Oracle Separations for Quantum Statistical Zero-Knowledge

Sanketh Menda[1] and John Watrous[1,2]

[1] *Institute for Quantum Computing and School of Computer Science*
University of Waterloo, Canada

[2] *Canadian Institute for Advanced Research*
Toronto, Canada

October 29, 2018

Abstract

This paper investigates the power of quantum statistical zero knowledge interactive proof systems in the relativized setting. We prove the existence of an oracle relative to which quantum statistical zero knowledge does not contain UP ∩ coUP, and we prove that quantum statistical zero knowledge does not contain UP relative to a random oracle with probability 1. As UP is trivially contained in NP relative to every oracle, it follows that there exists an oracle relative to which quantum statistical zero knowledge does not contain NP ∩ coNP, and that quantum statistical zero knowledge does not contain NP relative to a random oracle with probability 1. Our proofs of these statements rely on a bound on output state discrimination for relativized quantum circuits based on the quantum adversary method of Ambainis [Amb02], following a technique similar to one used by Ben-David and Kothari [BDK17] to prove limitations on a query complexity variant of quantum statistical zero-knowledge.

1 Introduction

Interactive proof systems, first introduced by Goldwasser, Micali, and Rackoff [GMR85, GMR89] and Babai [Bab85, BM88], form a cornerstone of complexity theory. Many variants of interactive proof systems have been studied, including quantum statistical zero-knowledge interactive proof systems [Wat02, Kob03, Wat09, HMW13, GHMW15, Che16], which are the topic of this paper.

An interactive proof system has the property of being *statistical zero-knowledge* if the prover does not "leak" statistically significant knowledge to a computationally bounded verifier on positive problem inputs. It is known that the class QSZK of decision problems having quantum statistical zero-knowledge interactive proof systems is closed under complementation and is contained in QIP(2), the class of decision problems having (not necessarily zero-knowledge) quantum interactive proof systems in which precisely two messages are exchanged between the prover and verifier [Wat02]. Unlike its classical counterpart SZK, however, it is not known if the containment of NP in QSZK has unexpected complexity theoretic consequences. (The containment of NP in SZK implies that the polynomial-time hierarchy collapses to AM [For89, AH87, BHZ87].)

In this paper we consider QSZK in a relativized setting, with the aim of proving limitations on the power of this class. We prove two results along these lines. First, we prove that there exists an oracle relative to which UP $\cap$ coUP is not contained in QSZK, where UP is a restricted variant of NP containing decision problems recognized by a polynomial-time nondeterministic Turing machine with no more than one accepting computation path on every valid input. Second, we prove that with respect to a random oracle, the class UP is not contained in QSZK with probability 1.

Our proofs make use of the positive weights quantum adversary method of Ambainis [Amb02]. The positive weights quantum adversary method is known to not always give tight bounds on quantum query complexity, see [AS04, Zha05, ŠS06], but it suffices for our needs. Ben-David and Kothari [BDK17] recently observed that the positive weights quantum adversary method can be used to prove limitations on a query complexity variant of quantum statistical zero-knowledge. Also, a related notion of state conversion in query complexity was investigated by Lee et al. [LMR$^+$11].

2 Preliminaries

In this section, we summarize relevant concepts regarding complexity theory and quantum computation, with which we assume the reader is generally familiar.

Measures of distance between quantum states

We define the *trace norm* $\| A \|_1$ of an operator A as the sum of its singular values (with no pre-factor of 1/2), and we define the *fidelity* between quantum states ρ and σ as

$$\mathrm{F}(\rho,\sigma) := \| \sqrt{\rho}\sqrt{\sigma} \|_1 \tag{1}$$

(with the right-hand side not being squared). Uhlmann's theorem [Uhl76] implies that the fidelity between two states ρ and σ is given by

$$\mathrm{F}(\rho,\sigma) = \max_{|\psi\rangle,|\phi\rangle} |\langle \psi | \phi \rangle|, \tag{2}$$

where the maximization is over all purifications $|\psi\rangle$ and $|\phi\rangle$ of ρ and σ, respectively. For quantum states ρ and σ, their trace distance and fidelity are related by the Fuchs-van de Graaf inequalities [FvdG99] as follows:

$$2 - 2\,\mathrm{F}(\rho,\sigma) \leq \|\rho - \sigma\|_1 \leq 2\sqrt{1 - \mathrm{F}(\rho,\sigma)^2}. \tag{3}$$

Quantum circuits

The results we prove in this paper are not sensitive to the specific gate set one chooses to adopt when discussing quantum circuits. Nevertheless, for the sake of simplicity and concreteness, we may assume that quantum circuits in this paper are composed of *Hadamard*, *Toffoli*, and *phase shift* gates, as well as *query gates* (discussed below).

We may also assume that the introduction of new, initialized qubits into a quantum circuit are represented by *auxiliary qubit* gates, which have no inputs and output one qubit in the state $|0\rangle$; and circuits may also include *erasure* gates, which take one qubit as input and have no outputs, effectively tracing out their input qubit. Of course these gates can be removed from a circuit, provided that a suitable number of qubits in the state $|0\rangle$ are provided as part of the input into the circuit and that the qubits that would have gone into erasure gates are traced-out once the computation is finished—and this is what we mean when we refer to a *unitary purification* of a given circuit. It is, however, convenient to view auxiliary qubit gates and erasure gates as being gates, so that we may speak of circuits that have no inputs and output some number of qubits in a possibly mixed state. We refer the reader to [AKN98] and [Wat11] for further details on quantum circuits acting on mixed states.

Oracles and relativization

Throughout the paper we denote the binary alphabet by $\Sigma = \{0,1\}$. An *oracle* is any subset $A \subseteq \Sigma^*$ of binary strings, to which membership queries are made available at unit cost. We will use the term *black box* to refer to the restriction of an oracle to strings of a single, fixed length.

