with m being the messages exchanged in all prover sessions, and $\mathcal{O}$ is the random oracle.
- Assuming the existence of a simulator S that interacts with V^* in a black-box manner, the goal is to convert S into a malicious prover P^*.
- The malicious prover P^* first samples a random index $i \leftarrow [t]$ and embeds the real-time interaction into the i-th session.

A key question is whether post-quantum public-coin unbounded-parallel black-box zero-knowledge proofs/arguments are still impossible. We aim to achieve at least one of the following results:

- A statement on whether it is possible to construct post-quantum public-coin unbounded-parallel black-box zero-knowledge proofs/arguments (likely to be impossible).
- A counterexample showing that this type of proof does not work, not necessarily by constructing post-quantum proofs/arguments but perhaps by demonstrating some trivial protocols where a simulator cannot be converted into a malicious prover.

4 Preliminaries

4.1 Notations

For most of the time, we use $A, B, \cdots$ to denote sets and distributions. We use $\mathcal{A}, \mathcal{B}, \cdots$ to denote algorithms and oracles.

4.2 Distribution and Entropy

Let X, Y be sets, then $\mathsf{Func}_{X \to Y}$ denotes the set of all functions mapping from X to Y. Let D be a distribution over X.

Definition 4.1 (Entropy)**.** *Let D be an arbitrary distribution over set X, the entropy of D is*

$$\mathrm{H}(D) = -\sum_{x \in X} D(x) \log_2 D(x).$$

Definition 4.2 (Top Sum and Top Weight Supports)**.** *Let D be an arbitrary distribution over set X, define $\mathsf{TS}_t(D)$ be the sum of weights on the largest t supports*

$$\mathsf{TS}_t(D) = \sum_{x \in \mathsf{TWS}_t(D)} D(x)$$

where $\mathsf{TWS}_t(D) \subseteq X$ is an subset of size t such that $\forall x \in X - \mathsf{TWS}_t(D), \mathsf{TWS}_t(D) \in \mathsf{TWS}_t(D)$ we have $D(x) \leq D(\mathsf{TWS}_t(D))$.

Lemma 4.3. *Let D be an arbitrary distribution over X that satisfies $\mathsf{TS}_t(D) \leq a$, then the entropy of distribution D, $H(D) \geq (1-a)(\log t - \log a)$.*

Proof. Let $\mathsf{TWS}_t(D)$ be the set defined above. For any $x \in X - \mathsf{TWS}_t(D)$, $D(x) \leq \frac{a}{t}$ thus

$$\mathrm{H}(D) \geq (1-a)(\log t - \log a).$$

Proof of convexity is omitted. □

Lemma 4.4 (Shearer's inequality)**.** *Let D be a distribution over sets $X_1 \times X_2 \times \cdot \times X_n$, let D_i be the marginal distribution of D over X_i. Then*

$$\mathrm{H}(D) \leq \sum_{i=1}^{n} \mathrm{H}(D_i).$$

Lemma 4.5 (Pinsker's inequality)**.** *Let P, Q be two distribution over set X, we have*

$$\|P - Q\|_1 \leq \sqrt{2 D_{\mathrm{KL}}(P, Q)}$$

where D_{KL} is the Kullback–Leibler divergence. When Q is the uniform distribution over X, it becomes

$$\|P - Q\|_1 \leq \sqrt{2(\log |X| - \mathrm{H}(P))}.$$

4.3 Quantum Random Oracles

Lemma 4.6 ([Zha12a], Corollary 7.5)**.** *Let X, Y, T be three sets. Let $\mathcal{A}$ be any quantum algorithm that makes q queries to an oracle $\mathcal{O} : X \to Y$. We have*

$$\left| \Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{X\to Y}} \left[\mathcal{A}^{\mathcal{O}}(\cdot) = 1\right] - \Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{T\to Y} \circ \mathsf{Func}_{X\to T}} \left[\mathcal{A}^{\mathcal{O}}(\cdot) = 1\right] \right| \leq O(q^3/|T|).$$

Lemma 4.7 ([Zha12b])**.** *Let X, Y be two sets. Let $\mathcal{A}$ be any quantum algorithm that makes q queries to an oracle $\mathcal{O} : X \to Y$. Let $\mathsf{Func}^{2q}_{X\to Y}$ be a distribution over $\mathsf{Func}_{X\to Y}$ such that it is $2q$-wise independent. We have*

$$\Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{X\to Y}} \left[\mathcal{A}^{\mathcal{O}}(\cdot) = 1\right] = \Pr_{\mathcal{O} \leftarrow \mathsf{Func}^{2q}_{X\to Y}} \left[\mathcal{A}^{\mathcal{O}}(\cdot) = 1\right].$$

4.4 Zero-knowledge Protocols

Definition 4.8 (Proof/Argument System)**.** *A protocol $\Pi = (P, V)$ is a post-quantum proof/argument for language $\mathcal{L} \in \mathsf{NP}$ if it satisfies:*

- ***Completeness:*** *For all $x \in \mathcal{L}$ there exists w such that*

$$\Pr\left[\langle P(x, w), V(x)\rangle = 1\right] = 1.$$

- ***Soundness:*** *For all $x \notin \mathcal{L}$, all possibly unbounded/QPT prover P^* and all non-uniform advice $|\psi_x\rangle$,*

$$\Pr\left[\mathsf{Output}_V\langle P^*(x, |\psi_x\rangle), V(x)\rangle = 1\right] < \mathsf{negl}.$$

Definition 4.9 (Zero-knowledge)**.** *A post-quantum proof/argument for language $\mathcal{L} \in \mathsf{NP}$ is*

- ***Black-box Zero-knowledge:*** *if there exists a black-box expected polynomial time quantum simulator $\mathcal{S}$ such that for all non-uniform QPT verifier $V^*(|\psi\rangle)$,*

$$\mathsf{View}_V\langle P(x), V^*(x, |\psi\rangle)\rangle \approx_c \mathcal{S}^{V^*(|\psi\rangle)}.$$

- ***Black-box Weak Zero-knowledge:*** *if there exists a black-box BQP simulator $\mathcal{S}$ such that for all non-uniform QPT verifier V^*, all distinguisher $\mathcal{D}$ and all $0 < \epsilon < 1$,*

$$\left|\Pr\left[\mathcal{D}\left(\mathsf{View}_V\langle P(x), V^*(x, |\psi\rangle)\rangle\right) = 1\right] - \Pr\left[\mathcal{D}\left(\mathcal{S}^{V^*(|\psi\rangle)}(x, 1^{1/\epsilon})\right) = 1\right]\right| \leq \epsilon.$$

- ***ϵ-Zero-knowledge:*** *if there exists a black-box BQP simulator $\mathcal{S}$ such that for all non-uniform QPT verifier V^*, all distinguisher $\mathcal{D}$,*

$$\left|\Pr\left[\mathcal{D}\left(\mathsf{View}_V\langle P(x), V^*(x, |\psi\rangle)\rangle\right) = 1\right] - \Pr\left[\mathcal{D}\left(\mathcal{S}^{V^*(|\psi\rangle)}(x)\right) = 1\right]\right| \leq \epsilon.$$

5 Impossibility on Public-coin Unbounded-parallel Zero-knowledge Proofs

To prove the impossibility of public-coin unbounded-parallel zero-knowledge proofs, consider a simulator $\mathcal{S}$ that produces a transcript indistinguishable from a real transcript. The central idea of our proof is that, if the number of sessions is sufficiently large, there must exist at least one session where the distribution of the verifier's messages output by the simulator has nearly maximum entropy. In this session, we can construct a malicious prover that uses the transcript to convince the verifier of any $x \notin \mathcal{L}$.

To establish this, we need to demonstrate that the total entropy of the verifier's messages is large [claim 5.5]. This involves invoking Zhandry's lemma on quantum random oracles to argue that the entropy is sufficiently high.

