Limits on the power of quantum statistical zero-knowledge

John Watrous
Department of Computer Science
University of Calgary
Calgary, Alberta, Canada
jwatrous@cpsc.ucalgary.ca

January 16, 2003

Abstract

In this paper we propose a definition for honest verifier quantum statistical zero-knowledge interactive proof systems and study the resulting complexity class, which we denote $\mathrm{QSZK}_{\mathrm{HV}}$. We prove several facts regarding this class, including:

- The following problem is a complete promise problem for $\mathrm{QSZK}_{\mathrm{HV}}$: given instructions for preparing two mixed quantum states, are the states close together or far apart with respect to the trace norm? This problem is a quantum generalization of the complete promise problem of Sahai and Vadhan [34] for (classical) statistical zero-knowledge.
- Any polynomial-round honest verifier quantum statistical zero-knowledge proof system can be simulated by a two-message (i.e., one-round) honest verifier quantum statistical zero-knowledge proof system. Similarly, any polynomial-round honest verifier quantum statistical zero-knowledge proof system can be simulated by a three-message public-coin honest verifier quantum statistical zero-knowledge proof system.
- $\mathrm{QSZK}_{\mathrm{HV}}$ is closed under complement.
- $\mathrm{QSZK}_{\mathrm{HV}} \subseteq \mathrm{PSPACE}$. (At present it is not known if arbitrary quantum interactive proof systems can be simulated in PSPACE, even for one-round proof systems.)

These facts establish close connections between classical statistical zero-knowledge and our definition for quantum statistical zero-knowledge, and give some insight regarding the effect of this zero-knowledge restriction on quantum interactive proof systems. The relationship between our definition and possible definitions of general (i.e., not necessarily honest) quantum statistical zero-knowledge are also discussed.

1 Introduction

In recent years there has been an effort to better understand the potential advantages offered by computational models based on the laws of quantum physics as opposed to classical physics. Examples of such advantages include: polynomial-time quantum algorithms for factoring, computing discrete logarithms, and various believed-to-be intractable group-theoretic and number-theoretic problems [11, 22, 23, 24, 28, 35, 40]; information-theoretically secure quantum key-distribution [8, 36]; and exponentially more efficient quantum than classical communication-complexity protocols [33]. Equally important for understanding the power of quantum models are upper bounds and impossibility proofs, such as the containment of BQP in PP [2, 15], the impossibility of quantum bit commitment [27], and the existence of oracles and black-box problems relative to which quantum computers have limited power [1, 5, 6, 7, 15].

In this paper we consider the potential advantages of quantum variants of zero-knowledge proof systems. Zero-knowledge proof systems were first defined by Goldwasser, Micali, and Rackoff [20] in 1985, are have since been studied extensively in complexity theory and cryptography. Familiarity with the basics of zero-knowledge proof systems is assumed in this paper. For a recent survey on zero-knowledge, see Goldreich [16].

Several notions of zero-knowledge have been studied, but we will only consider *statistical* zero-knowledge in this paper. Moreover, we will focus on *honest verifier* statistical zero-knowledge, which means that it need only be possible for a polynomial-time simulator to approximate the view of a verifier that follows the specified protocol (as opposed to a verifier that may intentionally deviate from the specified protocol in order to gain knowledge). In the classical case it was proved by Goldreich, Sahai and Vadhan [18] that any honest verifier statistical zero-knowledge proof system can be transformed into a statistical zero-knowledge proof system against any verifier.

The class of languages having statistical zero-knowledge proof systems is denoted SZK; it is known that SZK is closed under complement [32], that SZK $\subseteq$ AM [4, 14], and that SZK has natural complete promise problems [19, 34]. Several interesting problems such as Graph Isomorphism and Quadratic Residuosity are known to be contained in SZK but are not known to be in BPP [17, 20]. For a comprehensive discussion of statistical zero-knowledge see Vadhan [38].

To our knowledge, no formal definitions for quantum zero-knowledge proof systems have previously appeared in the literature. However, the question of whether quantum information allows for an extension of the class of problems having zero-knowledge proofs has been addressed by several researchers. For example, one of the motivations for investigating the possibility of quantum bit commitment is its applicability to zero-knowledge proof systems. The primary reason for the lack of formal definitions seems to be that difficulties arise when classical definitions for zero-knowledge are translated to the quantum setting in the most straightforward ways. For a further discussion of these problems, see van de Graaf [21].

The goal of this paper is not to attempt to resolve these difficulties, or to propose a definition for quantum zero-knowledge that is intended to be satisfying from a cryptographic point of view. Rather, our goal is to study the complexity-theoretic aspects of a simple definition of quantum zero-knowledge based on the notion of an honest verifier. Our primary motives for considering this definition are:

- Our definition is likely to be weaker than any sensible definition for quantum statistical zero-knowledge when the honest verifier assumption is absent. By this we mean that, with respect to a wide range of possible definitions for general quantum statistical zero-knowledge, any general quantum statistical zero-knowledge proof system would also satisfy our honest verifier definition. Consequently, the upper bounds we prove on the power of honest verifier quantum zero-knowledge proof systems also hold for any such general verifier definition.

- We hope that by investigating simple notions of quantum zero-knowledge we are taking steps toward the study and understanding of more cryptographically meaningful formal definitions of quantum zero-knowledge.

- We are interested in the effect of zero-knowledge type restrictions on the power of quantum interactive proof systems from a purely complexity-theoretic point of view. Indeed, we are able to prove some interesting facts about quantum statistical zero-knowledge proofs that are not known to hold for arbitrary quantum interactive proofs, such as containment in PSPACE and parallelizability to two messages.

Our approach for studying a quantum variant of honest verifier statistical zero-knowledge parallels the approach of Sahai and Vadhan [34] for the classical case, which is based on the identification of a complete promise problem for the class SZK. We identify a complete promise problem for quantum statistical zero-knowledge that generalizes Sahai and Vadhan's complete promise problem to the quantum setting. The problem, which we call the Quantum State Distinguishability problem, may be informally stated as follows: given instructions for preparing two mixed quantum states, are the states close together or far apart with respect to the trace norm? The trace norm is an extension of the total variation or statistical difference norm to quantum states, and gives a natural way of measuring distances between quantum states. By instructions for preparing a mixed quantum state we mean the description of a quantum circuit that produces the mixed state on some specified subset of its qubits, assuming all qubits are initially in the $|0\rangle$ state. The promise in this promise problem guarantees that the two mixed states given are indeed either close together or far apart.

Several facts about quantum statistical zero-knowledge proof systems and the resulting complexity class, which we denote $\mathrm{QSZK}_{\mathrm{HV}}$, may be derived from the completeness of this problem. In particular, we prove that $\mathrm{QSZK}_{\mathrm{HV}}$ is closed under complement, that $\mathrm{QSZK}_{\mathrm{HV}} \subseteq \mathrm{PSPACE}$ (which is not known to hold for quantum interactive proof systems if the zero-knowledge condition is dropped, even in the case of one-round proof systems), and that any honest verifier quantum statistical zero-knowledge proof system can be parallelized to a one-round honest verifier quantum statistical zero-knowledge proof system. The one-round proof system may be taken to have exponentially small completeness and soundness error, or alternately may be taken to be a proof system in which the prover sends only one qubit to the verifier (in order to achieve completeness and soundness error exponentially close to 0 and 1/2, respectively). We also prove that any problem in $\mathrm{QSZK}_{\mathrm{HV}}$ has a three-message *public-coin* honest-verifier statistical zero-knowledge proof system, wherein the *verifier* needs to send only a single coin-flip to the prover (again, in order to achieve completeness and soundness error exponentially close to 0 and 1/2, respectively).

The remainder of the paper is organized as follows. In Section 2 we discuss relevant background information for the paper, including a brief discussion of the quantum formalism, quantum circuits, and quantum interactive proof systems. In Section 3 we give our definition for honest verifier quantum statistical zero-knowledge and for the Quantum State Distinguishability promise problem. In Section 4 we describe honest verifier quantum statistical zero-knowledge protocols for Quantum State Distinguishability and its complement, and in Section 5 we prove that any promise problem having an honest verifier quantum statistical zero-knowledge proof system Karp-reduces to Quantum State Distinguishability. In Section 6 we discuss some consequences of the completeness of Quantum State Distinguishability for $\mathrm{QSZK}_{\mathrm{HV}}$. We conclude with Section 7, which mentions some open problems relating to this paper.

2 Preliminaries

The purpose of this section is to review relevant background information for this paper, including a discussion of some useful aspects of the quantum formalism, quantum circuits, and quantum interactive proof systems. With the exception of quantum interactive proof systems this is done primarily in order to make clear the notation that we will use and to provide references containing more detailed discussions of these topics. The discussion of quantum interactive proof systems is intended to provide a suitable introduction for readers already familiar with the study of (classical) interactive proof systems.

2.1 The quantum formalism

Detailed discussions of the quantum formalism can be found in Nielsen and Chuang [31] and Kitaev, Shen and Vyalyi [25].

Recall that a *pure quantum state* of an n-qubit quantum system can be represented as a unit vector in the Hilbert space $\mathcal{H}$ that consists of all linear mappings from $\{0,1\}^n$ to the complex numbers. Corresponding to each pure state $|\psi\rangle \in \mathcal{H}$ is a linear functional $\langle\psi|$ that maps each vector $|\phi\rangle$ to the inner product $\langle\psi|\phi\rangle$ (conjugate-linear in the first argument). A *mixed state* of a quantum system is a state that may be described by a distribution on (not necessarily orthogonal) pure states. A collection $\{(p_k, |\psi_k\rangle)\}$ such that $0 \leq p_k$, $\sum_k p_k = 1$, and each $|\psi_k\rangle$ is a pure state is called a *mixture*: for each k, the system is in state $|\psi_k\rangle$ with probability p_k. For a given mixture $\{(p_k, |\psi_k\rangle)\}$, we associate a *density matrix* ρ having operator representation $\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$. Necessary and sufficient conditions for a given matrix ρ to be a density matrix (i.e., to represent some mixed state) are (i) ρ must be positive semidefinite, and (ii) ρ must have unit trace. Two mixtures can be distinguished (in a statistical sense) if and only if they yield different density matrices, and for this reason we interpret a given density matrix ρ as being a canonical representation of a given mixed state.