With respect to a given black box $B \subseteq \Sigma^n$, the corresponding query gate K_B is the $(n+1)$-qubit unitary gate defined by the following action on the standard basis:

$$K_B|x\rangle|a\rangle = \begin{cases} |x\rangle|\neg a\rangle & \text{if } x \in B \\ |x\rangle|a\rangle & \text{if } x \notin B, \end{cases} \tag{4}$$

for all $x \in \{0,1\}^n$ and $a \in \{0,1\}$. Equivalently, one may write

$$K_B = \sum_{x \in B} |x\rangle\langle x| \otimes X + \sum_{x \in \overline{B}} |x\rangle\langle x| \otimes \mathbb{1}, \tag{5}$$

where (in this case) $\mathbb{1}$ denotes the identity operator on a single qubit and X denotes a single-qubit NOT operation.

A *relativized circuit* is one that may include query gates (accessing one black box for each string length), and one views that such a circuit queries a given oracle if the query gates are consistent with the oracle.

Honest-verifier quantum statistical zero-knowledge

With respect to quantum statistical zero-knowledge, we will focus on the *honest-verifier* definition of this class, which is simpler to state than the more cryptographically satisfying *general-verifier* definition. Although the two definitions are known to be equivalent

[Wat09], it is only the easier of the two containments needed to prove this equivalence that is relevant to our results. That is, because we prove that certain relativized languages are not contained in QSZK, no generality is lost in making use of the honest-verifier definition, as it gives rise to a complexity class that is at least as large as the one given by the general-verifier definition.

With respect to an oracle A, a language L is in QSZK^A if there exists a quantum interactive proof system (V, P) satisfying the following (somewhat informally stated) properties:

1. The verifier is *efficient*: V is specified by a polynomial-time generated family of tuples of quantum circuits that represent the verifier's actions. These circuits may make queries to the oracle A.

2. The proof system is *complete and sound*: on inputs in L, the prover P (which may also query the oracle A) causes V to accept with high probability, and on inputs not in L, no prover causes V to accept, except with small probability.

3. The proof system is *honest-verifier quantum statistical zero-knowledge*: on inputs in L, and assuming that one considers a unitary purification of V, the *view* of V (represented by the tensor product of the states it holds after each message exchange takes place) has negligible trace distance to a state that can be produced by a polynomial-time uniform family of quantum circuits that do not interact with a prover (but that may make queries to A).

For the purposes of this paper, it is not necessary for us to make use of the specific details of the definition just suggested—we instead rely on the existence of a complete promise problem for QSZK, known as QUANTUM STATE DISTINGUISHABILITY [Wat02]. A relativized version of this problem can be phrased as follows.

RELATIVIZED QUANTUM STATE DISTINGUISHABILITY (QSD^A)

Input: Relativized quantum circuits Q_0 and Q_1 that take no input qubits and produce output states on the same number of qubits. Let $\rho_0(A)$ and $\rho_1(A)$ denote the states produced by Q_0 and Q_1, respectively, when the query gates of these circuits operate in accordance with the oracle A.

Yes: (Q_0, Q_1) is a *yes-instance* of QSD^A, denoted $(Q_0, Q_1) \in \mathrm{QSD}^A_{\mathrm{yes}}$, if $\rho_0(A)$ and $\rho_1(A)$ are *far*:

$$\frac{1}{2}\left\|\rho_0(A) - \rho_1(A)\right\|_1 \geq \frac{2}{3}.$$

No: (Q_0, Q_1) is a *no-instance* of QSD^A, denoted $(Q_0, Q_1) \in \mathrm{QSD}^A_{\mathrm{no}}$, if $\rho_0(A)$ and $\rho_1(A)$ are *close*:

$$\frac{1}{2}\left\|\rho_0(A) - \rho_1(A)\right\|_1 \leq \frac{1}{3}.$$

Although the proof that QSD is complete for QSZK found in [Wat02] does not mention query gates, the proof does extend directly to the relativized setting; query gates can simply be treated in the same way as other gates within the context of this proof. The following theorem expresses this fact in a form that is convenient for the purposes of this paper. (In this theorem, Γ denotes an arbitrary alphabet over which languages are to be considered—but we will only need to concern ourselves with the unary alphabet $\Gamma = \{0\}$ in this paper.)

Theorem 1. *Let $L \subseteq \Gamma^*$ be a language and let $A \subseteq \Sigma^*$ be an oracle. The language L is contained in QSZK^A if and only if there exists a polynomial-time uniform family of pairs of relativized quantum circuits $\{(Q_0^x, Q_1^x) : x \in \Gamma^*\}$ with these properties:*

1. *If $x \in L$, then $(Q_0^x, Q_1^x) \in \mathrm{QSD}^A_{\mathrm{yes}}$, and*

2. *If $x \notin L$, then $(Q_0^x, Q_1^x) \in \mathrm{QSD}^A_{\mathrm{no}}$.*

Unambiguous polynomial-time

Finally, the class UP, which stands for *unambiguous polynomial time*, is a restricted variant of NP that was first defined by Valiant [Val76]. A language L is in the class UP if there exists a polynomial-time nondeterministic Turing machine M satisfying these conditions:

1. If $x \in L$, then M has exactly one accepting computation path on input x.

2. If $x \notin L$, then M has no accepting computation paths on input x.

Relativized variants of UP are defined in the natural way, by allowing the machine M to make oracle queries.

3 Adversary bound for output state discrimination

In this section we prove a lemma that will be used to prove that certain problems fall outside of QSZK relative to some oracles.

Before proving the main lemma, we will prove a somewhat more basic lemma that implies that a quantum circuit must, on average, make a large number of queries to a black box in order to produce output states that allow one to discriminate between an empty black box and a black box containing one string. The proof makes use of the positive weights quantum adversary method [Amb02].