Lemma 5.1. *Let $\mathsf{Func}_{X\to Y}$ be the set of functions mapping set X to set Y. Let $S_x \subseteq Y$ be a subset of Y such that $|S_x| = s \ll |Y|$ for all $x \in X$. We abuse the notation and let $\mathsf{Func}_{X\to Y-S_x}$ be the set of functions f mapping X to Y such that $f(x) \notin S_x$ for all $x \in X$. For any quantum algorithm $\mathcal{A}$ (possibly inefficient) that makes $q \ll s$ queries to an oracle $\mathcal{O}$, we have*

$$\left| \Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{X\to Y}} \left[\mathcal{A}^{\mathcal{O}}(\cdot) = 1\right] - \Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{X\to Y-S_x}} \left[\mathcal{A}^{\mathcal{O}}(\cdot) = 1\right] \right| \leq O(\sqrt{q^3 s/|Y|}).$$

Proof. Let S be an arbitrary subset of Y with size s. Suppose there exists quantum algorithm $\mathcal{A}$ such that

$$\left| \Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{X\to Y}} \left[\mathcal{A}^{\mathcal{O}}(\cdot) = 1\right] - \Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{X\to Y-S_x}} \left[\mathcal{A}^{\mathcal{O}}(\cdot) = 1\right] \right| > O(\sqrt{q^3 s/|Y|}).$$

Then there exists an quantum algorithm $\mathcal{B}$ that makes $2q$ queries to $\mathcal{O}$ such that

$$\left| \Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{X\to Y}} \left[\mathcal{B}^{\mathcal{O}}(\cdot) = 1\right] - \Pr_{\mathcal{O} \xleftarrow{\$} \mathsf{Func}_{X\to Y-S}} \left[\mathcal{B}^{\mathcal{O}}(\cdot) = 1\right] \right| > O(\sqrt{q^3 s/|Y|}).$$

$\mathcal{B}$ works as follows. Let $g_x : Y - S_x \to Y - S$ be a possibly inefficient one-to-one mapping. Whenever $\mathcal{A}$ calls the oracle

$$\sum_{x\in X} \alpha_x |x\rangle |y\rangle \to \sum_{x\in X} \alpha_x |x\rangle |\mathcal{O}(x) \oplus y\rangle .$$

$\mathcal{B}$ does the following:

1. Queries the oracle in a temporary empty register

$$\sum_{x\in X} \alpha_x |x\rangle |y\rangle \to \sum_{x\in X} \alpha_x |x\rangle |y\rangle |\mathcal{O}(x)\rangle .$$

2. Maps the response to $Y - S$

$$\sum_{x\in X} \alpha_x |x\rangle |y\rangle |\mathcal{O}(x)\rangle \to \sum_{x\in X} \alpha_x |x\rangle |y\rangle |g_x(\mathcal{O}(x))\rangle .$$

3. Updates the answer register

$$\sum_{x\in X}\alpha_x\left|x\right\rangle\left|y\right\rangle\left|\mathcal{O}(x)\right\rangle\rightarrow\sum_{x\in X}\alpha_x\left|x\right\rangle\left|y\oplus g_x(\mathcal{O}(x))\right\rangle\left|g_x(\mathcal{O}(x))\right\rangle.$$

4. Queries the oracle again and uncompute the temporary register, then trace out the temporary register

$$\sum_{x\in X}\alpha_x\left|x\right\rangle\left|y\oplus g_x(\mathcal{O}(x))\right\rangle\left|g_x(\mathcal{O}(x))\right\rangle\rightarrow\sum_{x\in X}\alpha_x\left|x\right\rangle\left|y\oplus g_x(\mathcal{O}(x))\right\rangle.$$

So we only need to prove that

$$\left|\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{X\to Y}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]-\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{X\to Y-S}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]\right|\leq O(\sqrt{q^3s/\left|Y\right|}).$$

Let T be a set of size t, by lemma 4.6,

$$\left|\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{X\to Y}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]-\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{T\to Y}\circ\mathsf{Func}_{X\to T}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]\right|\leq O(q^3/t).$$

$$\left|\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{X\to Y-S}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]-\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{T\to Y-S}\circ\mathsf{Func}_{X\to T}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]\right|\leq O(q^3/t).$$

Note that the probability that a random function $\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{T\to Y}$ is in $\mathsf{Func}_{T\to Y-S}$ is $\left(1-\frac{s}{|Y|}\right)^t=1-O(st/\left|Y\right|)$. Thus

$$\left|\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{T\to Y}\circ\mathsf{Func}_{X\to T}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]-\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{T\to Y-S}\circ\mathsf{Func}_{X\to T}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]\right|\leq O(st/\left|Y\right|).$$

Let $t=O(\sqrt{q^3\left|Y\right|/s})$, combine above inequalities we get

$$\left|\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{X\to Y}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]-\Pr_{\mathcal{O}\xleftarrow{\$}\mathsf{Func}_{X\to Y-S}}\left[\mathcal{B}^{\mathcal{O}}(\cdot)=1\right]\right|\leq O(\sqrt{q^3s/\left|Y\right|}).$$

□

Theorem 5.2. *Let $(P,\mathcal{V})$ be a post-quantum black-box ϵ zero-knowledge **proof** for any* NP *language $\mathcal{L}$ where all messages are $m(\lambda)$ bits long and contains $\ell(\lambda)$ rounds, if it satisfies:*

- *It is a public-coin protocol.*
- *It remains black-box ϵ-zero-knowledge under k-parallel repetition when $k\geq\lambda^2\ell^5$ and $\epsilon\leq\frac{1}{mk}$.*

Then $\mathcal{L}$ is in BQP.

Proof. Fix any instance x. Let the message space be $M = \{0,1\}^{mk}$, the number of rounds be $\ell = \ell(\lambda)$. Define $M^{\leq i} = \cup_{i=1}^{i} M^i$. Let $\mathcal{S}$ be the black-box simulator and let the number of queries made by the simulator be $q = q(k, \lambda)$. Let U be the uniform random distribution over $\mathsf{Func}_{M^{\leq \ell} \to M}$. Consider a k-parallel malicious verifier $\mathcal{V}^{\mathcal{O}}$ which has quantum access to a random oracle $\mathcal{O} : M^{\leq \ell} \to M$ sampled from U. $\mathcal{V}^{\mathcal{O}}$ on prover's message $m' \in M^{\leq \ell}$ responses $\mathcal{O}(m')$. Let $\mathcal{V}^U$ denote the verifier that samples an oracle $\mathcal{O}$ according to U and runs $\mathcal{V}^{\mathcal{O}}$.
Let $D^{\mathcal{V}}$ denote the distribution of the output of $\mathcal{S}^{\mathcal{V}}(x)$. Some notations are defined below

- Let $D^{\mathcal{V}}(\mathsf{L})$ where L is a list of message indices be the marginal distribution on those indices. For example, $D^{\mathcal{V}}(r_1, r_2)$ is the marginal distribution of (r_1, r_2) and $D^{\mathcal{V}}(m_1, \cdots, m_i)$ is the marginal distribution of $(m_1, \cdots, m_i)$.
- For a predicate E, let $D^{\mathcal{V}}(\cdot|\mathsf{E})$ be the distribution of the output of $\mathcal{S}^{\mathcal{V}}(x)$ condition on the outputted transcript satisfies E. For example, for fixed $m_i^*, r_j^* \in M$, $D^{\mathcal{V}}(\cdot|m_i = m_i^*, r_j = r_j^*)$ is the distribution of the output of $\mathcal{S}^{\mathcal{V}}(x)$ condition on the i-th prover's message being m_i^* and the j-th verifier's message being r_j^*. Combine this with the first bullet, we use $D^{\mathcal{V}}(\mathsf{L}|\mathsf{E})$ to denote the marginal distribution on L condition on E. For example, for fixed m_1^*, we use $D^{\mathcal{V}}(r_1|m_1 = m_1^*)$ to denote the marginal distribution of r_1 condition on the first prover's message being m_1^*.
- For two predicates E and F, let $d^{\mathcal{V}}(\mathsf{E}|\mathsf{F})$ be the probability that $\mathcal{S}^{\mathcal{V}}(x)$ outputs a transcript that satisfies E condition on that transcript satisfies F. For example, for fixed m_1^*, r_1^*, we use $d^{\mathcal{V}}(r_1 = r_1^*|m_1 = m_1^*)$ to denote the probability that a transcript with the prover's first message being m_1^* has response r_1^* from the verifier.