For a given Hilbert space $\mathcal{H}$, let $\mathbf{L}(\mathcal{H})$ denote the set of linear operators on $\mathcal{H}$, let $\mathbf{D}(\mathcal{H})$ denote the set of positive semidefinite operators on $\mathcal{H}$ having unit trace (so that $\mathbf{D}(\mathcal{H})$ may be identified with the set of mixed states of a given system), let $\mathbf{U}(\mathcal{H})$ denote the set of unitary operators on $\mathcal{H}$, and let $\mathbf{P}(\mathcal{H})$ denote the set of projection operators on $\mathcal{H}$. (Here, and throughout this paper, whenever we refer to some Hilbert space it is implicitly assumed that the Hilbert space is finite dimensional.)

Given Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, we define a mapping $\operatorname{tr}_{\mathcal{K}} : \mathbf{L}(\mathcal{H} \otimes \mathcal{K}) \to \mathbf{L}(\mathcal{H})$ as follows:

$$\operatorname{tr}_{\mathcal{K}} X = \sum_{j=1}^{n} (I \otimes \langle e_j|) X (I \otimes |e_j\rangle),$$

where $\{|e_1\rangle, \ldots, |e_n\rangle\}$ is any orthonormal basis of $\mathcal{K}$. This mapping is known as the *partial trace*, and has the following intuitive meaning: given a mixed state $\rho \in \mathbf{D}(\mathcal{H} \otimes \mathcal{K})$ of a bipartite system (meaning that the first part of the system corresponds to $\mathcal{H}$ and the second part to $\mathcal{K}$), $\operatorname{tr}_{\mathcal{K}} \rho$ is the mixed state of the first part of the system obtained by discarding (or simply not considering) the second part of the system. To say that a particular part of a quantum system is *traced out* means that the partial trace is performed, removing this part of the system from consideration.

A *purification* of a mixed state $\rho \in \mathbf{D}(\mathcal{H})$ is any pure state $|\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ for some Hilbert space $\mathcal{K}$ such that $\operatorname{tr}_{\mathcal{K}} |\psi\rangle\langle\psi| = \rho$. Such a purification always exists provided $\dim(\mathcal{K}) \geq \operatorname{rank}(\rho)$. The following theorem concerning purifications is critical for the study of quantum interactive proof systems. A proof may be found in Section 2.5 of Nielsen and Chuang [31].

Theorem 1 *Let $|\phi\rangle, |\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ satisfy $\operatorname{tr}_{\mathcal{K}} |\phi\rangle\langle\phi| = \operatorname{tr}_{\mathcal{K}} |\psi\rangle\langle\psi| = \rho$ for some $\rho \in \mathbf{D}(\mathcal{H})$. Then there exists $U \in \mathbf{U}(\mathcal{K})$ such that $(I \otimes U)|\phi\rangle = |\psi\rangle$.*

For $X \in \mathbf{L}(\mathcal{H})$ define

$$\|X\|_{\mathrm{tr}} = \frac{1}{2} \operatorname{tr} \sqrt{X^\dagger X}.$$

(For given positive semidefinite matrix A, recall that $\sqrt{A}$ is the unique positive semidefinite matrix that squares to give A.) Alternately, $\|X\|_{\mathrm{tr}}$ may be defined as one-half of the sum of the singular values of X. The function $\|\cdot\|_{\mathrm{tr}}$ is a norm called the *trace norm*, and generalizes the norm induced by the statistical difference or total variation distance (i.e., one-half the ℓ_1 norm). The trace norm gives a useful way of measuring distances between mixed states. For any $\rho, \xi \in \mathbf{D}(\mathcal{H})$ we have $\|\rho - \xi\|_{\mathrm{tr}} = \max_A \operatorname{tr} A(\rho - \xi)$, where the maximum is over all positive semidefinite $A \in \mathbf{L}(\mathcal{H})$ with $\|A\| \leq 1$. This maximum value is always achieved by some projection $A \in \mathbf{P}(\mathcal{H})$.

Another way of discussing differences between mixed states, which turns out to be very useful for studying quantum interactive proof systems, is given by the *fidelity.* Given two positive semidefinite operators $X, Y \in \mathbf{L}(\mathcal{H})$, define the fidelity of Y and Y by

$$F(X, Y) = \operatorname{tr} \sqrt{\sqrt{X}\, Y \sqrt{X}}.$$

If ρ and ξ are identical, then $F(\rho, \xi) = 1$, while if ρ and ξ are perfectly distinguishable then $F(\rho, \xi) = 0$. The fidelity does not, however, immediately give rise to a proper notion of distance, but it is often more convenient to work with than the trace norm. It follows easily from the definition that the fidelity is multiplicative with respect to tensor products: for any $\rho_1, \xi_1 \in \mathbf{D}(\mathcal{H})$ and $\rho_2, \xi_2 \in \mathbf{D}(\mathcal{K})$ we have $F(\rho_1 \otimes \rho_2, \xi_1 \otimes \xi_2) = F(\rho_1, \xi_1)\, F(\rho_2, \xi_2)$.

An alternate description of the fidelity is given by Uhlmann's Theorem:

Theorem 2 (Uhlmann's Theorem) *Let $\rho, \xi \in \mathbf{D}(\mathcal{H})$, and let $\mathcal{K}$ be such that there exist purifications $|\phi_0\rangle, |\psi_0\rangle \in \mathcal{H} \otimes \mathcal{K}$ of ρ and ξ, respectively (i.e., $\operatorname{tr}_{\mathcal{K}} |\phi_0\rangle\langle\phi_0| = \rho$ and $\operatorname{tr}_{\mathcal{K}} |\psi_0\rangle\langle\psi_0| = \xi$). Then*

$$F(\rho, \xi) = \max_{|\phi\rangle, |\psi\rangle} |\langle\phi|\psi\rangle|,$$

where the maximum is over all purifications $|\phi\rangle, |\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ of ρ and ξ, respectively.

A proof of this theorem can be found in Section 9.2.2 of Nielsen and Chuang [31].

Next, we mention an inequality concerning the fidelity that will be useful later.

Lemma 3 *For any $\rho, \xi, \sigma \in \mathbf{D}(\mathcal{H})$, we have $F(\rho, \sigma)^2 + F(\sigma, \xi)^2 \leq 1 + F(\rho, \xi)$.*

Proofs of this lemma appear in Refs. [29, 37].

Finally, the following theorem gives a useful relation between the trace norm and the fidelity that will be used several times. A proof may be found in Section 9.2.3 of Nielsen and Chuang [31].

Theorem 4 *For all $\rho, \xi \in \mathbf{D}(\mathcal{H})$ we have*

$$1 - F(\rho, \xi) \leq \|\rho - \xi\|_{\mathrm{tr}} \leq \sqrt{1 - F(\rho, \xi)^2}.$$

Additional facts concerning the trace norm and the fidelity will be discussed later as they are needed.

2.2 Quantum circuits

Quantum interactive proof systems are based on the quantum circuit model. Although this model can be defined in a very general way that allows non-unitary quantum operations (such as measurements) to be performed by quantum circuits, we will restrict our attention to the simpler and more common model where only unitary gates are permitted. Nielsen and Chuang [31] contains a detailed discussion of the unitary quantum circuit model, while a discussion of the more general model (including a proof that the two models are equivalent in power) can be found in Aharonov, Kitaev, and Nisan [3].

Although the unitary quantum circuit model has been shown to be equivalent in power to the more general quantum circuit model, we hasten to add that our definition for honest verifier quantum statistical zero-knowledge is based on the unitary quantum circuit model. We do not have a proof that our definition is insensitive to the difference between unitary and non-unitary quantum circuits, and our proof of the hardness of the Quantum State Distinguishability problem for $\mathrm{QSZK_{HV}}$ relies on the assumption that the verifier only applies unitary transformations.

We use the following notion of a uniform family of quantum circuits. A family $\{Q_x\}$ of quantum circuits is said to be *polynomial-time uniformly generated* if there exists a deterministic procedure that, on input x, outputs a description of Q_x and runs in time polynomial in x. It is assumed that the number of gates in any circuit is not more than the length of that circuit's description (i.e., no compact descriptions of large circuits are allowed), so that Q_x must have size polynomial in $|x|$. We also assume that quantum circuits are composed of gates from some reasonable, universal, finite set of (unitary) gates. By "reasonable" we mean, for instance, that gates cannot be defined by matrices with non-computable, or difficult to compute, entries. In fact, it will be helpful later to use the fact that any quantum circuit composed of gates from any reasonable set of basis gates can be efficiently simulated by a quantum circuit consisting only of gates from a finite collection whose corresponding matrices have only entries with rational real and imaginary parts. (See, for instance, Section 4.5.3 in Nielsen and Chuang.) It should be noted that our notion of uniformity is somewhat nonstandard, since we allow an input x to be given to the procedure generating the circuits rather than just $|x|$ written in unary (with x given as input to the circuit itself). This does not change the computational power for the resulting class of quantum circuits, however, and we find that this notion is more convenient than the usual notion of uniformity.

2.3 Quantum interactive proofs

Quantum interactive proofs were defined and studied in Refs. [26, 39]. As in the classical case, a quantum interactive proof system consists of two parties, a prover with unlimited computation power and a computationally bounded verifier. Quantum interactive proofs differ from classical interactive proofs in that the prover and verifier may send and process quantum information.