Lemma 2. *Let Q be a quantum circuit that takes no input and makes T queries to a n-bit black box and let $\rho(B)$ denote the output of Q when the black box is described by $B \subseteq \Sigma^n$. It holds that*

$$\frac{1}{2^n} \sum_{x \in \Sigma^n} \mathrm{F}\big(\rho(\{x\}), \rho(\varnothing)\big) \geq 1 - \frac{2T}{2^{n/2}}. \tag{6}$$

Proof. Let R be a unitary quantum circuit that purifies Q. For a given black box B, the unitary operator corresponding to the action of R can be expressed as

$$U_T(K_B \otimes \mathbb{1})U_{T-1}(K_B \otimes \mathbb{1}) \cdots U_1(K_B \otimes \mathbb{1})U_0 \tag{7}$$

where each U_t is a unitary operator that is independent of B (and K_B represents a query gate to B as already mentioned). Let $|\psi_t(B)\rangle$ represent the state immediately after the unitary operation U_t is performed, assuming the computation begins with all qubits initialized to the $|0\rangle$ state:

$$|\psi_t(B)\rangle = U_t(K_B \otimes \mathbb{1})U_{t-1}(K_B \otimes \mathbb{1}) \cdots U_0|0\cdots 0\rangle. \tag{8}$$

Next, define a progress function

$$f(t) = \sum_{x\in\Sigma^n} \bigl|\bigl\langle \psi_t(\{x\})\bigm|\psi_t(\varnothing)\bigr\rangle\bigr| \tag{9}$$

for all $t \in \{0,\ldots,T\}$. Because R purifies Q, it holds that $|\psi_T(B)\rangle$ purifies $\rho(B)$ (for any choice of a black box B), and therefore

$$f(T) \leq \sum_{x\in\Sigma^n} \mathrm{F}\bigl(\rho(\{x\}), \rho(\varnothing)\bigr) \tag{10}$$

by the fact that the fidelity function is non-decreasing under partial tracing. It holds that $|\psi_0(\{x\})\rangle = |\psi_0(\varnothing)\rangle$, and therefore

$$f(0) = \sum_{x\in\Sigma^n} |\langle \psi_0(\{x\})|\psi_0(\varnothing)\rangle| = 2^n. \tag{11}$$

As

$$|\psi_{t+1}(B)\rangle = U_{t+1}(K_B \otimes \mathbb{1})|\psi_t(B)\rangle, \tag{12}$$

it follows that

$$\bigl|\bigl\langle \psi_{t+1}(\{x\})\bigm|\psi_{t+1}(\varnothing)\bigr\rangle\bigr| = \bigl|\bigl\langle \psi_t(\{x\})\bigm|K_{\{x\}} \otimes \mathbb{1}\bigm|\psi_t(\varnothing)\bigr\rangle\bigr|. \tag{13}$$

Making use of the expression (5), one finds that

$$\begin{aligned}&\bigl\langle \psi_t(\{x\})\bigm|K_{\{x\}} \otimes \mathbb{1}\bigm|\psi_t(\varnothing)\bigr\rangle\\ &\quad= \bigl\langle \psi_t(\{x\})\bigm||x\rangle\langle x| \otimes (X - \mathbb{1}) \otimes \mathbb{1}\bigm|\psi_t(\varnothing)\bigr\rangle + \bigl\langle \psi_t(\{x\})\bigm|\psi_t(\varnothing)\bigr\rangle.\end{aligned} \tag{14}$$

Therefore, by the Cauchy–Schwarz and triangle inequalities, and making use of the fact that $\|X - \mathbb{1}\| = 2$, one obtains

$$\begin{aligned}&\bigl|\bigl\langle \psi_t(\{x\})\bigm|K_{\{x\}} \otimes \mathbb{1}\bigm|\psi_t(\varnothing)\bigr\rangle\bigr|\\ &\quad\geq \bigl|\bigl\langle \psi_t(\{x\})\bigm|\psi_t(\varnothing)\bigr\rangle\bigr| - 2\bigl\|(\langle x| \otimes \mathbb{1} \otimes \mathbb{1})\bigl|\psi_t(\varnothing)\bigr\rangle\bigr\|.\end{aligned} \tag{15}$$

Using the Cauchy–Schwarz inequality again, it follows that

$$\begin{aligned}f(t+1) &= \sum_{x\in\Sigma^n} \bigl|\bigl\langle \psi_{t+1}(\{x\})\bigm|\psi_{t+1}(\varnothing)\bigr\rangle\bigr|\\ &\geq f(t) - 2\sum_{x\in\Sigma^n} \bigl\|(\langle x| \otimes \mathbb{1} \otimes \mathbb{1})\bigl|\psi_t(\varnothing)\bigr\rangle\bigr\|\\ &\geq f(t) - 2 \cdot 2^{n/2}.\end{aligned} \tag{16}$$

Consequently,

$$f(T) = f(0) + \sum_{t=0}^{T-1} (f(t+1) - f(t)) \geq 2^n - 2T \cdot 2^{n/2}. \tag{17}$$

Finally, using relation (10) one finds that

$$\frac{1}{2^n} \sum_{x \in \Sigma^n} \mathrm{F}\big(\rho(\{x\}), \rho(\varnothing)\big) \geq 1 - \frac{2T}{2^{n/2}} \tag{18}$$

as required. □

We now present the main lemma, which is proved through the use of Lemma 2 along with standard arguments.

Lemma 3 (Main Lemma)**.** *Let Q_0 and Q_1 be quantum circuits, both taking no input qubits, producing the same number of output qubits, and making at most T queries to an n-bit black box, and let $\rho_0(B)$ and $\rho_1(B)$ denote the output states of these circuits when the black box is described by $B \subseteq \Sigma^n$. If T and n satisfy*

$$T \leq \frac{\sqrt{2^n}}{20736}, \tag{19}$$

then there are at least $\frac{2}{3}2^n$ distinct choices of a string $x \in \Sigma^n$ such that

$$\left| \frac{1}{2} \big\| \rho_0(\{x\}) - \rho_1(\{x\}) \big\|_1 - \frac{1}{2} \big\| \rho_0(\varnothing) - \rho_1(\varnothing) \big\|_1 \right| < \frac{1}{6}. \tag{20}$$