For all oracle distribution U^*, let $p_{\mathsf{Invalid}}^{U^*}$ be the probability that $\mathcal{S}^{\mathcal{V}^{\mathcal{O}}}$ outputs a transcript that contradicts the oracle $\mathcal{O}$ when $\mathcal{O}$ is sampled according U^*. That is,

$$p_{\mathsf{Invalid}}^{U^*} = \mathsf{Pr}\left[\exists i \in [\ell] \text{ s.t. } \mathcal{O}(m_1, \cdots, m_i) \neq r_i \middle| (m_1, r_1, \cdots, m_\ell, r_\ell) \leftarrow \mathcal{S}^{\mathcal{V}^{U^*}}(x)\right].$$

Claim 5.3. *The probability that $\mathcal{S}^{\mathcal{V}^U}$ outputs a transcript that contradicts its oracle is small. That is, $p_{\mathsf{Invalid}}^{U} \leq \epsilon$.*

Proof. If not, a trivial distinguisher that checks whether the outputting transcript is valid will distinguish the output of the simulator and the actual view of the verifier. □

Claim 5.4. *The sum of top $t = 2^{mk - \log^2 q}$ weights of the distribution of r_i condition on the prover's previous message being $(m_1^*, \cdots, m_i^*)$ is small **on average** over all $(m_1^*, \cdots, m_i^*) \in M^{\leq i}$. More specifically, for all $i \in [\ell]$,*

$$\sum_{(m_1^*, \cdots, m_i^*) \in M^i} d^{\mathcal{V}^U}\begin{pmatrix} m_1 = m_1^* \\ \cdots \\ m_i = m_i^* \end{pmatrix} \mathsf{TS}_t\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{matrix} m_1 = m_1^* \\ \cdots \\ m_i = m_i^* \end{matrix}\right)\right) \leq O\left(\sqrt{q^3/q^{\log q}}\right) + p_{\mathsf{Invalid}}^{U}.$$

Proof. The idea behind the proof is that if the distribution $D^{\mathcal{V}^U}\left(r_i \middle| \begin{matrix} m_1 = m_1^* \\ \cdots \\ m_i = m_i^* \end{matrix}\right)$ has too much weight on a small fraction of supports then after we modify the random oracle such that the oracle never answer those supports, either the output distribution of $\mathcal{S}$ or p_{Invalid} has to change dramatically.

But according to lemma 5.1, both quantities can only change slightly.
Now we formally prove this claim using lemma 5.1. For fixed $i \in [\ell]$ and $(m_1^*, \cdots, m_i^*) \in M^i$, let $S_{(m_1^*,\cdots,m_i^*)} = \mathsf{TWS}_t\left(D^{\mathcal{V}^U}(r_i|m_1 = m_1^*, \cdots, m_i^*)\right) \subset M$ be set of supports that has the t largest weight. Let U_i' be the uniform distribution over $\mathsf{Func}_{M^{\le\ell}\to M - S_{(m_1,\cdots,m_i)}}$. Consider the output distribution of $\mathcal{S}^{\mathcal{V}^{U'}}(x)$. By lemma 5.1, we have

$$\left\| D^{\mathcal{V}^U} - D^{\mathcal{V}^{U_i'}} \right\|_1 \le O(\sqrt{q^3 t/|M|}) = O\left(\sqrt{q^3/q^{\log q}}\right) \tag{1}$$

and

$$\left| p^U_{\mathsf{Invalid}} - p^{U_i'}_{\mathsf{Invalid}} \right| \le O\left(\sqrt{q^3/q^{\log q}}\right). \tag{2}$$

Note that only the image for i-round messages are punctured, but the lemma still holds since we can view a single random oracle as ℓ oracles, one for each round. For any outputted transcript $(m_1^*, r_1^*, \cdots, m_\ell^*, r_\ell^*)$ of $\mathcal{S}^{\mathcal{V}^{\mathcal{O}}}(x)$ where $\mathcal{O} \leftarrow U_i'$ if $r_i^* \in \mathsf{TWS}_t\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right)$ then by the definition of U_i' this transcript contradicts the oracle $\mathcal{O}$. By the definition of $p^{U_i'}_{\mathsf{Invalid}}$,

$$\sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^{U_i'}}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathsf{Pr}\left[r_i^* \in \mathsf{TWS}_t\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \middle| (m_1^*, r_1^*, \cdots, m_\ell^*, r_\ell^*) \leftarrow \mathcal{S}^{\mathcal{V}^{U_i'}}(x)\right] \le p^{U_i'}_{\mathsf{Invalid}}.$$

By eq. (1), eq. (2) and section 5, we have

$$\begin{aligned}
&\sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathsf{TS}_t\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \\
=&\sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathsf{Pr}\left[r_i^* \in \mathsf{TWS}_t\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \middle| (m_1^*, r_1^*, \cdots, m_\ell^*, r_\ell^*) \leftarrow \mathcal{S}^{\mathcal{V}^U}(x)\right] \\
\le&\sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^{U_i'}}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathsf{Pr}\left[r_i^* \in \mathsf{TWS}_t\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \middle| (m_1^*, r_1^*, \cdots, m_\ell^*, r_\ell^*) \leftarrow \mathcal{S}^{\mathcal{V}^U}(x)\right] \\
&+O\left(\sqrt{q^3/q^{\log q}}\right) \\
\le&\sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^{U_i'}}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathsf{Pr}\left[r_i^* \in \mathsf{TWS}_t\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \middle| (m_1^*, r_1^*, \cdots, m_\ell^*, r_\ell^*) \leftarrow \mathcal{S}^{\mathcal{V}^{U_i'}}(x)\right] \\
&+O\left(\sqrt{q^3/q^{\log q}}\right) \\
\le&O\left(\sqrt{q^3/q^{\log q}}\right) + p^{U_i'}_{\mathsf{Invalid}} \le O\left(\sqrt{q^3/q^{\log q}}\right) + p^U_{\mathsf{Invalid}}.
\end{aligned}$$

□

Claim 5.5. *The entropy on the verifier's response in $D^{\mathcal{V}^U}$ is large. More specifically, for all $i \in [\ell]$,*

$$\sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathrm{H}\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \ge \left(1 - O\left(\sqrt{q^3/q^{\log q}}\right) - p^U_{\mathsf{Invalid}}\right)(mk - \log^2 q).$$

Proof. For fixed $i \in [\ell]$ and $(m_1^*, \cdots, m_i^*) \in M^i$, let $s_{m_1^*,\cdots,m_i^*} = \mathsf{TS}_t\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right)$, by lemma 4.3 we have,

$$\mathrm{H}\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \geq (1 - s_{m_1^*,\cdots,m_i^*})\left(\log t - \log s_{m_1^*,\cdots,m_i^*}\right) \geq (1 - s_{m_1^*,\cdots,m_i^*}) \log t.$$

Plug in $t = 2^{mk - \log^2 q}$ we have

$$\begin{aligned}
&\sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathrm{H}\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \\
\geq &\sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) (1 - s_{m_1^*,\cdots,m_i^*}) \log t \\
\geq &\left(1 - O\left(\sqrt{q^3/q^{\log q}}\right) - p^U_{\mathsf{Invalid}}\right)(mk - \log^2 q).
\end{aligned}$$

□

Let r_{ij} be the response of the j-th verifier session on the i-th round.