Formally, a quantum verifier is a polynomial-time computable mapping V where, for each input string x, $V(x)$ is interpreted as an encoding of a $k(|x|)$-tuple $(V(x)_1, \ldots, V(x)_{k(|x|)})$ of quantum circuits. These circuits represent the actions of the verifier at the different stages of the protocol, and are assumed to obey the properties of polynomial-time uniformly generated quantum circuits as discussed in the previous section. The qubits upon which each circuit $V(x)_j$ acts are divided into two sets: $q_{\mathcal{V}}(|x|)$ qubits that are private to the verifier and $q_{\mathcal{M}}(|x|)$ qubits that represent the communication channel between the prover and verifier. The Hilbert space corresponding to the $q_{\mathcal{V}}(|x|)$ qubits representing the verifier's private memory is denoted $\mathcal{V}$ and the space corresponding

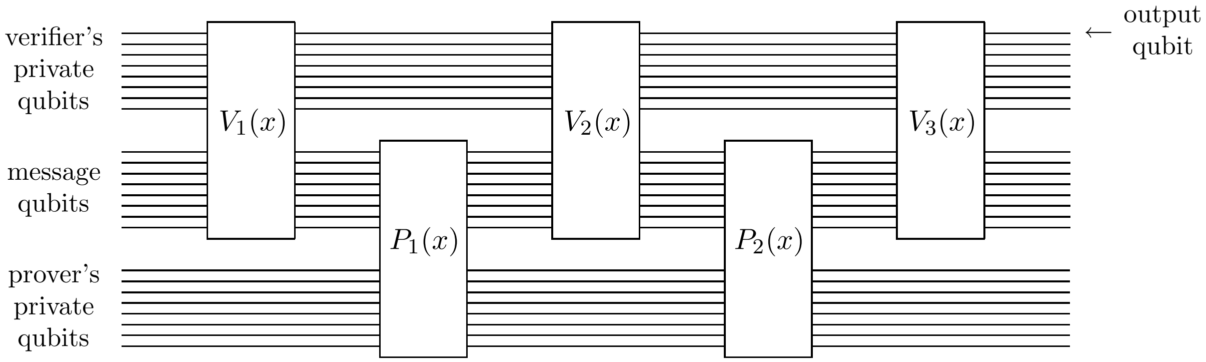

Figure 1: Quantum circuit for a 4-message quantum interactive proof system

to the communication channel is denoted $\mathcal{M}$. One of the verifier's private qubits is designated as the output qubit, which indicates whether the verifier accepts or rejects.

A quantum prover P is a function mapping each input x to an $l(|x|)$-tuple $(P(x)_1, \ldots, P(x)_{l(|x|)})$ of quantum circuits. Each of these circuits acts on $q_{\mathcal{M}}(|x|) + q_{\mathcal{P}}(|x|)$ qubits: $q_{\mathcal{P}}(|x|)$ qubits that are private to the prover and $q_{\mathcal{M}}(|x|)$ qubits representing the communication channel. The Hilbert space corresponding to the prover's private qubits is denoted $\mathcal{P}$. Unlike the verifier, no restrictions are placed on the complexity of the mapping P, the gates from which each $P(x)_j$ is composed, or on the size of each $P(x)_j$, so in general we may simply view each $P(x)_j$ as an arbitrary unitary transformation.

A verifier V and a prover P are compatible if for all inputs x we have (i) each $V(x)_i$ and $P(x)_j$ agree on the number $q_{\mathcal{M}}(|x|)$ of message qubits upon which they act, and (ii) $k(|x|) = \lfloor m(|x|)/2+1 \rfloor$ and $l(|x|) = \lfloor m(|x|)/2 + 1/2 \rfloor$ for some $m(|x|)$ (representing the number of messages exchanged). We say that V is an m-message verifier and P is an m-message prover in this case. Whenever we discuss an interaction between a prover and verifier, we naturally assume they are compatible.

Given a verifier V, a prover P, and an input x, we define a quantum circuit $(V(x), P(x))$ acting on $q(|x|) = q_{\mathcal{V}}(|x|) + q_{\mathcal{M}}(|x|) + q_{\mathcal{P}}(|x|)$ qubits as follows. If $m(|x|)$ is even, circuits

$$V(x)_1,\ P(x)_1,\ \ldots,\ P(x)_{m(|x|)/2},\ V(x)_{m(|x|)/2+1}$$

are applied in sequence, each to the $q_{\mathcal{V}}(|x|) + q_{\mathcal{M}}(|x|)$ verifier/message qubits or to the $q_{\mathcal{M}}(|x|) + q_{\mathcal{P}}(|x|)$ message/prover qubits accordingly. This situation is illustrated in Figure 1 for the case $m(|x|) = 4$. If $m(|x|)$ is odd the situation is similar, except that the prover applies the first circuit, so circuits $P(x)_1,\ V(x)_1,\ \ldots,\ P(x)_{(m(|x|)+1)/2},\ V(x)_{(m(|x|)+1)/2}$ are applied in sequence. Thus, it is assumed that the prover always sends the last message (since there would be no point for the verifier to send a message without a response). For readability we generally drop the arguments x and $|x|$ in the various functions/circuits described above when it is understood (e.g., we write V_j and P_j to denote $V(x)_j$ and $P(x)_j$ for each j, and we write m to denote $m(|x|)$).

At a given instant, the state of the qubits in the circuit (V, P) is a unit vector in the space $\mathcal{V} \otimes \mathcal{M} \otimes \mathcal{P}$. Throughout this paper, we assume that operators acting on subsystems of a given system are extended to the entire system by tensoring with the identity. For instance, for a 4-message proof system as illustrated in Figure 1, the state of the system after all circuits have been applied is $V_3\, P_2\, V_2\, P_1\, V_1\, |0^q\rangle$.

Now, for a given input x, the probability that the pair (V, P) accepts x is defined to be the probability that an observation of the verifier's output qubit (in the $\{|0\rangle, |1\rangle\}$ basis) yields the value 1, after the circuit $(V(x), P(x))$ is applied to a collection of $q(|x|)$ qubits each initially in the $|0\rangle$ state. Formally, let $\Pi_{\mathrm{acc}} \in \mathbf{P}(\mathcal{V} \otimes \mathcal{M})$ denote the projection onto states of the verifier's qubits and message qubits[1] for which the accept qubit is set to 1. Thus, the probability of acceptance for the proof system in Figure 1 is $\|\Pi_{\mathrm{acc}} V_3\, P_2\, V_2\, P_1\, V_1\, |0^q\rangle\|^2$. Define $\mathit{max\_accept}(V(x))$ (the maximum acceptance probability of $V(x)$) to be the probability that (V, P) accepts x maximized over all possible m-message provers P.

A language A is said to have an m-message quantum interactive proof system with completeness error ε_c and soundness error ε_s, where ε_c and ε_s may be functions of the input length, if the exists an m-message verifier V such that (i) if $x \in A$ then $\mathit{max\_accept}(V(x)) \geq 1 - \varepsilon_c(|x|)$, and (ii) if $x \notin A$ then $\mathit{max\_accept}(V(x)) \leq \varepsilon_s(|x|)$. We also say that (V, P) is a quantum interactive proof system for A with completeness error ε_c and soundness error ε_s if V satisfies these properties and P is a prover that succeeds in convincing V to accept with probability at least $1 - \varepsilon_c(|x|)$ when $x \in A$.

Various facts concerning quantum interactive proofs are proved in Kitaev and Watrous [26]. In particular, if we let QIP(m) denote the class of languages having quantum interactive proof systems with m messages in total, then PSPACE $\subseteq$ QIP(3) $=$ QIP(*poly*) $\subseteq$ EXP. Quantum interactive proof systems are robust with respect to completeness and soundness errors.

3 Definitions

In this section we give our definition for honest verifier quantum statistical zero-knowledge and discuss the relation of this definition to the classical definition. We also define the Quantum State Distinguishability problem, which is discussed in later sections.

3.1 Honest verifier quantum statistical zero-knowledge

In the classical case, the zero-knowledge property concerns the distribution of possible conversations between the prover and verifier from the verifier's point of view. In the quantum case, we cannot consider the verifier's view of the *entire* interaction in terms of a single quantum state in any physically meaningful way, so instead we consider the mixed quantum state of the verifier's private qubits together with the message qubits at various times during the protocol. This gives a reasonably natural way of characterizing the verifier's view of the interaction.

It will be sufficient to consider the verifier's view after each message is sent (since the verifier's views at all other times are easily obtained from the views after each message is sent by running the verifier's circuits). The zero-knowledge property will be that the mixed states representing the verifier's view after each message is sent should be approximable to within negligible trace distance by a polynomial-size (uniformly generated) quantum circuit on accepted inputs. We formalize this notion as follows.

First, given a collection $\{\rho_y\}$ of mixed quantum states, we say that the collection is *polynomial-time preparable* if there exists a polynomial-time uniformly generated family $\{Q_y\}$ of quantum

[1] It will be convenient later to have Π_{acc} defined as a projection on $\mathcal{V} \otimes \mathcal{M}$ rather than $\mathcal{V} \otimes \mathcal{M} \otimes \mathcal{P}$. Since we implicitly tensor with the identity on $\mathcal{P}$ when we want to view any operator on $\mathcal{V} \otimes \mathcal{M}$ as acting on $\mathcal{V} \otimes \mathcal{M} \otimes \mathcal{P}$, there is nothing lost in defining Π_{acc} on $\mathcal{V} \otimes \mathcal{M}$ in this way.

circuits, each having a specified collection of output qubits, such that the following holds. For each y, the state ρ_y is the mixed state obtained by running Q_y with all input qubits initialized to the $|0\rangle$ state and then tracing out all non-output qubits. By $\{Q_y\}$ polynomial-time uniform, we mean that there exists a polynomial time deterministic Turing machine that, on input y, outputs a description of Q_y.

Next, given a verifier V and a prover P, we define $\mathrm{view}_{V,P}(x, j)$ to be the reduced state of the verifier and message qubits after j messages have been sent during an execution of the proof system (V, P) on input x. In other words, this is the state obtained by taking the state of the entire system immediately after j messages have been sent and tracing out the prover's private qubits.

Finally, given a verifier V and a prover P, we say that the pair (V, P) is an *honest verifier quantum statistical zero-knowledge proof system* for a language A if

1. (V, P) is a quantum interactive proof system for A, and
2. there exists a polynomial-time preparable set $\{\sigma_{x,j}\}$ and a negligible function δ such that

   $$\|\sigma_{x,j} - \mathrm{view}_{V,P}(x, j)\|_{\mathrm{tr}} < \delta(|x|)$$

   for every $x \in A$ and each message number j.