Proof. Define sets $S_0, S_1 \subseteq \Sigma^n$ as follows:

$$\begin{aligned} S_0 &= \left\{ x \in \Sigma^n \;:\; \big\| \rho_0(\{x\}) - \rho_0(\varnothing) \big\|_1 < \frac{1}{6} \right\}, \\ S_1 &= \left\{ x \in \Sigma^n \;:\; \big\| \rho_1(\{x\}) - \rho_1(\varnothing) \big\|_1 < \frac{1}{6} \right\}. \end{aligned} \tag{21}$$

Also define $S = S_0 \cap S_1$. For every $x \in S$, it follows from the triangle inequality that

$$\begin{aligned} &\left| \frac{1}{2} \big\| \rho_0(\{x\}) - \rho_1(\{x\}) \big\|_1 - \frac{1}{2} \big\| \rho_0(\varnothing) - \rho_1(\varnothing) \big\|_1 \right| \\ &\qquad \leq \frac{1}{2} \big\| \rho_0(\{x\}) - \rho_0(\varnothing) \big\|_1 + \frac{1}{2} \big\| \rho_1(\{x\}) - \rho_1(\varnothing) \big\|_1 < \frac{1}{6}. \end{aligned} \tag{22}$$

We will now prove that

$$|S_0| \geq \frac{5}{6} 2^n \quad \text{and} \quad |S_1| \geq \frac{5}{6} 2^n. \tag{23}$$

The same argument, applied separately to Q_0 and Q_1, establishes both inequalities, so let us focus on Q_0 and prove the first inequality. By making use of the Fuchs–van de Graaf inequalities, we conclude from Lemma 2 that

$$\frac{1}{2^n} \sum_{x \in \Sigma^n} \big\| \rho_0(\{x\}) - \rho_0(\varnothing) \big\|_1 \leq 4 \sqrt{\frac{T}{2^{n/2}}}. \tag{24}$$

Considering only those strings not contained in S_0 yields

$$\frac{2^n - |S_0|}{6 \cdot 2^n} \leq \frac{1}{2^n} \sum_{x \notin S_0} \left\| \rho_0(\{x\}) - \rho_0(\varnothing) \right\|_1 \leq 4\sqrt{\frac{T}{2^{n/2}}}, \tag{25}$$

which yields the required bound given the assumptions of the lemma.

By the union bound there are at most $2^n/3$ strings that are either not contained in S_0 or not contained in S_1, which implies that $|S| \geq \frac{2}{3}2^n$, as required. □

4 Oracle separations

In this section we apply the main lemma proved in the previous section to prove the existence of oracles that establish limitations on the class QSZK.

4.1 Separating UP intersect coUP from QSZK

We begin by proving the existence of an oracle relative to which UP ∩ coUP is not contained in QSZK.

Theorem 4. *There exists an oracle A for which* $(\mathrm{UP}^A \cap \mathrm{coUP}^A) \not\subseteq \mathrm{QSZK}^A$.

The remainder of this subsection is devoted to a proof of this theorem, divided according to the main steps of the proof.

Problem specification and inclusion in UP intersect coUP

The basic idea of the proof is to consider oracles that contain exactly one string of each length along with the computational problem of determining the first bit of this unique string for a given length. Fortnow and Rogers [FR99] proved that a BQP machine requires an exponential number of queries to solve this problem, and our proof represents an extension of this argument to QSZK. In essence, our proof replaces their use of Corollary 3.4 of Bennett et al. [BBBV97] with Lemma 3.

The set of oracles under consideration is

$$\mathcal{A} = \{A \subseteq \Sigma^* : |A \cap \Sigma^n| = 1 \text{ for all } n \geq 1\}, \tag{26}$$

and, for any oracle $A \in \mathcal{A}$, we define a language $L(A)$ over the single-letter alphabet $\Gamma = \{0\}$ as

$$L(A) = \left\{0^n : \ n \geq 1 \text{ and } 1y \in A \text{ for some } y \in \Sigma^{n-1}\right\}. \tag{27}$$

Our aim is to prove that there exists an oracle $A \in \mathcal{A}$ such that

$$L(A) \in \mathrm{UP}^A \cap \mathrm{coUP}^A \tag{28}$$

but

$$L(A) \notin \mathrm{QSZK}^A. \tag{29}$$

```
On input $w = 0^n$:
    Reject if $n = 0$.
    Nondeterministically choose a string $y \in \Sigma^{n-1}$.
    If $1y \in A$ then accept else reject.
```

Figure 1: Nondeterministic decision procedure for $L(A)$.

Observe that for any oracle $A \in \mathcal{A}$, the inclusion $L(A) \in \mathrm{UP}^A$ is established by the simple nondeterministic procedure presented in Figure 1. As there is exactly one string of each positive length in an oracle $A \in \mathcal{A}$, it holds that the compliment of the language $L(A)$ is given by

$$\overline{L(A)} = \bigl\{0^n : n = 0 \text{ or } 0y \in A \text{ for some } y \in \Sigma^{n-1}\bigr\}. \tag{30}$$

A similar argument to the one just presented reveals that $\overline{L(A)} \in \mathrm{UP}^A$ for any $A \in \mathcal{A}$, and therefore $L(A) \in (\mathrm{UP}^A \cap \mathrm{coUP}^A)$ for all oracles $A \in \mathcal{A}$.

Black box separation

It remains to prove that there exists an oracle $A \in \mathcal{A}$ for which $L(A) \notin \mathrm{QSZK}^A$. This is done in two steps, the first of which is a black box separation based on the main lemma proved in the previous section.

Fix a positive integer n, and let Q_0 and Q_1 be quantum circuits that take no input and make at most T queries to an n-bit black box, and let $\rho_0(B)$ and $\rho_1(B)$ denote the output states of these circuits when the black box is described by $B \subseteq \Sigma^n$. We will say that the pair (Q_0, Q_1) is *incorrect* for a given choice of a black box $B = \{x\} \subset \Sigma^n$ if either of these conditions hold:

1. $x \in 1\,\Sigma^{n-1}$ and
$$\frac{1}{2}\bigl\|\rho_0(\{x\}) - \rho_1(\{x\})\bigr\|_1 < \frac{2}{3}. \tag{31}$$
2. $x \in 0\,\Sigma^{n-1}$ and
$$\frac{1}{2}\bigl\|\rho_0(\{x\}) - \rho_1(\{x\})\bigr\|_1 > \frac{1}{3}. \tag{32}$$

Otherwise the pair (Q_0, Q_1) is *correct* for B.