Claim 5.6. *There exists some sessions that the verifier's response in that session is close to uniform random. More specifically, there exists $j \in [k]$, such that*

$$\sum_{i=1}^{\ell} \sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \left\| D^{\mathcal{V}^U}\left(r_{ij} \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) - U_M \right\|_1 \leq O\left(\ell \sqrt{p_{\mathsf{Invalid}} m + \frac{\log^2 q}{k}}\right)$$

where U_m is the uniform distribution over M.

Proof. By claim 5.5,

$$\begin{aligned}
&\sum_{i=1}^{\ell} \sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \sum_{j=1}^{k} \mathrm{H}\left(D^{\mathcal{V}^U}\left(r_{ij} \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \\
\geq &\sum_{i=1}^{\ell} \sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathrm{H}\left(D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \\
\geq &\left(1 - O\left(\sqrt{q^3/q^{\log q}}\right) - p_{\mathsf{Invalid}}\right)(mk - \log^2 q)\ell.
\end{aligned}$$

The first inequality follows from lemma 4.4. Thus there must exists $j \in [k]$ such that

$$\begin{aligned}
&\sum_{i=1}^{\ell} \sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\left(\begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right) \mathrm{H}\left(D^{\mathcal{V}^U}\left(r_{ij} \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)\right) \\
\geq &\left(1 - O\left(\sqrt{q^3/q^{\log q}}\right) - p_{\mathsf{Invalid}}\right)\left(m - \frac{\log^2 q}{k}\right)\ell.
\end{aligned}$$

By lemma 4.5 and concavity,

$$\begin{aligned}
&\sum_{i=1}^{\ell} \sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\begin{pmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{pmatrix} \left\| D^{\mathcal{V}^U}\left(r_{ij} \middle| \begin{matrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{matrix} \right) - U_M \right\|_1 \\
\leq &\sum_{i=1}^{\ell} \sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\begin{pmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{pmatrix} \sqrt{2m - 2\left(1 - O\left(\sqrt{q^3/q^{\log q}}\right) - p_{\mathsf{Invalid}}\right)\left(m - \frac{\log^2 q}{k}\right)} \\
\leq &O\left(\ell\sqrt{p_{\mathsf{Invalid}} m + \frac{\log^2 q}{k}}\right).
\end{aligned}$$

□

Now let's prove the theorem by constructing a BQP algorithm $\mathcal{D}$ that decides whether $x \in \mathcal{L}$ or not. Let U_{qw} be the uniform distribution over $\mathsf{Func}^{2q}_{M^{\leq \ell} \to M}$ and $\mathcal{D}$ runs $\mathcal{S}^{\mathcal{V}^{U_{\mathsf{qw}}}}(x)$, checks whether the outputting transcript is accepted by the verifier and accepts iff the verifier accepts.

Claim 5.7. *If $x \in L$, $\mathcal{D}$ accepts with probability $\geq 1 - \epsilon - \mathsf{negl} \geq 1 - \frac{1}{m\ell^5\lambda^2} - \mathsf{negl}$.*

Proof. Directly follows from the definition of the simulator. □

Claim 5.8. *If $x \notin L$, $\mathcal{D}$ accepts with probability $\leq 1 - \frac{2}{m\ell^5\lambda^2}$.*

Proof. We prove by contradiction. Suppose that $\mathcal{S}^{\mathcal{V}^{U_{\mathsf{qw}}}}(x)$ outputs an accepting transcript with probability at least $1 - \frac{2}{m\ell^5\lambda^2}$. By claim 5.6 there exists $j \in [k]$, such that

$$\sum_{i=1}^{\ell} \sum_{(m_1^*,\cdots,m_i^*)\in M^i} d^{\mathcal{V}^U}\begin{pmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{pmatrix} \left\| D^{\mathcal{V}^U}\left(r_{ij} \middle| \begin{matrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{matrix} \right) - U_M \right\|_1 \leq O\left(\ell\sqrt{p_{\mathsf{Invalid}} m + \frac{\log^2 q}{k}}\right) \leq O\left(\frac{1}{\lambda\ell^2}\right). \tag{3}$$

The second inequality follows from claim 5.3. Now we construct a malicious prover P^* that breaks the soundness of the protocol. The idea of P^* is that since it has unbounded power it can send messages according to distribution D and use session j to convince the external verifier. P^* on round i does the following:

1. Samples m_i^* according to the distribution $D^{\mathcal{V}^U}\left(m_i \middle| \begin{matrix} m_1=m_1^*, r_1=r_1^*, \\ \cdots \\ m_{i-1}=m_{i-1}^*, r_{i-1}=r_{i-1}^* \end{matrix}\right)$ and sends m_{ij}^* to the external verifier and receive r_{ij}^*.
2. Samples the rest of r_i^* according to the distribution $D^{\mathcal{V}^U}\left(r_i \middle| \begin{matrix} m_1=m_1^*, r_1=r_1^*, \\ \cdots \\ m_i=m_i^*, r_{ij}=r_{ij}^* \end{matrix}\right)$.

Claim 5.9. *If $\mathcal{S}^{\mathcal{V}^{U_{\mathsf{qw}}}}(x)$ outputs an accepting transcript with probability at least $1 - \frac{2}{m\ell^5\lambda^2}$. Then P^* convinces a honest verifier with probability at least $1 - \frac{2}{m\ell^5\lambda^2} - O\left(\frac{1}{\lambda\ell^2}\right)$.*

Proof. We prove this by hybrid argument. Consider the following hybrids:
Hybrid 0: In this hybrid, the malicious prover does exactly as above. Let p_0 denote the probability that it convinces the verifier.
Hybrid 1: In this hybrid, the malicious prover does the following on round i:

1. Samples m_i^* according to the distribution $D^{\mathcal{V}^U}\left(m_i \middle| \begin{smallmatrix} m_1=m_1^*, r_1=r_1^*, \\ \cdots \\ m_{i-1}=m_{i-1}^*, r_{i-1}=r_{i-1}^* \end{smallmatrix}\right)$ and samples r_{ij}^* according to the distribution $D^{\mathcal{V}^U}\left(r_{ij} \middle| \begin{smallmatrix} m_1=m_1^* \\ \cdots \\ m_i=m_i^* \end{smallmatrix}\right)$.

2. Samples the rest of r_i^* according to the distribution $D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^*, r_1=r_1^*, \\ \cdots \\ m_i=m_i^*, r_{ij}=r_{ij}^* \end{smallmatrix}\right)$.

Let p_1 denote the probability that $(m_1^*, r_1^*, \cdots, m_\ell^*, r_\ell^*)$ is an accepting transcript. By eq. (3), $|p_0 - p_1| \leq O\left(\frac{1}{\lambda \ell^2}\right)$.

Hybrid 2:In this hybrid, the malicious prover does the following on round i:

1. Samples m_i^* according to the distribution $D^{\mathcal{V}^U}\left(m_i \middle| \begin{smallmatrix} m_1=m_1^*, r_1=r_1^*, \\ \cdots \\ m_{i-1}=m_{i-1}^*, r_{i-1}=r_{i-1}^* \end{smallmatrix}\right)$ and samples r_{ij}^* according to the distribution $D^{\mathcal{V}^U}\left(r_{ij} \middle| \begin{smallmatrix} m_1=m_1^*, r_1=r_1^* \\ \cdots \\ r_{i-1}=r_{i-1}^*, m_i=m_i^* \end{smallmatrix}\right)$.

2. Samples the rest of r_i^* according to the distribution $D^{\mathcal{V}^U}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^*, r_1=r_1^*, \\ \cdots \\ m_i=m_i^*, r_{ij}=r_{ij}^* \end{smallmatrix}\right)$.

Let p_2 denote the probability that $(m_1^*, r_1^*, \cdots, m_\ell^*, r_\ell^*)$ is an accepting transcript. Since all r_i are independent, $p_1 = p_2$.