As usual, δ negligible means that for every polynomial p, $\delta(n) < 1/p(n)$ for sufficiently large n. The polynomial-time preparable set $\{\sigma_{x,j}\}$ corresponds to the output of a polynomial-time simulator. The completeness and soundness error of an honest verifier quantum statistical zero-knowledge proof system are determined by the underlying proof system.

Finally we define $\mathrm{QSZK}_{\mathrm{HV}}$ (honest verifier quantum statistical zero-knowledge) to be the class of languages having honest verifier quantum statistical zero-knowledge proof systems with completeness and soundness error at most $1/3$. We note that sequential repetition of honest verifier quantum statistical zero-knowledge proof systems reduces completeness and soundness error exponentially while preserving the zero-knowledge property. Thus, we may equivalently define $\mathrm{QSZK}_{\mathrm{HV}}$ to be the class of languages having honest verifier quantum statistical zero-knowledge proof systems with completeness and soundness error at most $2^{-p(n)}$ for any chosen polynomial p.

3.2 Notes on the definition

Aside from the obvious difference of quantum vs. classical information, our definition differs from the standard definition for classical honest-verifier statistical zero-knowledge in the following sense. In the classical case, the simulator randomly outputs a transcript representing the *entire interaction* between the prover and verifier, while our definition requires only that the view of the verifier *at each instant* can be approximated by a simulator. The main reason for this difference is that the notion of a transcript of a quantum interaction is counter to the nature of quantum information—in general, there is no physically meaningful way to define a transcript of a quantum interaction. For instance, if a verifier tries to copy the value of every qubit it touches during the protocol in order to produce such a transcript, this would be tantamount to the verifier performing measurements that would likely spoil the properties of the protocol.

Thus, our definition is not a direct quantum analogue of the standard classical definition, and it is not clear how to give such a direct analogue that is physically meaningful. However, rather than trying to give a direct quantum analogue of the classical definition, our aim has been to provide a definition that (i) is weaker than as wide a range of definitions for general quantum statistical zero-knowledge as possible (for the purposes of proving upper bounds), (ii) satisfies the intuitive notion of honest verifier statistical zero-knowledge, and (iii) is as simple as possible. While our definition is not a direct analogue of the classical definition, our results do suggest that our definition yields a complexity class that is a natural quantum analogue of classical statistical zero-knowledge.

Finally, we reiterate at this point that our definition requires the verifier's actions to be described by unitary transformations rather than arbitrary quantum transformations. The condition that a simulator must only approximate the view of the verifier at each instant is, in this sense, not as weak as it might be for the general transformation case. From the point of view that our definition should be weaker than any reasonable notion of general verifier quantum statistical zero-knowledge, the restriction to unitary verifiers is not unnatural (since, at the very least, a cheating verifier could simulate the honest verifier's non-unitary transformations with unitary transformations).

3.3 Quantum state distinguishability

Recall that a promise problem consists of two disjoint sets $A_{\text{yes}}, A_{\text{no}} \subseteq \Sigma^*$. The associated computational task is as follows: we are given some $x \in A_{\text{yes}} \cup A_{\text{no}}$, and the goal is to accept if $x \in A_{\text{yes}}$ and to reject if $x \in A_{\text{no}}$. Thus, the input is *promised* to be an element of $A_{\text{yes}} \cup A_{\text{no}}$, and no requirement is made in case the input string is not in $A_{\text{yes}} \cup A_{\text{no}}$. Ordinary decision problems are a special case of promise problem where $A_{\text{yes}} \cup A_{\text{no}} = \Sigma^*$. See Even, Selman, and Yacobi [13] for further information on promise problems. Our above definition for QSZK_{HV} is stated in terms of decision problems, but may be rephrased in terms of promise problems in the straightforward way.

We will focus on the following promise problem, which is parameterized by constants α and β satisfying $0 \leq \alpha < \beta \leq 1$. (We consider only a restricted version of this problem where $\alpha < \beta^2$.)

(α, β)-Quantum State Distinguishability ((α, β)-QSD)

Input: Quantum circuits Q_0 and Q_1 acting on m qubits and having k specified output qubits.

Promise: Let ρ_i denote the mixed state obtained by running Q_i on state $|0^m\rangle$ and discarding (tracing out) the non-output qubits. Then either $\|\rho_0 - \rho_1\|_{\text{tr}} \geq \beta$ or $\|\rho_0 - \rho_1\|_{\text{tr}} \leq \alpha$.

Output: Accept when $\|\rho_0 - \rho_1\|_{\text{tr}} \geq \beta$, and reject when $\|\rho_0 - \rho_1\|_{\text{tr}} \leq \alpha$.

4 Protocols for state distinguishability

In this section we give honest-verifier quantum statistical zero-knowledge protocols for the (α, β)-QSD problem and its complement. As a consequence, we have that (α, β)-QSD and its complement are in QSZK_{HV} provided that α and β are constants satisfying $\alpha < \beta^2$.

4.1 Manipulating trace distance

Our protocols for (α, β)-QSD and its complement rely heavily on constructions for manipulating trace distances between outputs of quantum circuits. These constructions were developed by Sahai and Vadhan [34] in the classical case, and generalize to the quantum case with very few changes. The following theorem describes the main consequence of the constructions.

Theorem 5 *Let α and β satisfy $0 \leq \alpha < \beta^2 \leq 1$. Then there is a deterministic polynomial-time procedure that, on input $(Q_0, Q_1, 1^n)$ where Q_0 and Q_1 are descriptions of quantum circuits specifying mixed states ρ_0 and ρ_1, outputs descriptions of quantum circuits (R_0, R_1) (each having size polynomial in n and in the size of Q_0 and Q_1) specifying mixed states ξ_0 and ξ_1 satisfying the following.*

$$\begin{aligned} \|\rho_0 - \rho_1\|_{\mathrm{tr}} \;\leq\; \alpha \quad &\Rightarrow \quad \|\xi_0 - \xi_1\|_{\mathrm{tr}} \;\leq\; 2^{-n}, \\ \|\rho_0 - \rho_1\|_{\mathrm{tr}} \;\geq\; \beta \quad &\Rightarrow \quad \|\xi_0 - \xi_1\|_{\mathrm{tr}} \;\geq\; 1 - 2^{-n}. \end{aligned}$$

The remainder of this subsection is devoted to a proof of this theorem.

Proposition 6 *Let $\rho_0, \rho_1 \in \mathbf{D}(\mathcal{H})$ and $\xi_0, \xi_1 \in \mathbf{D}(\mathcal{K})$. Define*

$$\begin{aligned} \gamma_0 &= \frac{1}{2}(\rho_0 \otimes \xi_0) + \frac{1}{2}(\rho_1 \otimes \xi_1), \\ \gamma_1 &= \frac{1}{2}(\rho_0 \otimes \xi_1) + \frac{1}{2}(\rho_1 \otimes \xi_0). \end{aligned}$$

Then $\|\gamma_0 - \gamma_1\|_{\mathrm{tr}} = \|\rho_0 - \rho_1\|_{\mathrm{tr}} \, \|\xi_0 - \xi_1\|_{\mathrm{tr}}$.

Proof. We have

$$\begin{aligned} \|\gamma_0 - \gamma_1\|_{\mathrm{tr}} &= \left\| \frac{1}{2}(\rho_0 \otimes \xi_0) + \frac{1}{2}(\rho_1 \otimes \xi_1) - \frac{1}{2}(\rho_0 \otimes \xi_1) - \frac{1}{2}(\rho_1 \otimes \xi_0) \right\|_{\mathrm{tr}} \\ &= \left\| \frac{1}{2}(\rho_0 - \rho_1) \otimes (\xi_0 - \xi_1) \right\|_{\mathrm{tr}} \\ &= \|\rho_0 - \rho_1\|_{\mathrm{tr}} \cdot \|\xi_0 - \xi_1\|_{\mathrm{tr}} \end{aligned}$$

as required. ■

Lemma 7 *There is a deterministic polynomial-time procedure that, on input $(Q_0, Q_1, 1^r)$ where Q_0 and Q_1 are quantum circuits each having k specified output qubits, outputs (R_0, R_1), where R_0 and R_1 are quantum circuits each having rk specified output qubits and satisfy the following. Letting ρ_0, ρ_1, ξ_0, and ξ_1 denote the mixed states obtained by running Q_0, Q_1, R_0, and R_1 with all inputs in the $|0\rangle$ state and tracing out the output qubits, we have $\|\xi_0 - \xi_1\|_{\mathrm{tr}} = \|\rho_0 - \rho_1\|_{\mathrm{tr}}^r$.*

Proof. The circuit R_0 operates as follows: choose $b_1, \ldots, b_{r-1} \in \{0, 1\}$ independently and uniformly, set $b_r = b_1 \oplus \cdots \oplus b_{r-1}$, and output the state $\rho_{b_1} \otimes \cdots \otimes \rho_{b_r}$ (by running $Q_{b_1}, \ldots, Q_{b_r}$ on r separate collections of k qubits). The circuit R_1 operates similarly, except b_r is flipped: randomly choose $b_1, \ldots, b_{r-1} \in \{0, 1\}$ uniformly, set $b_r = 1 \oplus b_1 \oplus \cdots \oplus b_{r-1}$, and output the state $\rho_{b_1} \otimes \cdots \otimes \rho_{b_r}$. In both cases, the random choices are easily implemented using the Hadamard transform, and the construction of the circuits is straightforward. The required inequality $\|\xi_0 - \xi_1\|_{\mathrm{tr}} = \|\rho_{-}\rho_1\|_{\mathrm{tr}}^r$ follows from Proposition 6 along with a simple proof by induction. ■

Lemma 8 *Let ρ, $\xi \in \mathbf{D}(\mathcal{H})$ satisfy $\|\rho - \xi\|_{\mathrm{tr}} = \varepsilon$. Then*

$$1 - e^{-k\varepsilon^2/2} \;<\; \|\rho^{\otimes k} - \xi^{\otimes k}\|_{\mathrm{tr}} \;\leq\; k\varepsilon.$$