Now suppose that T and n satisfy

$$T \leq \frac{\sqrt{2^n}}{20736}, \tag{33}$$

and define

$$S = \left\{x \in \Sigma^n \;:\; \left|\frac{1}{2}\bigl\|\rho_0(\{x\}) - \rho_1(\{x\})\bigr\|_1 - \frac{1}{2}\bigl\|\rho_0(\varnothing) - \rho_1(\varnothing)\bigr\|_1\right| < \frac{1}{6}\right\}. \tag{34}$$

By Lemma 3, the set S has cardinality at least $\frac{2}{3}2^n$. If it is the case that

$$\frac{1}{2}\big\|\rho_0(\varnothing) - \rho_1(\varnothing)\big\|_1 \leq \frac{1}{2}, \tag{35}$$

then there must therefore exist at least $\frac{2}{3}2^n - \frac{1}{2}2^n = \frac{1}{6}2^n$ choices of $x \in \Sigma^n$ such that the first condition listed above holds. Similarly, if it is the case that

$$\frac{1}{2}\big\|\rho_0(\varnothing) - \rho_1(\varnothing)\big\|_1 \geq \frac{1}{2}, \tag{36}$$

then there must therefore exist at least $\frac{2}{3}2^n - \frac{1}{2}2^n = \frac{1}{6}2^n$ choices of $x \in \Sigma^n$ such that the second condition listed above holds. One of the two implicants (35) and (36) must hold, establishing that (Q_0, Q_1) is incorrect for a uniformly chosen black box $B = \{x\} \subset \Sigma^n$ with probability at least $1/6$.

Oracle existence

To prove the existence of an oracle $A \in \mathcal{A}$ for which $L(A) \not\in \text{QSZK}^A$, we use the probabilistic method, along the lines of the random oracle methodology of Bennett and Gill [BG81]. Suppose that

$$\mathcal{Q} = \big\{ \big(Q_0^n, Q_1^n\big) \,:\, n \in \mathbb{N} \big\} \tag{37}$$

is a polynomial-time uniform family of pairs of relativized quantum circuits, and consider the performance of these circuits on an oracle $A \in \mathcal{A}$ chosen uniformly—meaning that for each positive integer n, one string of length n is selected uniformly and included in A, with the random selections being independent for different choices of n.

Let $\rho_0^n(A)$ and $\rho_1^n(A)$ and denote the states output by Q_0^n and Q_1^n, respectively, when the query gates in these circuits operate in a way that is consistent with the oracle A. For a given choice of $A \in \mathcal{A}$, the pair (Q_0^n, Q_1^n) therefore incorrectly determines membership of 1^n in $L(A)$, with respect to the characterization of QSZK^A given by Theorem 1, if either of these conditions hold:

1. $A \cap \Sigma^n = \{1y\}$ for some $y \in \Sigma^{n-1}$ and
$$\frac{1}{2}\big\|\rho_0^n(A) - \rho_1^n(A)\big\|_1 < \frac{2}{3}. \tag{38}$$

2. $A \cap \Sigma^n = \{0y\}$ for some $y \in \Sigma^{n-1}$ and
$$\frac{1}{2}\big\|\rho_0^n(A) - \rho_1^n(A)\big\|_1 > \frac{1}{3}. \tag{39}$$

We will also say that $\mathcal{Q}$ is *incorrect* for $A \in \mathcal{A}$ if (Q_0^n, Q_1^n) is incorrect for A for at least one choice of a positive integer n. Our aim is to prove that $\mathcal{Q}$ is incorrect with probability 1.

A small inconvenience arises at this point, which is that the circuits Q_0^n and Q_1^n are permitted to include query gates for lengths *different* from n, and therefore the events that (Q_0^n, Q_1^n) is incorrect for different choices of n are not necessarily independent. (Of course it is evident from the definition of $L(A)$ that this possibility is not helpful for solving the problem at hand, but the point must be addressed nevertheless.) This inconvenience can be circumvented by making use of a general result of Bennett and Gill, but in the present case a simple way to proceed is to define a new family

$$\mathcal{R} = \left\{ \left(R_0^n, R_1^n\right) \,:\, n \in \mathbb{N} \right\} \tag{40}$$

of quantum circuits that is identical to $\mathcal{Q}$ except that, for each n, each of the query gates of Q_0^n and Q_1^n having size different from n are hard-coded. The hard-codings are chosen so that the probability that $\left(R_0^n, R_1^n\right)$ is incorrect for a random choice of a black box $B = \{x\} \subset \Sigma^n$ is minimized. It is evident that the probability that $\mathcal{Q}$ is incorrect is no smaller than the probability that $\mathcal{R}$ is incorrect, for a random choice of $A \in \mathcal{A}$, and we have independence among the events that $\left(R_0^n, R_1^n\right)$ is incorrect for a random choice of $A \in \mathcal{A}$ over all choices of n.

By the assumption that $\mathcal{Q}$ is polynomial-time uniform, the circuits R_0^n and R_1^n include a number of n-bit query gates that is polynomial in n. For all but finitely many choices of n, it must therefore hold that the number T of n-bit queries made by either R_0^n or R_1^n must satisfy the bound $T \leq \sqrt{2^n}/20736$. This implies that for all but finitely many choices of n, the pair $\left(R_0^n, R_1^n\right)$ is incorrect with probability at least $1/6$. By the independence of these events for different choices of n, it follows that $\mathcal{R}$, and therefore $\mathcal{Q}$, is incorrect with probability 1.