Hybrid 3:In this hybrid, the malicious prover does the following on round i:

1. Samples m_i^* according to the distribution $D^{\mathcal{V}^{U_{\mathsf{qw}}}}\left(m_i \middle| \begin{smallmatrix} m_1=m_1^*, r_1=r_1^*, \\ \cdots \\ m_{i-1}=m_{i-1}^*, r_{i-1}=r_{i-1}^* \end{smallmatrix}\right)$ and samples r_{ij}^* according to the distribution $D^{\mathcal{V}^{U_{\mathsf{qw}}}}\left(r_{ij} \middle| \begin{smallmatrix} m_1=m_1^*, r_1=r_1^* \\ \cdots \\ r_{i-1}=r_{i-1}^*, m_i=m_i^* \end{smallmatrix}\right)$.

2. Samples the rest of r_i^* according to the distribution $D^{\mathcal{V}^{U_{\mathsf{qw}}}}\left(r_i \middle| \begin{smallmatrix} m_1=m_1^*, r_1=r_1^*, \\ \cdots \\ m_i=m_i^*, r_{ij}=r_{ij}^* \end{smallmatrix}\right)$.

Let p_3 denote the probability that $(m_1^*, r_1^*, \cdots, m_\ell^*, r_\ell^*)$ is an accepting transcript. By lemma 4.7, $p_2 = p_3$. Note that the output distribution in this hybrid is exactly $D^{\mathcal{V}^{U_{\mathsf{qw}}}}$ thus $p_3 \geq 1 - \frac{2}{m\ell^5\lambda^2}$ by our assumption and $p_0 \geq 1 - \frac{2}{m\ell^5\lambda^2} - O\left(\frac{1}{\lambda\ell^2}\right)$.

□

By contradiction we prove the claim. □

Combine claim 5.7 and claim 5.8 we prove the theorem. □

6 Why it is Hard to Prove Impossibilities on Public-coin Unbounded-Parallel Zero-knowledge Arguments

To demonstrate the impossibility of constructing a malicious prover from a simulator that emulates the view of a random oracle verifier, consider the following argument. The key insight is that a quantum algorithm's method for extracting information from a random oracle may differ fundamentally from that of a classical algorithm.

Consider a trivial protocol where the verifier accepts any input, with ℓ rounds, and each message is in $\{0,1\}^m$. Let k denote the number of repetitions. Define $\mathbf{M}_i$ as the register storing the i-th prover's message and $\mathbf{R}$ as the register storing the verifier's response. The simulator performs the following steps:

- For each $i = 1, \ldots, \ell$:
  1. Sample two random strings $s_0, s_1 \in \{0,1\}^m$, where the first bit of s_b is b. Let U be a unitary operator that maps $|s_0\rangle$ to $|0^m\rangle$ and $|s_1\rangle$ to $|0^{m-1}1\rangle$. Prepare the state $|(m_1, \ldots, m_{i-1}, s_0)\rangle$ on $\mathbf{M}_i$ and query the verifier oracle V_U to obtain $V(m_1, \ldots, m_{i-1}, s_0)$. Then query the inverse verifier oracle $V_U^\dagger$ to reset the verifier's state.
  2. For each $j = 1, \ldots, 2km$:
     (a) Prepare the state
     $$\frac{1}{\sqrt{2}}\left(|(m_1, \ldots, m_{i-1}, s_0)\rangle_{\mathbf{M}_i} + |(m_1, \ldots, m_{i-1}, s_1)\rangle_{\mathbf{M}_i}\right)$$
     on the query register $\mathbf{M}_i$.
     (b) Query the verifier oracle U_V. The state now looks like
     $$\frac{1}{\sqrt{2}}\left(|(m_1, \ldots, m_{i-1}, s_0)\rangle_{\mathbf{M}_i} |V(m_1, \ldots, m_{i-1}, s_0)\rangle_{\mathbf{R}} + |(m_1, \ldots, m_{i-1}, s_1)\rangle_{\mathbf{M}_i} |V(m_1, \ldots, m_{i-1}, s_1)\rangle_{\mathbf{R}}\right.$$
     (c) Apply the unitary U to $\mathbf{M}_i$ and ignore all qubits except the last qubit of $\mathbf{M}_i$. The state then becomes
     $$\frac{1}{\sqrt{2}}\left(|0\rangle |V(m_1, \ldots, m_{i-1}, s_0)\rangle_{\mathbf{R}} + |1\rangle |V(m_1, \ldots, m_{i-1}, s_1)\rangle_{\mathbf{R}}\right).$$
     (d) Apply Hadamard gates to all qubits, then measure all qubits in the computational basis. This yields a random string $d_{ij} \in \{0,1\}^{mk}$ and a value $y_{ij} = \langle d_{ij}, V(m_1, \ldots, m_{i-1}, s_0) \oplus V(m_1, \ldots, m_{i-1}, s_1)\rangle$.
     (e) Prepare the state $|(m_1, \ldots, m_{i-1}, s_0)\rangle_{\mathbf{M}_i} |V(m_1, \ldots, m_{i-1}, s_0)\rangle$ and query the inverse verifier oracle $V_U^\dagger$ to reset the oracle state.

     Use the collected data d_{ij} and y_{ij} for $j = 1, \ldots, 2km$ along with $V(m_1, \ldots, m_{i-1}, s_0)$ to solve for $V(m_1, \ldots, m_{i-1}, s_1)$ by solving a system of linear equations.

Note that regardless of which query is measured, the probability of correctly identifying the prover's message is at most $\frac{1}{2}$. Consequently, for any malicious prover using the simulator in a 'black-box' manner (which measures only ℓ queries, a reasonable assumption), the probability of outputting a distribution identical to the simulator's output distribution is at most $\frac{1}{2^\ell}$.

7 Other Results: Impossibility on Straight Line Simulator

For a $\ell = \ell(\lambda)$ round protocol (2ℓ messages in total with the prover sends the first message and the verifier sends the last message) with each message in $\{0,1\}^m$. Registers are defined as follows:

1. $\mathbf{M} = \otimes_{i=1}^{\ell} \mathbf{M}_i$ where $\mathbf{M}$ is the register that the simulator stores its queries on. $\mathbf{M}_i$ stores the i-th message of the prover and has basis $\{0,1\}^m \cup \{\perp\}$.

2. $\mathbf{R}$ is the register that stores the verifier's answer and has basis $\{0,1\}^m \cup \{\perp\}$. If the verifier aborts this register will be on state $|\perp\rangle$.

3. $\mathbf{F} = \mathbf{F} \otimes \mathbf{I} \otimes \mathbf{B} \otimes \mathbf{W} \otimes \mathbf{C}$ is the verifier's private register.

   - $\mathbf{F}$ is the register that stores a 'random oracle' that decides whether the verifier aborts for each prover's query and the randomness used for the honest classical verifier. It is initialized to
   $$|\psi_\varepsilon\rangle_{\mathbf{F}} = \sum_{s \in \mathcal{S}, F \in \mathsf{Func}_{(\{0,1\}^m \cup \{\perp\})^\ell \to \{0,1\}}} \sqrt{\frac{D_\varepsilon(F)}{|S|}} |s, F\rangle_{\mathbf{F}}$$
   where $\mathcal{S}$ is the randomness space for the honest verifier and D_ε is the weight function such that the distribution of F respect to D_ε is the distribution of functions where for each x, $F(x)$ is an independent bit with ε probability to be 1 and otherwise 0. For now, ε is set to $1/\ell$.
   - $\mathbf{I}$ is the round register that stores which round the verifier is in. It is initialized to $|0\rangle_{\mathbf{I}}$.
   - $\mathbf{B}$ is the qubit register that indicates whether the verifier aborts, its is $|1\rangle$ iff the verifier aborts. It is initialized to $|0\rangle_{\mathbf{B}}$
   - $\mathbf{W}$ is the work register of the verifier. It is initialized to $|\mathbf{0}\rangle_{\mathbf{W}}$.
   - $\mathbf{C} = \otimes_{i=1}^{\ell} \mathbf{C}_i$ is the register that the verifier uses it to copy and dequantize the simulator's messages and responses. $\mathbf{C}_i$ will store the message and response on round i. It is initialized to $|\mathbf{0}\rangle_{\mathbf{C}}$.