Proof. For the first inequality, we have

$$\|\rho^{\otimes k} - \xi^{\otimes k}\|_{\mathrm{tr}} \geq 1 - F(\rho^{\otimes k}, \xi^{\otimes k}) = 1 - F(\rho,\xi)^k \geq 1 - \left(\sqrt{1 - \|\rho - \xi\|_{\mathrm{tr}}^2}\right)^k$$
$$= 1 - (1-\varepsilon^2)^{\frac{k}{2}} = 1 - (1-\varepsilon^2)^{\frac{1}{\varepsilon^2}\cdot\frac{k\varepsilon^2}{2}} > 1 - e^{-\frac{k\varepsilon^2}{2}}.$$

In order to prove the second inequality, first note that for arbitrary states σ, σ', τ, and τ' we have

$$\begin{aligned}\|\sigma \otimes \sigma' - \tau \otimes \tau'\|_{\mathrm{tr}} &\leq \|\sigma \otimes \sigma' - \tau \otimes \sigma'\|_{\mathrm{tr}} + \|\tau \otimes \sigma' - \tau \otimes \tau'\|_{\mathrm{tr}} \\ &= \|(\sigma - \tau) \otimes \sigma'\|_{\mathrm{tr}} + \|\tau \otimes (\sigma' - \tau')\|_{\mathrm{tr}} \\ &= \|\sigma - \tau\|_{\mathrm{tr}} + \|\sigma' - \tau'\|_{\mathrm{tr}}.\end{aligned}$$

Based on this fact, the second inequality follows easily by induction. ■

Lemma 9 *There is a deterministic polynomial-time procedure that, on input $(Q_0, Q_1, 1^r)$ where Q_0 and Q_1 are quantum circuits each having k specified output qubits, outputs (R_0, R_1), where R_0 and R_1 are quantum circuits each having rk specified output qubits and satisfy the following. Letting ρ_0, ρ_1, ξ_0, and ξ_1 denote the mixed states obtained by running Q_0, Q_1, R_0, and R_1 with all inputs in the $|0\rangle$ state and tracing out the non-output qubits, we have*

$$1 - \exp\left(-\frac{r}{2}\|\rho_0 - \rho_1\|_{\mathrm{tr}}^2\right) \leq \|\xi_0 - \xi_1\|_{\mathrm{tr}} \leq r\,\|\rho_0 - \rho_1\|_{\mathrm{tr}}.$$

Proof. R_0 and R_1 are each simply obtained by running r independent copies of Q_0 and Q_1, respectively. Thus $\xi_i = \rho_i^{\otimes r}$ for $i = 0, 1$. The bounds on $\|\xi_0 - \xi_1\|_{\mathrm{tr}}$ follow from Lemma 8. ■

Proof of Theorem 5. We assume Q_0 and Q_1 each act on m qubits and have k specified output qubits for some choice of m and k.

Apply the construction in Lemma 7 to $(Q_0, Q_1, 1^r)$, where $r = \lceil \log(8n)/\log(\beta^2/\alpha)\rceil$. The result is circuits Q_0' and Q_1' that produce states ρ_0' and ρ_1' satisfying

$$\begin{aligned}\|\rho_0 - \rho_1\|_{\mathrm{tr}} < \alpha &\Rightarrow \|\rho_0' - \rho_1'\|_{\mathrm{tr}} < \alpha^r \\ \|\rho_0 - \rho_1\|_{\mathrm{tr}} > \beta &\Rightarrow \|\rho_0' - \rho_1'\|_{\mathrm{tr}} > \beta^r.\end{aligned}$$

Now apply the construction from Lemma 9 to $(Q_0', Q_1', 1^s)$, where $s = \lfloor \alpha^{-r}/2 \rfloor$. This results in circuits Q_0'' and Q_1'' that produce ρ_0'' and ρ_1'' such that

$$\begin{aligned}\|\rho_0 - \rho_1\|_{\mathrm{tr}} < \alpha &\Rightarrow \|\rho_0'' - \rho_1''\|_{\mathrm{tr}} < \alpha^r \alpha^{-r}/2 = 1/2, \\ \|\rho_0 - \rho_1\|_{\mathrm{tr}} > \beta &\Rightarrow \|\rho_0'' - \rho_1''\|_{\mathrm{tr}} > 1 - \exp\left(-\frac{s}{2}\beta^{2r}\right) \geq 1 - e^{-2n+1}.\end{aligned}$$

Finally, again apply the construction from Lemma 7, this time to $(Q_0'', Q_1'', 1^n)$. This results in circuits R_0 and R_1 that produce states ξ_0 and ξ_1 satisfying

$$\begin{aligned}\|\rho_0 - \rho_1\|_{\mathrm{tr}} < \alpha &\Rightarrow \|\xi_0 - \xi_1\|_{\mathrm{tr}} < 2^{-n}, \\ \|\rho_0 - \rho_1\|_{\mathrm{tr}} > \beta &\Rightarrow \|\xi_0 - \xi_1\|_{\mathrm{tr}} > \left(1 - e^{-2n+1}\right)^n > 1 - 2^{-n}.\end{aligned}$$

The circuits R_0 and R_1 have size polynomial in n and the size of Q_0 and Q_1 as required. ■

Verifier:

Apply the construction of Theorem 5 to $(Q_0, Q_1, 1^n)$ for n exceeding the length of the input (Q_0, Q_1). Let R_0 and R_1 denote the constructed circuits, and ξ_0 and ξ_1 the associated mixed states. Choose $a \in \{0, 1\}$ uniformly and send ξ_a to the prover.

Prover:

Perform the optimal measurement for distinguishing ξ_0 and ξ_1. Let b be 0 if the measurement indicates the state is ξ_0, and 1 if the measurement indicates the state is ξ_1. Send b to the verifier.

Verifier:

Accept if $a = b$ and reject otherwise.

Figure 2: Protocol for QSD

4.2 Distance test

Now we describe a quantum statistical zero-knowledge protocol for (α, β)-QSD. The prover's goal in this case is to prove that the mixed states produced by the given circuits Q_0 and Q_1 are far apart in the trace norm metric. The protocol is described in Figure 2, including a description of the (honest) prover.

This protocol is identical in principle to the well-known Graph Non-isomorphism protocol of Goldreich, Micali, and Wigderson [17] and Quadratic Non-residuosity protocol of Goldwasser, Micali, and Rackoff [20]. If the states are indeed far apart, the prover can determine which state was sent by performing an appropriate measurement, while if the states are close together, the prover cannot reliably tell the difference between the states because there does not exist a measurement that distinguishes them. By requiring that the verifier first apply the construction of Theorem 5, an exponentially small completeness error is achieved, which ensures that the zero-knowledge property holds.

Based on this protocol, we have the following theorem.

Theorem 10 *Let α and β satisfy $0 \leq \alpha < \beta^2 \leq 1$. Then (α, β)-QSD $\in$ QSZK$_{HV}$.*

Proof. First we discuss the completeness and soundness of the proof system, then prove that the zero-knowledge property holds.

For the completeness property of the protocol, we assume that the prover receives one of ξ_0 or ξ_1, where $\|\xi_0 - \xi_1\|_{\mathrm{tr}} > 1 - 2^{-n}$, and thus can distinguish the two cases with probability of error bounded by 2^{-n} by performing an appropriate measurement. Specifically, the prover can apply the measurement described by orthogonal projections $\{\Pi_0, \Pi_1\}$ where Π_0 maximizes $\operatorname{tr} \Pi_0(\xi_0 - \xi_1)$ and $\Pi_1 = I - \Pi_0$. This gives an outcome of 0 with probability at least $1 - 2^{-n}$ in case the verifier sent ξ_0 and gives an outcome of 1 with probability at least $1 - 2^{-n}$ in case the verifier sent ξ_1. This will cause the verifier to accept with probability at least $1 - 2^{-n}$.

For the soundness condition, we assume the prover receives either ξ_0 or ξ_1 where $\|\xi_0 - \xi_1\|_{\mathrm{tr}} < 2^{-n}$, and then the prover returns a single bit to the verifier. There is no loss of generality in

Verifier:

Apply the construction of Theorem 5 to $(Q_0, Q_1, 1^n)$ for n exceeding the length of the input (Q_0, Q_1). Let R_0 and R_1 denote the constructed circuits, and ξ_0 and ξ_1 the associated mixed states. Let t be the number of qubits on which R_0 and R_1 act. Apply R_0 to $|0^t\rangle$ and send the prover **only** the non-output qubits (that is, the qubits that would be traced-out to yield ξ_0).

Prover:

Apply unitary transformation U (described below) to the qubits sent by the verifier, then send these qubits back to the verifier.

Verifier:

Apply $R_1^\dagger$ to the output qubits of R_0 (which were not send to the prover in the first message) together with the qubits received from the prover. Measure the resulting qubits: *accept* if the result is 0^t, and *reject* otherwise.

Figure 3: Protocol for co-QSD

assuming that the bit sent by the prover is measured immediately upon being received by the verifier, since this would not change the verifier's decision to accept or reject. Thus, we may treat this bit as being the outcome of a measurement of whichever state ξ_0 or ξ_1 was initially sent by the verifier. Since the trace distance between these two states is at most 2^{-n}, no measurement can distinguish the states with bias exceeding 2^{-n}. Consequently the prover has probability at most $1/2 + 2^{-n}$ of correctly answering $\tilde{b} = b$.

Finally, the zero-knowledge property is straightforward—the state of the verifier and message qubits after the first message is obtained by applying V_1 (the verifier's first transformation), and the state of the verifier and message qubits after the prover's response is approximated by applying V_1, tracing out the message qubits, then setting $\tilde{b}$ to b. Since the completeness error is exponentially small, this gives a negligible error for the simulator. ■

4.3 Closeness test

Next we give a protocol for the complement of (α,β)-QSD. Unlike the previous protocol this protocol has no obvious classical analogue and relies on non-classical properties of quantum states. A description of the protocol is given in Figure 3.