Finally, because there are countably many polynomial-time uniform families of pairs of relativized quantum circuits, there exists an oracle $A \in \mathcal{A}$ for which $L(A) \notin \text{QSZK}^A$, as this is true for a random $A \in \mathcal{A}$ with probability 1.

4.2 Random oracle separation

Next we consider the relationship between UP and QSZK relative to a *random oracle*, meaning that each individual string is included in the oracle with probability $1/2$, independent of every other string. Specifically, we prove that relative to a random oracle, UP is not contained in QSZK with probability 1. Our proof follows the methodology introduced by Beigel [Bei89], who proved various random oracle separations involving UP and its variants.

Theorem 5. *For a random oracle A, it holds that $\text{UP}^A \not\subseteq \text{QSZK}^A$ with probability 1.*

Again, the remainder of the subsection is devoted to a proof of this theorem, divided according to the main steps of the proof.

Problem specification and inclusion in UP (with probability 1)

For a given positive integer n, we define $m = \lfloor \log(n) \rfloor$ and $N = 2^m$, and for the remainder of the proof we will always consider m and N to be defined in this way, as functions of a given positive integer n.

For every positive integer n and for every $k \in \{0, \ldots, 2^{N-m}\}$, define $\mathcal{B}_k^n$ to be the set of all black boxes $B \subseteq \Sigma^n$ for which there exist precisely k distinct choices of $x \in \Sigma^{N-m}$ such that $xy0^{n-N} \in B$ for all $y \in \Sigma^m$. More succinctly,

$$\mathcal{B}_k^n = \Big\{ B \subseteq \Sigma^n \,:\, \big|\big\{ x \in \Sigma^{N-m} \,:\, x\Sigma^m 0^{n-N} \subseteq B \big\}\big| = k \Big\}. \tag{41}$$

These sets define a partition

$$\mathcal{B}_0^n \cup \mathcal{B}_1^n \cup \cdots \cup \mathcal{B}_{2^{N-m}}^n \tag{42}$$

of the set of all n-bit black boxes.

Next, define a language $L(A)$, for every oracle $A \subseteq \Sigma^*$, as follows:

$$L(A) = \big\{ 0^n \,:\, n \geq 1 \text{ and } A \cap \Sigma^n \in \mathcal{B}_1^n \big\}. \tag{43}$$

It will be proved, with respect to a random choice of A, that $L(A) \in \text{UP}^A$ with probability 1 and $L(A) \in \text{QSZK}^A$ with probability 0. The first step of the proof, which establishes that $L(A) \in \text{UP}^A$ with probability 1, is a special case of the results of Beigel [Bei89]. We include a proof, both for completeness and because the concepts and notation required for the proof are useful in the second step of the proof that concerns QSZK.

First, fix a positive integer n, consider a random choice of $B \subseteq \Sigma^n$ (where each string of length n is independently included in B with probability $1/2$), and define indicator random variables

$$Z_x = \begin{cases} 1 & \text{if } x\Sigma^m 0^{n-N} \subseteq B \\ 0 & \text{otherwise.} \end{cases} \tag{44}$$

for every string $x \in \Sigma^{N-m}$. Also define

$$Z = \sum_{x \in \Sigma^{N-m}} Z_x, \tag{45}$$

and observe that the value taken by the random variable Z corresponds to the index of the set in the partition (42) to which B belongs. It is the case that

$$\text{E}[Z_x] = 2^{-2^m} = 2^{-N} \tag{46}$$

for every $x \in \Sigma^{N-m}$, and moreover

$$\Pr(Z = 0) = \big(1 - 2^{-N}\big)^{2^{N-m}} > 1 - \frac{1}{N}, \tag{47}$$

where the inequality follows from the fact that $2^N > 2^m = N$ together with the observation that the function $k \mapsto (1 - 1/k)^k$ is strictly increasing. We also have

$$\Pr(Z = 1) = 2^{N-m} \cdot 2^{-N} \cdot \big(1 - 2^{-N}\big)^{2^{N-m}-1} > \frac{1}{N} - \frac{1}{N^2}, \tag{48}$$

and therefore

$$\Pr(Z \geq 2) < \frac{1}{N^2} \leq \frac{4}{n^2}. \tag{49}$$

```
On input $w = 0^n$:
    If $n = 0$ or $n$ is one of the finitely many values for which $A \cap \Sigma^n \notin \mathcal{B}_0^n \cup \mathcal{B}_1^n$,
    then *reject*.
    Let $m = \lfloor \log(n) \rfloor$ and $N = 2^m$.
    Nondeterministically choose a string $x \in \Sigma^{N-m}$.
    If $xy0^{n-N} \in A$ for all $y \in \Sigma^m$ then *accept*, else *reject*.
```

Figure 2: A polynomial-time nondeterministic decision procedure for $L(A)$ with either 0 or 1 accepting computation, provided that there are finitely many values of n for which $A \cap \Sigma^n \notin \mathcal{B}_0^n \cup \mathcal{B}_1^n$.

Now, the series

$$\sum_{n=1}^{\infty} \frac{4}{n^2} \tag{50}$$

converges, so it follows from the Borel–Cantelli lemma that for a random oracle $A \subseteq \Sigma^*$, with probability 1 there are at most finitely many values of n for which

$$A \cap \Sigma^n \notin \mathcal{B}_0^n \cup \mathcal{B}_1^n. \tag{51}$$

It therefore holds with probability 1 that $L(A) \in \mathrm{UP}^A$, for if there are finitely many values of n for which (51) holds, then membership in $L(A)$ can be decided through the nondeterministic decision procedure described in Figure 2.

Black box separation

It remains to prove that $L(A) \in \mathrm{QSZK}^A$ with probability 0. The first step toward proving this fact is to consider a simple way of modifying queries made by quantum circuits.