For any verifier V^*, it can be represent as an unitary U over register $\mathbf{M} \otimes \mathbf{R} \otimes \mathbf{F}$. For simplicity, we assume that the verifier output accept/reject via setting the first qubit of the $\mathbf{R}$ to 1/0 for the last round.

Definition 7.1 (Straight Line Simulation)**.** *A post-quantum zero-knowledge protocol (P, V) for a language $\mathcal{L}$ is said to have a straight line simulator iff there exists a simulator $\mathcal{S}$ that satisfy the zero-knowledge property and $\mathcal{S}$ only queries U (without control registers) but not $U^\dagger$.*

Theorem 7.2. *If a post-quantum argument $\Pi = (P, V)$ for a language $\mathcal{L}$ is ϵ-Zero-knowledge for some $\epsilon < O\left(\frac{1}{\lambda \ell^2}\right)$ and the simulator is straightline, then $\mathcal{L} \in \mathsf{BQP}$.*

Proof. For any post-quantum zero-knowledge protocol (P, V) for a language $\mathcal{L}$ and a straight line simulator $\mathcal{S}$ for it. W.L.O.G we assume that $\mathcal{S}$ queries U for exactly ℓ times. Define a verifier V_0^* that works as follows:

Algorithm 1: V_0^*

Input: Register $\mathbf{M} \otimes \mathbf{R} \otimes \mathbf{F}$

1 Run the actual classical verifier V and xor its output on register $\mathbf{R}$ condition on $\mathbf{B}$ is $|0\rangle$. Apply bit-flip on $\mathbf{B}$ if the classical verifier aborts on this query (due to incorrect format), this query does not match previous queries (check using $\mathbf{C}$) or $F(m_1, m_2, \cdots, m_\ell) = 1$ where F is the random function on $\mathbf{F}$ and $m_1, m_2, \cdots, m_\ell$ are the queries made by the prover on $\mathbf{M}$.

2 Apply the copy unitary $U_{copy} = \sum_{i=1}^{\ell} |i\rangle\langle i|_{\mathbf{I}} \otimes U_{copy}^{(i)}$ condition on $\mathbf{B}$ is $|0\rangle$ where $U_{copy}^{(i)} : |m\rangle_{\mathbf{M}_i} |r\rangle_{\mathbf{R}} |u\rangle_{\mathbf{C}_i} \to |m\rangle_{\mathbf{M}_i} |r\rangle_{\mathbf{R}} |u \oplus (m,r)\rangle_{\mathbf{C}_i}$ copies the i-the message of the prover and the response of the verifier to the verifier's private register.

We prove the following claims that says, the output of the simulator must be 'the message that it queries the verifier oracle and its response'. If it is not the case, the simulator never knows whether the verifier aborts.

Claim 7.3. *Let p_{suc} be the probability that $\mathcal{S}$ outputs a transcript that the verifier does not abort and accepts Then $\left| p_{suc} - \left(1 - \frac{1}{\ell}\right)^{\ell} \right| \leq O\left(\frac{1}{\lambda \ell^2}\right)$. Thus $p_{suc} = \Omega(1)$ is non-negligible.*

Proof. Follows by the definition of the simulator. Note that V_0^* aborts with probability $\left(1 - \frac{1}{\ell}\right)^{\ell}$. □

Claim 7.4. *Condition on the simulator outputting a non-aborting transcript, say $(m_1, r_1, m_2, r_2, \cdots, m_\ell, r_\ell)$, let $|\phi\rangle_{\boldsymbol{F}}$ be the state of the verifier at the end of the simulation. Let*

$$\Pi_{eq} = I_{\mathsf{Tr}_{\boldsymbol{C}}(\boldsymbol{F})} \otimes_{i=1}^{\ell} |m_i, r_i\rangle\langle m_i, r_i|_{\boldsymbol{C}_i}$$

be the projector that projects onto the subspace where the output of the simulator is just its queries and response. Then we have,

$$\| \Pi_{eq} |\phi\rangle_{\boldsymbol{F}} \|_2^2 \geq 1 - O\left(\frac{1}{\lambda}\right).$$

Proof. Let

$$\Pi_i^m = I_{\mathsf{Tr}_{\mathbf{C}_i}(\mathbf{F})} \otimes \sum_{r_i} |m_i, r_i\rangle\langle m_i, r_i|_{\mathbf{C}_i}$$

and

$$\Pi_i^r = I_{\mathsf{Tr}_{\mathbf{C}_i}(\mathbf{F})} \otimes |m_i, r_i\rangle\langle m_i, r_i|_{\mathbf{C}_i}$$

be the projectors that projects onto the subspace where the i-th query/query-and-response of the output of the simulator is just its query/query-and-response of the verifier oracle. Now consider a verifier V_1^* that measures $\mathbf{F}$ at the start and measures $\mathbf{C}$ after each query.

Algorithm 2: V_1^*

Input: Register $\mathbf{M} \otimes \mathbf{R} \otimes \mathbf{F}$

1 Run the actual classical verifier V and xor its output on register $\mathbf{R}$ condition on $\mathbf{B}$ is $|0\rangle$. Apply bit-flip on $\mathbf{B}$ if the classical verifier aborts on this query (due to incorrect format), this query does not match previous queries (check using $\mathbf{C}$) or $F(m_1, m_2, \cdots, m_\ell) = 1$ where F is the random function on $\mathbf{F}$ and $m_1, m_2, \cdots, m_\ell$ are the queries made by the prover on $\mathbf{M}$.

2 Apply the copy unitary $U_{copy} = \sum_{i=1}^{\ell} |i\rangle \langle i|_{\mathbf{I}} \otimes U_{copy}^{(i)}$ condition on $\mathbf{B}$ is $|0\rangle$ where $U_{copy}^{(i)} : |m\rangle_{\mathbf{M}_i} |r\rangle_{\mathbf{R}} |u\rangle_{\mathbf{C}_i} \to |m\rangle_{\mathbf{M}_i} |r\rangle_{\mathbf{R}} |u \oplus (m, r)\rangle_{\mathbf{C}_i}$ copies the i-the message of the prover and the response of the verifier to the verifier's private register. Then perform a measurement on $\mathbf{C}_i$ under computational basis. Let the result be $(m_i^{queried}, r_i^{queried})$.

Note that this verifier cannot be represented as an unitary but since our simulator is straightline we can still consider the output of $\mathcal{S}^{V_1^*}$. Since $\mathbf{C}$ and $\mathbf{F}$ can be viewed as control register for all other operators, measurements on $\mathbf{C}_i$ and $\mathbf{F}$ commute with other unitaries and measurements performed by the simulator, the verifier oracle, Π_i^m and Π_i^r for all $i \in [\ell]$. Thus the output distributions $\mathcal{S}^{V_0^*}$ and $\mathcal{S}^{V_1^*}$ are identical. But note that in $\mathcal{S}^{V_1^*}$ all queries are classical since V_1^* measures $\mathbf{C}$. Let $m_i^{queried}, r_i^{queried}$ be the measurement outcomes for the measurement on $\mathbf{C}_i$. If there exists $i \in [n]$ such that

$$\|\Pi_i^m |\phi\rangle_{\mathbf{F}}\|_2^2 \leq 1 - \Omega\left(\frac{1}{\lambda\ell}\right).$$

Then in the execution of $S^{V_1^*}$,

$$\mathsf{Pr}\left[m_i \neq m_i^{queried} \mid S^{V_1^*} \text{ outputting a non-aborting output}\right] \geq \Omega\left(\frac{1}{\lambda\ell}\right).$$