The correctness of the protocol is related to the following fact concerning bipartite quantum states. Suppose that $|\phi\rangle$ and $|\phi'\rangle$ are pure quantum states in some tensor product space $\mathcal{H}\otimes\mathcal{K}$ such that the two states give the same mixed state when the second system is traced-out: $\mathrm{tr}_{\mathcal{K}}\,|\phi\rangle\langle\phi| = \mathrm{tr}_{\mathcal{K}}\,|\phi'\rangle\langle\phi'| = \rho$. Then by Theorem 1 there must exist a unitary operator U acting only on the traced-out space $\mathcal{K}$ such that $(I\otimes U)|\phi\rangle = |\phi'\rangle$. In case $\rho = \mathrm{tr}_{\mathcal{K}}\,|\phi\rangle\langle\phi|$ and $\rho' = \mathrm{tr}_{\mathcal{K}}\,|\phi'\rangle\langle\phi'|$ are not identical, but are close together in the trace norm metric, an approximate version of this fact holds: there exists a unitary operator U acting on $\mathcal{K}$ such that $(I\otimes U)|\phi\rangle$ and $|\phi'\rangle$ are close in Euclidean norm. For the above protocol, the states $|\phi\rangle$ and $|\phi'\rangle$ are the states produced by R_0 and R_1, $\mathcal{K}$ is the space corresponding to the qubits sent to the prover, and U corresponds to the action of the prover. We formalize this argument in the proof of the following theorem.

Theorem 11 *Let α and β satisfy $0 \leq \alpha < \beta^2 \leq 1$. Then (α, β)-QSD $\in$ co-QSZK$_{HV}$.*

The proof of this theorem will require the following lemma, which is the approximate version of Theorem 1 mentioned previously.

Lemma 12 *Let $\rho, \xi \in \mathbf{D}(\mathcal{H})$ satisfy $F(\rho, \xi) \geq 1 - \varepsilon$ and let $|\phi\rangle, |\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ be purifications of ρ and ξ, respectively, i.e., $\operatorname{tr}_{\mathcal{K}} |\phi\rangle\langle\phi| = \rho$ and $\operatorname{tr}_{\mathcal{K}} |\psi\rangle\langle\psi| = \xi$. Then there exists $U \in \mathbf{U}(\mathcal{K})$ such that*

$$\|(I \otimes U)|\phi\rangle - |\psi\rangle\| \leq \sqrt{2\varepsilon}.$$

Proof. By Theorem 2 we have

$$F(\rho, \xi) = \max_{|\phi_0\rangle, |\psi_0\rangle} |\langle\phi_0|\psi_0\rangle|,$$

where the maximum is over all purifications $|\phi_0\rangle, |\psi_0\rangle \in \mathcal{K}$ of ρ and ξ, respectively. Let $|\phi_0\rangle$ and $|\psi_0\rangle$ be pure states achieving this maximum, and assume without loss of generality that $\langle\phi_0|\psi_0\rangle$ is a nonnegative real number.

Since $|\phi\rangle$ and $|\phi_0\rangle$ are both purifications of ρ, we have by Theorem 1 that there exists $V \in \mathbf{U}(\mathcal{K})$ such that $|\phi_0\rangle = (I \otimes V)|\phi\rangle$. Similarly, there exists $W \in \mathbf{U}(\mathcal{K})$ such that $|\psi_0\rangle = (I \otimes W)|\psi\rangle$.

Define $U = V^{\dagger}W$. Then

$$\|(I \otimes U)|\phi\rangle - |\psi\rangle\| = \|(I \otimes W)|\phi\rangle - (I \otimes V)|\psi\rangle\| = \||\phi_0\rangle - |\psi_0\rangle\| = \sqrt{2 - 2\langle\phi_0|\psi_0\rangle} \leq \sqrt{2\varepsilon}$$

as required. ■

Proof of Theorem 11. First let us consider the completeness condition. If $(Q_0, Q_1) \not\in (\alpha, \beta)$-QSD then we have $\|\xi_0 - \xi_1\|_{\text{tr}} < 2^{-n}$ and thus $F(\xi_0, \xi_1) > 1 - 2^{-n}$. The states $R_0|0^t\rangle$ and $R_1|0^t\rangle$ are purifications of ξ_0 and ξ_1, respectively, and consequently there exists a unitary transformation U acting only on the non-output qubits of $R_0|0^t\rangle$ (i.e., the qubits sent to the prover) such that $\|(I \otimes U)R_0|0^t\rangle - R_1|0^t\rangle\| \leq 2^{-(n-1)/2}$. This is the transformation U performed by the honest prover. This causes the verifier to accept with probability

$$|\langle 0^t|R_1^{\dagger}(I \otimes U)R_0|0^t\rangle|^2 \geq \left(1 - \frac{1}{2}\|R_1|0^t\rangle - (I \otimes U)R_0|0^t\rangle\|^2\right)^2 > 1 - 2^{-n+1}.$$

The soundness of the proof system may be proved as follows. Assume $(Q_0, Q_1) \in (\alpha, \beta)$-QSD. Then we have $\|\xi_0 - \xi_1\|_{\text{tr}} > 1 - 2^{-n}$, and thus $F(\xi_0, \xi_1) < 2^{-(n-1)/2}$. The verifier prepares $R_0|0^t\rangle$ and sends the non-output qubits to the prover. The most general action of the prover is to apply some arbitrary unitary transformation to the qubits sent by the verifier along with any number of its own private qubits, and then return some number of these qubits to the verifier. As usual, let $\mathcal{V}$ and $\mathcal{M}$ denote the Hilbert spaces corresponding to the verifier's qubits and the message qubits, respectively, and let σ denote the mixed state of the verifier's private qubits and the message qubits immediately after the prover has sent its message. It must be the case that $\operatorname{tr}_{\mathcal{M}} \sigma = \xi_0$, since the reduced state of the verifier's qubits cannot be modified by the prover, as the prover does not have access to these qubits. The verifier applies $R_1^\dagger$ and measures, which causes the verifier to accept with probability

$$\langle 0^t | R_1^\dagger \sigma R_1 | 0^t \rangle = F(R_1|0^t\rangle\langle 0^t|R_1^\dagger, \sigma)^2 \leq F(\operatorname{tr}_{\mathcal{M}} R_1|0^t\rangle\langle 0^t|R_1^\dagger, \operatorname{tr}_{\mathcal{M}} \sigma)^2 = F(\xi_1, \xi_0)^2 < 2^{-n+1}.$$

Finally, the zero-knowledge property is fairly straightforward. We define a simulator that outputs $R_0|0^t\rangle$ for the verifier's view as the first message is being sent and $R_1|0^t\rangle$ for the verifier's view after the second message. The simulator is perfect for the first message, and has trace distance at most 2^{-n+1} from the view of the verifier interacting with the prover defined above for the second message. ■

4.4 Alternate public-coin closeness test

Okamoto [32] proved that any language having a classical interactive proof system that is statistical zero-knowledge against an honest verifier also has a *public-coin* proof system that is statistical zero-knowledge against an honest verifier. This turned out to be a key fact in the proof of Goldreich, Sahai and Vadhan [18] that honest-verifier statistical zero-knowledge equals statistical zero-knowledge (without the honest verifier assumption).

We prove that co-(α,β)-QSD, i.e., the complement of the (α,β)-QSD problem, has a public-coin honest-verifier quantum statistical zero-knowledge proof system. As this problem is complete for $\mathrm{QSZK_{HV}}$, as demonstrated in the next section, we have that any honest-verifier quantum statistical zero-knowledge proof system can be transformed into a public-coin honest-verifier quantum statistical zero-knowledge proof system. By a public-coin honest-verifier quantum statistical zero-knowledge proof system, we mean that each of the verifier's messages is given by a sequence of fair coin-flips. We stress that the verifier's messages are completely classical, and the verifier does not need to perform any computation, quantum or classical, until after all messages have been exchanged.

Our protocol only requires a single coin-flip on the part of the verifier (in order to achieve exponentially small completeness error and soundness error exponentially close to 1/2). The downside of the public-coin proof system we present is that three messages are required rather than two as for the previously presented proof system for this problem. In contrast to the classical case, the proof that our proof system functions correctly and has the required zero-knowledge property is easy, following from simple facts about quantum information. The protocol is described in Figure 4. Based on this protocol, we have the following theorem.

Theorem 13 *Let α and β satisfy $0 \leq \alpha < \beta^2 \leq 1$. Then there exists a three-message public-coin honest-verifier statistical zero-knowledge proof system for co-(α,β)-QSD.*

Proof. The completeness of the proof system is essentially the same as in the proof of Theorem 11; in case $(Q_0, Q_1) \notin (\alpha,\beta)$-QSD the honest prover will convince the verifier to accept with probability at least $1 - 2^{-n+1}$.

Now suppose $(Q_0, Q_1) \in (\alpha,\beta)$-QSD, so that $\|\xi_0 - \xi_1\|_{\mathrm{tr}} > 1 - 2^{-n}$, and thus $F(\xi_0, \xi_1) < 2^{-(n-1)/2}$. Let the reduced density operator of the qubits sent by the prover to the verifier in the first message be denoted σ. Then the maximum probability that the verifier can be made to accept is

$$\frac{1}{2}F(\xi_0, \sigma)^2 + \frac{1}{2}F(\xi_1, \sigma)^2.$$

Prover:

Apply the construction of Theorem 5 to $(Q_0, Q_1, 1^n)$ for n exceeding the length of the input (Q_0, Q_1). Let R_0 and R_1 denote the constructed circuits, and ξ_0 and ξ_1 the associated mixed states. Let t be the number of qubits on which R_0 and R_1 act. Apply R_0 to $|0^t\rangle$ and send the verifier the output qubits (that is, send the verifier ξ_0, and keep the qubits that purify this state).

Verifier:

Flip a fair coin and send the result to the prover. Denote the value of the coin-flip by $b \in \{0, 1\}$.

Prover:

If $b = 1$, apply the unitary transformation U (exactly as in the proof of Theorem 11) to the non-output qubits of R_0 that were not sent to the verifier in the first message, then send these qubits to the verifier. If $b = 0$, send these qubits to the verifier without performing any operation on them.

Verifier:

The verifier also applies the construction of Theorem 5 as in the prover's first step. Apply $R_b^\dagger$ to all of the qubits received from the prover (in both the first and second message) and measure them in the standard basis: *accept* if the result is 0^t, and *reject* otherwise.