Fix a positive integer n, along with an arbitrary subset $C \subseteq \Sigma^n$, let $m = \lfloor \log(n) \rfloor$ and $N = 2^m$ as before, and suppose that $B \subseteq \Sigma^{N-m}$ is a given black box. Using a single query to B, it is possible to design a circuit (into which C may be hard-coded) that exactly simulates a query to the set

$$C \cup B\Sigma^m 0^{n-N}. \tag{52}$$

Now assume that Q is a quantum circuit that takes no inputs and makes at most T queries to an n-bit black box, and as above suppose that $C \subseteq \Sigma^n$ is a fixed subset of strings of length n and $B \subseteq \Sigma^{N-m}$ is an $(N-m)$-bit black box. By replacing each query gate of Q with the circuit suggested above, one obtains a new circuit R that makes T queries to B and produces exactly the same output as Q when run on the black box (52).

Next, suppose that Q_0 and Q_1 are two quantum circuits that take no input and make at most T queries to an n-bit black box, and let $\rho_0(D)$ and $\rho_1(D)$ denote the outputs of these circuits on a given black box $D \subseteq \Sigma^n$. If it is the case that $T \leq \sqrt{2^{N-m}}/20736$, then for an arbitrary choice of $C \subseteq \Sigma^n$, there are at least $\frac{2}{3}2^{N-m}$ distinct choices of a string $x \in \Sigma^{N-m}$ such that

$$\left| \frac{1}{2} \bigl\| \rho_0(C \cup x\Sigma^m 0^{n-N}) - \rho_1(C \cup x\Sigma^m 0^{n-N}) \bigr\|_1 \right. \\ \left. - \frac{1}{2} \bigl\| \rho_0(C) - \rho_1(C) \bigr\|_1 \right| < \frac{1}{6}. \tag{53}$$

This follows from Lemma 3, together with the observation described in the previous paragraph. That is, for the circuits R_0 and R_1 resulting from Q_0 and Q_1 together with the given choice of C by the process above, we obtain output states $\sigma_0(B)$ and $\sigma_1(B)$ (on a given black box $B \subseteq \Sigma^{N-m}$) satisfying

$$\begin{aligned} \sigma_0(\varnothing) &= \rho_0(C), \\ \sigma_1(\varnothing) &= \rho_1(C), \\ \sigma_0(\{x\}) &= \rho_0(C \cup x\Sigma^m 0^{n-N}), \\ \sigma_1(\{x\}) &= \rho_1(C \cup x\Sigma^m 0^{n-N}). \end{aligned} \tag{54}$$

Moving closer to the language $L(A)$, we may say that a pair of quantum circuits (Q_0, Q_1) that makes n-bit queries to a black box $B \subseteq \Sigma^n$ is *incorrect* for B if one of the following two statements is satisfied:

1. $B \in \mathcal{B}_0^n$ and $\bigl\| \rho_0(B) - \rho_1(B) \bigr\|_1 > 1/3$.
2. $B \in \mathcal{B}_1^n$ and $\bigl\| \rho_0(B) - \rho_1(B) \bigr\|_1 < 2/3$.

Otherwise, (Q_0, Q_1) is *correct* for B.

Now consider a random choice of a black box $B \subseteq \Sigma^n$, with each string being included in B independently with probability 1/2. Suppose (Q_0, Q_1) is correct for a δ fraction of black boxes in $\mathcal{B}_0^n$. For each $C \in \mathcal{B}_0^n$ for which (Q_0, Q_1) is correct, there are at least $\frac{2}{3}2^{N-m}$ choices of $x \in \Sigma^{N-m}$ such that (Q_0, Q_1) is incorrect for $C \cup x\Sigma^m 0^{n-N}$ by the analysis above. Each element of $B \in \mathcal{B}_1^n$ can be obtained as $B = C \cup x\Sigma^m 0^{n-N}$ for $2^N - 1$ distinct sets $C \in \mathcal{B}_0^n$, and therefore (Q_0, Q_1) is incorrect for at least

$$(1-\delta)|\mathcal{B}_0^n| + \frac{2 \cdot \delta \cdot 2^{N-m}}{3 \cdot 2^N} |\mathcal{B}_0^n| \geq \frac{2|\mathcal{B}_0^n|}{3N} \tag{55}$$

distinct choices of $B \in \mathcal{B}_0^n \cup \mathcal{B}_1^n$. For a random choice of $B \subseteq \Sigma^n$, it therefore holds that (Q_0, Q_1) is incorrect with probability at least

$$\frac{2}{3N} - \frac{2}{3N^2} \geq \frac{2}{3n} - \frac{2}{3n^2} \geq \frac{1}{3n}, \tag{56}$$

provided $n \geq 2$.

Oracle existence

Suppose that

$$\mathcal{Q} = \{(Q_0^n, Q_1^n) \,:\, n \in \mathbb{N}\} \tag{57}$$

is a polynomial-time uniform family of pairs of relativized quantum circuits, and consider the performance of these circuits on a random oracle $A \subseteq \Sigma^*$. That is, we will consider the probability that each pair (Q_0^n, Q_1^n) correctly determines membership in $L(A)$, with respect to the characterization of QSZK given by Theorem 1.

A similar issue to the one discussed in the previous subsection now arises, due to the possibility for the circuits Q_0^n and Q_1^n to make queries to A on strings of length different from n, and the same argument allows for this issue to be circumvented. That is, there must exist a family

$$\mathcal{R} = \{(R_0^n, R_1^n) \,:\, n \in \mathbb{N}\} \tag{58}$$

of quantum circuits, where R_0^n and R_1^n include a number of n-bit query gates that is polynomial in n and include no query gates for strings of other lengths, such that the probability $\mathcal{Q}$ is incorrect is at least the probability $\mathcal{R}$ is incorrect, for a random oracle A.

For all but finitely many choices of n, it must therefore hold that the number T of n-bit queries made by either R_0^n or R_1^n must satisfy the bound $T \leq \sqrt{2^{N-m}}/20736$. This implies that for all but finitely many choices of n, the pair (R_0^n, R_1^n) incorrectly determines the membership $0^n \in L(A)$ with probability at least $1/(3n)$. The series

$$\sum_{n=1}^{\infty} \frac{1}{3n} \tag{59}$$

diverges, so by the second Borel–Cantelli lemma the collection $\mathcal{R}$ fails to compute $L(A)$ for at least one input 0^n (and in fact infinitely many such inputs) with probability 1 for a random choice of $A \subseteq \Sigma^*$. The family $\mathcal{Q}$ is therefore incorrect for a random oracle with probability 1.