But note that $F(m_1, m_2, \cdots, m_i)$ is information theoretically hidden from the simulator. This will cause a contradiction since the probability that $F(m_1, m_2, \cdots, m_i) = 1$ is $\frac{1}{\ell}$ independently, which means that with probability $\Omega\left(\frac{1}{\lambda\ell^2}\right)$ the simulator claims that the transcript it outputted is non-aborting but it is actually aborting. Thus for all $i \in [n]$, we have

$$\|\Pi_i^m |\phi\rangle_{\mathbf{F}}\|_2^2 > 1 - O\left(\frac{1}{\lambda\ell}\right). \tag{4}$$

If there exists $i \in [n]$ and a polynomial $p(\cdot)$ such that

$$\|\Pi_i^r |\phi\rangle_{\mathbf{F}}\|_2^2 \leq 1 - \Omega\left(\frac{1}{\lambda\ell}\right).$$

By eq. (4), we have

$$\mathsf{Pr}\left[m_i = m_i^{queried} \wedge r_i \neq r_i^{queried} \mid S^{V_1^*} \text{ outputting a non-aborting output}\right] \geq \Omega\left(\frac{1}{\lambda\ell}\right)$$

which will cause a contradiction since in $S^{V_1^*}$ the randomness for the honest verifier is measured and the response should be unique and the probability of outputting a wrong response should be at most $O\left(\frac{1}{\lambda\ell^2}\right)$ by definition. Thus for all $i \in [n]$, we have

$$\|\Pi_i^r |\phi\rangle_{\mathbf{F}}\|_2^2 > 1 - O\left(\frac{1}{\lambda\ell}\right).$$

By quantum union bound, we prove this claim. □

Now we consider a non-black-box simulator $\overline{\mathcal{S}}$ that runs $\mathcal{S}$ honestly but outputs the measurement result of $\mathbf{C}$ as the transcript (and the measurement result of $\mathbf{B}$ as the bit indicating whether the verifier aborts). By the above claim, this is a valid simulator. We will use this simulator to construct a quantum polynomial time algorithm $\mathcal{D}$ that decide whether an instance x is in $\mathcal{L}$ or not. $\mathcal{D}$ upon receiving its input instance x, initializes the verifier's private register as above and runs the simulator $\overline{\mathcal{S}}$ on input x with respect to V_1^*. This is because we only need to sample the value of $F(\cdot)$ when each classical query arrives. Finally, $\mathcal{D}$ accepts iff the transcript $(m_1^{queried}, r_1^{queried}, \cdots, m_\ell^{queried}, r_\ell^{queried})$ is both non-aborting (according to F) and accepted by the verifier.

Claim 7.5. *If $x \in \mathcal{L}$, $\mathcal{D}$ accepts with at least $p_{suc} - O\left(\frac{1}{\lambda}\right) \geq \Omega(1)$ probability.*

Proof. By claim 7.3, with probability p_{suc}, $(m_1, r_1, \cdots, m_\ell, r_\ell)$ is both non-aborting and accepted. By claim 7.4, with probability $1-O\left(\frac{1}{\lambda}\right)$, $(m_1, r_1, \cdots, m_\ell, r_\ell) = (m_1^{queried}, r_1^{queried}, \cdots, m_\ell^{queried}, r_\ell^{queried})$. By union bound we prove the claim. □

Claim 7.6. *If $x \notin \mathcal{L}$, $\mathcal{D}$ accepts with at most $O\left(\frac{1}{\lambda}\right) + \mathsf{negl}$ probability.*

Proof. We construct an adversary $\mathcal{B}$ that tries to break the soundness of the protocol with probability. $\mathcal{B}$ runs $\overline{S}$ with respect V_2^* that forwards queries to an outside honest verifier V.

Algorithm 3: V_2^*

Input: Register $\mathbf{M} \otimes \mathbf{R} \otimes \mathbf{F}$

1 Measures $\mathbf{I}$ to obtain the round number i. Measures $\mathbf{M}_i$ and forward the measurement result $m_i^{queried}$ to the external verifier V. Xor its response $r_i^{queried}$ on $\mathbf{R}$. Apply bit-flip on $\mathbf{B}$ if this query does not match previous queries (check using $\mathbf{C}$) or $F(m_1, m_2, \cdots, m_\ell) = 1$ where F is the random function on $\mathbf{F}$ and $m_1, m_2, \cdots, m_\ell$ are the queries made by the prover on $\mathbf{M}$.

2 Apply the copy unitary $U_{copy} = \sum_{i=1}^{\ell} |i\rangle\langle i|_{\mathbf{I}} \otimes U_{copy}^{(i)}$ condition on $\mathbf{B}$ is $|0\rangle$ where $U_{copy}^{(i)} : |m\rangle_{\mathbf{M}_i} |r\rangle_{\mathbf{R}} |u\rangle_{\mathbf{C}_i} \to |m\rangle_{\mathbf{M}_i} |r\rangle_{\mathbf{R}} |u \oplus (m, r)\rangle_{\mathbf{C}_i}$ copies the i-the message of the prover and the response of the verifier to the verifier's private register. Then perform a measurement on $\mathbf{C}_i$ under computational basis. Let the result be $(m_i^{queried}, r_i^{queried})$.

It is easy to see that $\mathcal{S}$ cannot tell the difference between $\overline{S}^{V_0^*}$, $\overline{S}^{V_1^*}$ and $\overline{S}^{V_2^*}$ are identical. On the other hand, by claim 7.4, with probability $1 - O\left(\frac{1}{\lambda}\right)$, the output of $\overline{S}^{V_2^*}$ is just its transcript with the external verifier V. By the soundness condition $\mathcal{B}$ should only succeed with negligible probability. By union bound, we prove the claim. □

By claim 7.5 and claim 7.6, the efficient quantum algorithm $\mathcal{D}$ decides $\mathcal{L}$, thus $\mathcal{L} \in \mathsf{BQP}$ and we are done with the proof. □

8 Other Results: Impossibility on FLS-type Public-coin Unbounded-Parallel Zero-knowledge Arguments

Definition 8.1 (FLS-type protocol [FLS99])**.** *A FLS-type protocol for* NP *language $\mathcal{L}$ is a protocol of the following form:*

- ***Public input:*** *$x \in \{0,1\}^n$ (the prover wish to prove $x \in \mathcal{L}$)*
- ***Prover's auxiliary input:*** *w (a witness for $x \in \mathcal{L}$)*
- ***Step 1: Generation Protocol:*** *The prover and the verifier first engage in a generation protocol Gen, let the transcript of this step be τ.*
- ***Step 2: WI argument:*** *The prover and the verifier then engage in a witness indistinguishable zero-knowledge argument that proves either $\tau \in \Delta$ or $x \in \mathcal{L}$ where Δ is a fixed language related to the generation protocol Gen.*

Definition 8.2 (Generation protocol)**.** *A protocol Gen between a prover P and a verifier V is called a generation protocol iff there exists a NP language Δ such that the protocol satisfies the following properties:*

- ***Soundness:*** *Let τ be the transcript of execution. If the verifier executes the protocol honestly then for **any** prover $\mathsf{Pr}[\tau \in \Delta] < \mathsf{negl}$.*
- ***Simulation of verifiers:*** *There exists a simulator $\mathcal{S}_{\mathsf{Gen}}$ that has black-box access to the verifier such that for any non-uniform QPT verifier $V^*(|\psi\rangle)$, $\mathcal{S}_{\mathsf{Gen}}$ runs for time polynomial in the running time of $V^*(|\psi\rangle)$ and $1/\epsilon$ where ϵ is the gap parameter, outputs a pair (v, σ) such that:*
  1. *Let p_d be the maximal distinguishing probability between v and the view of $V^*(|\psi\rangle)$ in the execution of Gen between an honest prover and $V^*(|\psi\rangle)$. That is, for all QPT distinguisher $\mathcal{D}$,*

$$\left| \Pr_{v \leftarrow \mathcal{S}_{\mathsf{Gen}}^{V^*(|\psi\rangle)}(x)} [\mathcal{D}(v) = 1] - \Pr_{v \leftarrow \mathsf{View}_V(P, V^*(|\psi\rangle))} [\mathcal{D}(v) = 1] \right| \leq p_d.$$

  2. *Let τ be the transcript obtained by view v. Then the probability that both $\tau \in \mathcal{L}$ and σ is a witness for it is at least $1 - \epsilon + p_d$.*

Theorem 8.3. *Let (P, V) be a post-quantum black-box weak zero-knowledge **argument** for any NP language $\mathcal{L}$, if it satisfies:*

- *It is a public-coin protocol.*
- *It remains black-box weak zero-knowledge under unbounded parallel repetition.*
- *It is an FLS-type protocol and the simulator is straightline in the WI argument step.*

Then $\mathcal{L}$ is in BQP.