Figure 4: Public-coin protocol for co-QSD

To see this, suppose first that $b = 1$. Let τ be the state of all of the qubits sent to the verifier. Then the probability the verifier accepts is $\langle 0^t | R_1^\dagger \tau R_1 | 0^t \rangle$. Since tracing out the qubits of τ sent by the prover in the second round gives σ and $R_1|0^t\rangle$ purifies ξ_1, the probability of acceptance is bounded by $F(\xi_1, \sigma)^2$ by Theorem 2.

In case $b = 0$ the maximum probability of acceptance is $F(\xi_0, \sigma)^2$, which follows by a similar argument, and gives the claimed bound. Now, by Lemma 3 we have

$$\frac{1}{2}F(\xi_0, \sigma)^2 + \frac{1}{2}F(\xi_1, \sigma)^2 \leq \frac{1}{2}(1 + F(\xi_0, \xi_1)) \leq \frac{1}{2} + 2^{-(n+1)/2}.$$

Finally, for the zero-knowledge property, define a simulator as follows. Output ξ_0 for the verifier's view after the first message is sent. In order to produce the view after the second and third message, choose $b \in \{0, 1\}$ uniformly and prepare $R_0|0^t\rangle$ or $R_1|0^t\rangle$ appropriately. The simulator is perfect for the first and second message, and has negligible trace distance from the view of the verifier interacting with the prover defined above for the third message. ■

5 $\mathrm{QSZK_{HV}}$–completeness of quantum state distinguishability

Given a promise problem $B = (B_{\mathrm{yes}}, B_{\mathrm{no}})$, we say that B is complete for $\mathrm{QSZK_{HV}}$ if (i) $B \in \mathrm{QSZK_{HV}}$, and (ii) for every promise problem $A = (A_{\mathrm{yes}}, A_{\mathrm{no}}) \in \mathrm{QSZK_{HV}}$ there exists a polynomial-time computable function f such that $x \in A_{\mathrm{yes}} \Rightarrow f(x) \in B_{\mathrm{yes}}$ and $x \in A_{\mathrm{no}} \Rightarrow f(x) \in B_{\mathrm{no}}$. In this section we prove that (α, β)-QSD is a complete promise problem for $\mathrm{QSZK_{HV}}$.

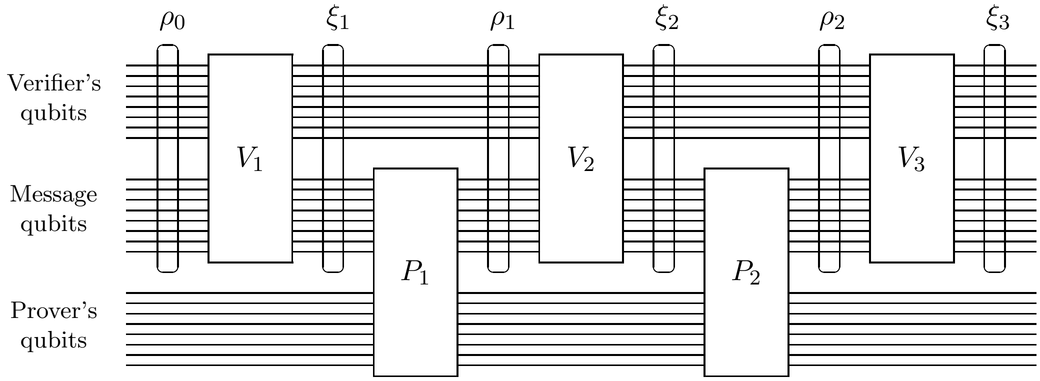

Figure 5: States $\rho_0, \ldots, \rho_{k-1}$ and $\xi_1, \ldots, \xi_k$ for $m = 4$.

Theorem 14 *Let α and β satisfy $0 < \alpha < \beta^2 < 1$. Then (α, β)-QSD is complete for QSZK$_{HV}$.*

Before giving the proof of this theorem, we will describe the main intuition behind the proof. We note that the main idea behind the proof is a familiar idea due to Fortnow [14] that has been used several times to prove results concerning statistical zero-knowledge.

Since we have already proved that (α, β)-QSD and its complement are contained in QSZK$_{\text{HV}}$ in the previous section, it will suffice to prove that co-(α, β)-QSD is hard for QSZK$_{\text{HV}}$.

Suppose that we have a quantum interactive proof system (V, P) for some promise problem A that is statistical zero-knowledge against an honest verifier. Assume that the completeness and soundness errors for this proof system are exponentially small; since sequential repetition can be used to reduce error exponentially, there is no loss of generality in making this assumption. Let $m = m(|x|)$ denote the total number of messages sent in this protocol, and assume without loss of generality that m is even. Let $k = m/2 + 1$, so that the verifier's actions on a fixed input x are described by circuits $V_1, \ldots, V_k$ and the prover's actions are described by circuits $P_1, \ldots, P_{k-1}$. The case where $m = 4$ is illustrated in Figure 5.

Let $\rho_0, \ldots, \rho_{k-1}$ and $\xi_1, \ldots, \xi_k$ correspond to the simulator's approximation to the reduced states of the verifier's qubits together with the message qubits at various times during the protocol as suggested by Figure 5. In the case that $x \in A_{\text{yes}}$, these states are therefore close to the actual views of the verifier at these instants in the protocol, while there is no guarantee in case $x \in A_{\text{no}}$.

We can impose some restrictions on the simulator without loss of generality. First, we may of course assume that $\rho_0 = |0 \cdots 0\rangle\langle 0 \cdots 0|$, and that $\xi_j = V_j \rho_{j-1} V_j^\dagger$ for $j = 1, \ldots, k$ (i.e., the simulator prepares each ξ_j by first preparing ρ_{j-1} and then applying V_j to this state). Lastly, we can assume that the output qubit of ξ_k is set to 1 with certainty. If a given simulator does not satisfy this last restriction, we can define a new simulator that produces ρ_{k-1} according to the following procedure: (i) produce some approximation ρ'_{k-1} by running the original simulator, (ii) apply V_k to this state, (iii) replace the output qubit with a qubit in state $|1\rangle$, and (iv) apply $V_k^\dagger$ to the resulting state to give ρ_{k-1}. The output qubit of $\xi_k = V_k \rho_{k-1} V_k^\dagger$ is therefore set to 1 with certainty. If $x \in A_{\text{yes}}$, then the new simulator will produce an output that has negligible distance from the output of the original simulator, and therefore the actual view of V, based on the fact that the proof system is assumed to have exponentially small completeness error.

Assume first that $x \in A_{\text{yes}}$. Since $x \in A_{\text{yes}}$, the simulator outputs states $\rho_0, \ldots, \rho_{k-1}$ and $\xi_1, \ldots, \xi_k$ that closely approximate the actual view of V. We see that in this case it must be that $\operatorname{tr}_{\mathcal{M}} \xi_j \approx \operatorname{tr}_{\mathcal{M}} \rho_j$ for $j = 1, \ldots, k-1$, since the prover cannot affect the reduced state of the verifier's qubits. Consequently, the two states

$$\begin{gathered} \operatorname{tr}_{\mathcal{M}} \rho_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \rho_{k-1} \\ \operatorname{tr}_{\mathcal{M}} \xi_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \xi_{k-1} \end{gathered} \tag{1}$$

are close together (within negligible trace distance). Based on the description of the simulator, it is possible to construct quantum circuits Q_0 and Q_1 that output the states $\operatorname{tr}_{\mathcal{M}} \rho_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \rho_{k-1}$ and $\operatorname{tr}_{\mathcal{M}} \xi_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \xi_{k-1}$, respectively. The descriptions of the circuits Q_0 and Q_1 will (essentially) be the instance of co-(α, β)-QSD produced by the reduction.

Now, assume that $x \in A_{\text{no}}$, and suppose we use the same method as above to get circuits Q_0 and Q_1. Since $x \in A_{\text{no}}$, however, we no longer have any guarantee that the states produced by these circuits are close together. Indeed, our goal will be to prove that the states produced by Q_0 and Q_1 are necessarily far apart. Fortunately, this is not difficult; based on the assumption that no prover can succeed in convincing the verifier to accept with high probability, we have that the states in Eq. 1 are far apart given any choice of $\rho_0, \ldots, \rho_{k-1}$ and $\xi_1, \ldots, \xi_k$ satisfying the various constraints that are imposed on the output of the simulator. Proving this claim is the main technical part of the proof, and corresponds to Lemma 15.

Finally, in order to achieve the required error bounds for co-(α, β)-QSD, the constructions used to prove Theorem 5 are applied to (Q_0, Q_1), resulting in (R_0, R_1) that are elements of (α, β)-QSD_{no} or (α, β)-QSD_{yes} appropriately.

We now proceed with a more formal proof of Theorem 14. Recall that we denote by Π_{acc} the projection onto states for which the verifier's output qubit is set to 1 (accept), and in addition we will denote by Π_{init} the projection onto the state for which all of the verifier and message qubits are set to 0 (which is the initial state of these qubits before the protocol begins). These projections are operators acting on $\mathcal{V} \otimes \mathcal{M}$, but by our convention they can also be viewed as operators acting on $\mathcal{V} \otimes \mathcal{M} \otimes \mathcal{P}$ by tensoring with the identity. We begin with the technical lemma needed to show that the states in Eq. 1 are far apart for the case $x \in A_{\text{no}}$.