Finally, because there are countably many polynomial-time uniform families of pairs of relativized quantum circuits, and each family correctly decides $L(A)$ with probability 0, we have that $L(A) \notin \mathrm{QSZK}^A$ with probability 1, as required.

Acknowledgments

We thank Shalev Ben-David and Robin Kothari for discussions and for sharing an early draft of their paper with us. We also thank Rajat Mittal and Abel Molina for helpful conversations. SM thanks IIT Kanpur, where part of this work was done, for their hospitality. This work was supported by Canada's NSERC.

References

[AH87] William Aiello and Johan Håstad. Perfect zero-knowledge languages can be recognized in two rounds. In *Proceedings of the 28th Annual IEEE Symposium* *on Foundations of Computer Science*, pages 439–448. IEEE Computer Society, 1987.

[AKN98] Dorit Aharonov, Alexei Y. Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on the Theory of Computing*, pages 20–30. ACM, 1998. arXiv:quant-ph/9806029.

[Amb02] Andris Ambainis. Quantum lower bounds by quantum arguments. *Journal of Computer and System Sciences*, 64(4):750–767, 2002. arXiv:quant-ph/0002066.

[AS04] Scott Aaronson and Yaoyun Shi. Quantum lower bounds for the collision and the element distinctness problems. *Journal of the ACM*, 51(4):595–605, July 2004.

[Bab85] László Babai. Trading group theory for randomness. In *Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, pages 421–429. ACM, 1985.

[BBBV97] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, and Umesh V. Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, 1997. arXiv:quant-ph/9701001.

[BDK17] Shalev Ben-David and Robin Kothari. Quantum sabotage complexity, zero-error algorithms, and statistical zero knowledge. Manuscript, 2017.

[Bei89] Richard Beigel. On the relativized power of additional accepting paths. In *Proceedings of the 4th Annual Structure in Complexity Theory Conference*, pages 216–224. IEEE Computer Society, 1989.

[BG81] Charles H. Bennett and John Gill. Relative to a random oracle A, $\mathrm{P}^{\mathrm{A}} \neq \mathrm{NP}^{\mathrm{A}} \neq \text{co-NP}^{\mathrm{A}}$ with probability 1. *SIAM Journal on Computing*, 10(1):96–113, 1981.

[BHZ87] Ravi B. Boppana, Johan Håstad, and Stathis Zachos. Does co-NP Have Short Interactive Proofs? *Information Processing Letters*, 25(2):127–132, 1987.

[BM88] László Babai and Shlomo Moran. Arthur-Merlin Games: A randomized proof system, and a hierarchy of complexity classes. *Journal of Computer and System Sciences*, 36(2):254–276, 1988.

[Che16] Lijie Chen. A Note on Oracle Separations for BQP, 2016. Manuscript. arXiv:1605.00619.

[For89] Lance Fortnow. The complexity of perfect zero-knowledge. *Advances in Computing Research*, 5:327–343, 1989.

[FR99] Lance Fortnow and John D. Rogers. Complexity limitations on quantum computation. *Journal of Computer and System Sciences*, 59(2):240–252, 1999. arXiv:cs.CC/9811023.

[FvdG99] Christopher A. Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999. arXiv:quant-ph/9712042.

[GHMW15] Gus Gutoski, Patrick Hayden, Kevin Milner, and Mark M. Wilde. Quantum interactive proofs and the complexity of separability testing. *Theory of Computing*, 11:59–103, 2015. arXiv:1308.5788.

[GMR85] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof-systems (extended abstract). In *Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, pages 291–304. ACM, 1985.

[GMR89] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal on Computing*, 18(1):186–208, 1989.

[HMW13] Patrick Hayden, Kevin Milner, and Mark M. Wilde. Two-message quantum interactive proofs and the quantum separability problem. In *Proceedings of the 28th Conference on Computational Complexity*, pages 156–167. IEEE Computer Society, 2013. arXiv:1211.6120.

[Kob03] Hirotada Kobayashi. Non-interactive quantum perfect and statistical zero-knowledge. In *Proceedings of the 14th International Symposium Algorithms and Computation ISAAC 2003*, volume 2906 of *Lecture Notes in Computer Science*, pages 178–188. Springer, 2003. arXiv:quant-ph/0207158.

[LMR$^+$11] Troy Lee, Rajat Mittal, Ben W. Reichardt, Robert Spalek, and Mario Szegedy. Quantum query complexity of state conversion. In *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science*, pages 344–353. IEEE Computer Society, 2011. arXiv:1011.3020.

[ŠS06] Robert Špalek and Mario Szegedy. All quantum adversary methods are equivalent. *Theory of Computing*, 2(1):1–18, 2006. arXiv:quant-ph/0409116.

[Uhl76] Armin Uhlmann. The “transition probability” in the state space of a $*$-algebra. *Reports on Mathematical Physics*, 9:273–279, April 1976.

[Val76] Leslie G. Valiant. Relative complexity of checking and evaluating. *Information Processing Letters*, 5(1):20–23, 1976.

[Wat02] John Watrous. Limits on the power of quantum statistical zero-knowledge. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 459–468. IEEE Computer Society, 2002. arXiv:quant-ph/0202111.

[Wat09] John Watrous. Zero-knowledge against quantum attacks. *SIAM Journal on Computing*, 39(1):25–58, 2009. arXiv:quant-ph/0511020.

[Wat11] John Watrous. Guest column: an introduction to quantum information and quantum circuits. *SIGACT News*, 42(2):52–67, 2011.

[Zha05] Shengyu Zhang. On the power of Ambainis lower bounds. *Theoretical Computer Science*, 339(2):241 – 256, 2005. `arXiv:quant-ph/0311060`.