Proof. Consider the malicious verifier $V_0^{\mathcal{O}}$ mentioned in proof for theorem 5.2. Let all parameters and notations be the same as what it is in the proof for theorem 5.2.

Claim 8.4. *Let τ_j be the transcript of the generation step in session j. There exists a session $j \in [k]$ such that*

$$\mathsf{Pr}[\tau_j \in \Delta] \leq O\left(\frac{1}{\lambda \ell^2}\right).$$

Proof. The claim follows from claim 5.6 and the soundness property of generation protocol. Consider a malicious prover P^* for the generation protocol that sends messages according to distribution D. More specifically, P^* on round i samples m_i according to the distribution $D\left(m_i \mid (m_1, r_1, \cdots, m_{i-1}, r_{i-1})\right.$ and sends m_{ij} to the verifier to receive r_{ij}, then samples r_i according to the distribution $D\left(r_i \mid (m_1, r_1, \cdots, m_i, r_{ij})\right)$. By the soundness property the transcript τ' between this prover and $V_0^{\mathcal{O}}$ should satisfy

$$\mathsf{Pr}\left[\tau' \in \Delta\right] < \mathsf{negl}.$$

Together with the fact that the distance between the distribution of τ and the distribution of τ' is $O\left(\frac{1}{\lambda \ell^2}\right)$ we prove the claim. □

Now we construct a black-box simulator S_{WI} that simulates the WI argument for any verifier $V^*(|\psi\rangle)$ in a straightline way. S_{WI} calls $\mathcal{S}$ with verifier $V' = V_0^{\mathcal{O}} \circ V^*(|\psi\rangle)$ which answers queries using $V_0^{\mathcal{O}}$ but substitute responses in the WI argument of session j by the response of $V^*(|\psi\rangle)$. Note that $\mathsf{Pr}\left[\tau_j \in \Delta\right] \leq O\left(\frac{1}{\lambda \ell^2}\right)$. If $\tau_j \notin \Delta$ and the verifier accepts in the WI argument then it must be the case that $x \in L$. By theorem 7.2, this means that $\mathcal{L} \in \mathsf{BQP}$. □

Corollary 8.5. *Construction 4.1 in [ACP20] is not an unbounded parallel zero-knowledge for all* NP *language with respect to black-box simulators.*

References

[ACP20] Prabhanjan Ananth, Kai-Min Chung, and Rolando L. La Placa. *On the Concurrent Composition of Quantum Zero-Knowledge*. Cryptology ePrint Archive, Paper 2020/1528. `https://eprint.iacr.org/2020/1528`. 2020. URL: `https://eprint.iacr.org/2020/1528` (cit. on p. 20).

[BDF+11] Dan Boneh, Özgür Dagdelen, Marc Fischlin, Anja Lehmann, Christian Schaffner, and Mark Zhandry. "Random Oracles in a Quantum World". In: *ASIACRYPT 2011*. Ed. by Dong Hoon Lee and Xiaoyun Wang. Vol. 7073. LNCS. Springer, Heidelberg, Dec. 2011, pp. 41–69 (cit. on p. 2).

[BR93] Mihir Bellare and Phillip Rogaway. "Random Oracles are Practical: A Paradigm for Designing Efficient Protocols". In: *CCS '93, Proceedings of the 1st ACM Conference on Computer and Communications Security, Fairfax, Virginia, USA, November 3-5, 1993*. Ed. by Dorothy E. Denning, Raymond Pyle, Ravi Ganesan, Ravi S. Sandhu, and Victoria Ashby. ACM, 1993, pp. 62–73. DOI: `10.1145/168588.168596`. URL: `https://doi.org/10.1145/168588.168596` (cit. on p. 2).

[CCLY21] Nai-Hui Chia, Kai-Min Chung, Qipeng Liu, and Takashi Yamakawa. *On the Impossibility of Post-Quantum Black-Box Zero-Knowledge in Constant Rounds*. Cryptology ePrint Archive, Paper 2021/376. `https://eprint.iacr.org/2021/376`. 2021. URL: `https://eprint.iacr.org/2021/376` (cit. on p. 3).

[DFM20] Jelle Don, Serge Fehr, and Christian Majenz. "The Measure-and-Reprogram Technique 2.0: Multi-round Fiat-Shamir and More". In: *Lecture Notes in Computer Science*. Springer International Publishing, 2020, pp. 602–631. ISBN: 9783030568771. DOI: `10.1007/978-3-030-56877-1_21`. URL: `http://dx.doi.org/10.1007/978-3-030-56877-1_21` (cit. on p. 3).

[DFMS19] Jelle Don, Serge Fehr, Christian Majenz, and Christian Schaffner. *Security of the Fiat-Shamir Transformation in the Quantum Random-Oracle Model*. Cryptology ePrint Archive, Paper 2019/190. https://eprint.iacr.org/2019/190. 2019. URL: https://eprint.iacr.org/2019/190 (cit. on p. 3).

[FLS99] Uriel Feige, Dror Lapidot, and Adi Shamir. "Multiple NonInteractive Zero Knowledge Proofs Under General Assumptions". In: *SIAM Journal on Computing* 29.1 (1999), pp. 1–28. DOI: 10.1137/S0097539792230010. eprint: https://doi.org/10.1137/S0097539792230010. URL: https://doi.org/10.1137/S0097539792230010 (cit. on p. 18).

[LZ19] Qipeng Liu and Mark Zhandry. *Revisiting Post-Quantum Fiat-Shamir*. Cryptology ePrint Archive, Paper 2019/262. https://eprint.iacr.org/2019/262. 2019. URL: https://eprint.iacr.org/2019/262 (cit. on p. 4).

[PTW11] Rafael Pass, Wei-Lung Dustin Tseng, and Douglas Wikström. "On the Composition of Public-Coin Zero-Knowledge Protocols". In: *SIAM Journal on Computing* 40.6 (2011), pp. 1529–1553. DOI: 10.1137/100811465. eprint: https://doi.org/10.1137/100811465. URL: https://doi.org/10.1137/100811465 (cit. on p. 4).

[YZ22] Takashi Yamakawa and Mark Zhandry. *Verifiable Quantum Advantage without Structure*. 2022. arXiv: 2204.02063 [quant-ph]. URL: https://arxiv.org/abs/2204.02063 (cit. on p. 2).

[Zha12a] Mark Zhandry. *How to Construct Quantum Random Functions*. Cryptology ePrint Archive, Paper 2012/182. https://eprint.iacr.org/2012/182. 2012. URL: https://eprint.iacr.org/2012/182 (cit. on p. 6).

[Zha12b] Mark Zhandry. *Secure Identity-Based Encryption in the Quantum Random Oracle Model*. Cryptology ePrint Archive, Paper 2012/076. https://eprint.iacr.org/2012/076. 2012. URL: https://eprint.iacr.org/2012/076 (cit. on p. 6).

[Zha18] Mark Zhandry. *How to Record Quantum Queries, and Applications to Quantum Indifferentiability*. Cryptology ePrint Archive, Paper 2018/276. https://eprint.iacr.org/2018/276. 2018. URL: https://eprint.iacr.org/2018/276 (cit. on p. 4).