Lemma 15 *Consider an m-message quantum interactive proof system for even m on some input string for which the maximum acceptance probability is ε. Let $V_1, \ldots, V_k$ denote the verifier's circuits (where $k = m/2 + 1$), let $\rho_0, \ldots, \rho_{k-1}$ be any mixed quantum states over $\mathcal{V} \otimes \mathcal{M}$, let $\xi_j = V_j \rho_{j-1} V_j^\dagger$ for $j = 1, \ldots, k$, and assume that* $\operatorname{tr}(\Pi_{\text{init}} \rho_0) = \operatorname{tr}(\Pi_{\text{acc}}\, \xi_k) = 1$. *Then*

$$\| \operatorname{tr}_{\mathcal{M}} \xi_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \xi_{k-1} - \operatorname{tr}_{\mathcal{M}} \rho_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \rho_{k-1} \|_{\text{tr}} \;\geq\; \frac{(1-\sqrt{\varepsilon})^2}{4(k-1)}.$$

Proof. Let $|\phi_0\rangle \in \mathcal{V} \otimes \mathcal{M} \otimes \mathcal{P}$ be the state in which all of the verifier, message, and prover qubits are set to 0. Since $\operatorname{tr}(\Pi_{\text{init}} \rho_0) = 1$, $|\phi_0\rangle$ is necessarily a purification of ρ_0. Let $|\phi_1\rangle, \ldots, |\phi_{k-1}\rangle \in \mathcal{V} \otimes \mathcal{M} \otimes \mathcal{P}$ be arbitrary purifications of $\rho_1, \ldots, \rho_{k-1}$, and set $|\psi_j\rangle = V_j |\phi_{j-1}\rangle$ for $j = 1, \ldots, k$. Since $\xi_j = V_j \rho_{j-1} V_j^\dagger$ for $j = 1, \ldots, k$, each $|\psi_j\rangle$ is a purification of ξ_j.

Let $\delta_j = 1 - F(\operatorname{tr}_{\mathcal{M}} \xi_j, \operatorname{tr}_{\mathcal{M}} \rho_j)$ for $j = 1, \ldots, k-1$. It follows from Lemma 12 that for each j there exists some unitary operator P_j acting on $\mathcal{M} \otimes \mathcal{P}$ such that $\| P_j |\psi_j\rangle - |\phi_j\rangle \| \leq \sqrt{2\delta_j}$. Now, for each $j = 1, \ldots, k$, we have

$$\begin{aligned}
&\|V_j P_{j-1} V_{j-1} \cdots P_1 V_1 |\phi_0\rangle - |\psi_j\rangle\| \\
&\qquad = \|P_{j-1} V_{j-1} \cdots P_1 V_1 |\phi_0\rangle - |\phi_{j-1}\rangle\| \\
&\qquad \leq \|P_{j-1} V_{j-1} \cdots P_1 V_1 |\phi_0\rangle - P_{j-1}|\psi_{j-1}\rangle\| + \|P_{j-1}|\psi_{j-1}\rangle - |\phi_{j-1}\rangle\| \\
&\qquad \leq \|V_{j-1} \cdots P_1 V_1 |\phi_0\rangle - |\psi_{j-1}\rangle\| + \sqrt{2\delta_{j-1}},
\end{aligned}$$

and therefore

$$\|V_k P_{k-1} V_{k-1} \cdots P_1 V_1 |\phi_0\rangle - |\psi_k\rangle\| \leq \sum_{j=1}^{k-1} \sqrt{2\delta_j}.$$

Since $\operatorname{tr}(\Pi_{\text{acc}} \xi_k) = 1$ we have $\|\Pi_{\text{acc}} |\psi_k\rangle\| = 1$, and so

$$\|\Pi_{\text{acc}} V_k P_{k-1} V_{k-1} \cdots P_1 V_1 |\phi_0\rangle\| \geq 1 - \sum_{j=1}^{k-1} \sqrt{2\delta_j}.$$

Since $\|\Pi_{\text{acc}} V_k P_{k-1} V_{k-1} \cdots P_1 V_1 |\phi_0\rangle\|^2$ is bounded above by the acceptance probability of the proof system, we have

$$\sum_{j=1}^{k-1} \sqrt{2\delta_j} \geq 1 - \sqrt{\varepsilon}. \tag{2}$$

It remains to use Eq. 2 to put a lower bound on the quantity of interest, which can be done as follows. The fidelity is multiplicative with respect to tensor products, so

$$F(\operatorname{tr}_{\mathcal{M}} \xi_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \xi_{k-1}, \operatorname{tr}_{\mathcal{M}} \rho_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \rho_{k-1}) = \prod_{i=1}^{k-1} F(\operatorname{tr}_{\mathcal{M}} \xi_i, \operatorname{tr}_{\mathcal{M}} \rho_i) = \prod_{i=1}^{k-1} (1 - \delta_i).$$

Subject to the constraint in Eq. 2, we have

$$\prod_{i=1}^{k-1} (1 - \delta_i) \leq \left(1 - \frac{(1-\sqrt{\varepsilon})^2}{2(k-1)^2}\right)^{k-1} \leq \exp\left(-\frac{(1-\sqrt{\varepsilon})^2}{2(k-1)}\right) \leq 1 - \frac{(1-\sqrt{\varepsilon})^2}{4(k-1)}.$$

Therefore, by Theorem 4,

$$\|\operatorname{tr}_{\mathcal{M}} \xi_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \xi_{k-1} - \operatorname{tr}_{\mathcal{M}} \rho_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \rho_{k-1}\|_{\text{tr}} \geq \frac{(1-\sqrt{\varepsilon})^2}{4(k-1)}$$

as required. ∎

Proof of Theorem 14. Let $A \in \text{QSZK}_{\text{HV}}$, and let (V, P) be an honest verifier quantum statistical zero-knowledge proof system for A with completeness and soundness error smaller than 2^{-n} for inputs of length n. Such a proof system exists, since sequential repetition reduces completeness and soundness errors exponentially while preserving the zero-knowledge property of honest verifier quantum statistical zero-knowledge proof systems. Let $m = m(|x|)$ be the number of messages exchanged by P and V. Without loss of generality we assume that the number of messages m is even for all x, adding an initial move where the verifier sends some arbitrary state to the prover if necessary. Thus, the verifier will apply transformations $V_1, \ldots, V_k$ for $k = m/2 + 1$, and will send the first message in the protocol. We let $\{\sigma_{x,j}\}$ correspond to the mixed states output by the simulator for (V, P) as discussed in Section 2. The quantum circuits that produce the states $\{\sigma_{x,j}\}$ are used implicitly in the reduction.

First, we describe, for any fixed input x, the following quantum states:

1. Let ρ_0 be the state in which all verifier and message qubits are in state $|0\rangle$.
2. Let ξ_k denote the state obtained by applying V_k to $\sigma_{x,m}$, discarding the output qubit, and replacing it with a qubit in state $|1\rangle$.
3. Let $\rho_i = \sigma_{x,2i}$ for $i = 1, \ldots, k-2$ and let $\rho_{k-1} = V_k^\dagger \xi_k V_k$.
4. Let $\xi_i = V_i \rho_{i-1} V_i^\dagger$ for $i = 1, \ldots, k-1$.

These states are illustrated in Figure 5 for the case $m = 4$ (meaning that these states will be close approximations to the illustrated states given an input $x \in A_{\text{yes}}$). Let Q_0 and Q_1 be quantum circuits that output states

$$\gamma_0 = \operatorname{tr}_{\mathcal{M}} \rho_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \rho_{k-1}$$
$$\gamma_1 = \operatorname{tr}_{\mathcal{M}} \xi_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \xi_{k-1}$$

respectively, assuming the input to these circuits is a suitable number of qubits in the $|0\rangle$ state. Such circuits can be constructed in polynomial time based on V and on the simulator for (V, P).

We claim that the following implications hold:

$$\begin{aligned} x \in A_{\text{yes}} \quad &\Rightarrow \quad \|\gamma_0 - \gamma_1\|_{\text{tr}} < \delta(|x|) \\ x \in A_{\text{no}} \quad &\Rightarrow \quad \|\gamma_0 - \gamma_1\|_{\text{tr}} > c/k \end{aligned}$$

where $\delta(|x|)$ is a negligible function (determined by the accuracy of the simulator for (V, P)) and $c > 0$ is some constant. The second implication follows from Lemma 15. To prove the first implication, consider states $\rho'_0, \ldots, \rho'_{k-1}$ and $\xi'_1, \ldots, \xi'_k$ obtained precisely as in the description of Q_0 and Q_1, except replacing $\sigma_{x,j}$ with $\text{view}_{V,P}(x, j)$, the actual view of the verifier V while interacting with P, for each x and j. We necessarily have $\operatorname{tr}_{\mathcal{M}} \xi'_i = \operatorname{tr}_{\mathcal{M}} \rho'_i$ for $i = 1, \ldots, k-2$. Since measuring the output qubit of $V_k \, \text{view}_{V,P}(x, m) V_k^\dagger$ gives 1 with probability at least $1 - 2^{-|x|}$, replacing the output qubit with a qubit in state $|1\rangle$ has a negligible effect on this state. Thus, the trace distance between $\operatorname{tr}_{\mathcal{M}} \rho'_{k-1}$ and $\operatorname{tr}_{\mathcal{M}} \xi'_{k-1}$, and therefore between $\operatorname{tr}_{\mathcal{M}} \rho'_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \rho'_{k-1}$ and $\operatorname{tr}_{\mathcal{M}} \xi'_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \xi'_{k-1}$, is negligible. Now, since the simulator deviates from $\text{view}_{V,P}$ by a negligible quantity on each input, the inequality

$$\| \operatorname{tr}_{\mathcal{M}} \rho_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \rho_{k-1} \; - \; \operatorname{tr}_{\mathcal{M}} \xi_1 \otimes \cdots \otimes \operatorname{tr}_{\mathcal{M}} \xi_{k-1} \|_{\text{tr}} < \delta(|x|)$$

for some negligible $\delta(|x|)$ follows from the triangle inequality.

Finally, by applying the constructions from Lemmas 7 and 9 to (Q_0, Q_1) appropriately results in circuits R_0 and R_1 that specify mixed states γ_0 and γ_1, respectively, such that (i) $x \in A_{\text{yes}} \Rightarrow \|\gamma_0 - \gamma_1\|_{\text{tr}} < \alpha$, and (ii) $x \in A_{\text{no}} \Rightarrow \|\gamma_0 - \gamma_1\|_{\text{tr}} > \beta$, for any chosen constants $\alpha, \beta \in (0, 1)$. Thus, $x \in A_{\text{yes}}$ implies $(R_0, R_1) \in (\alpha, \beta)\text{-QSD}_{\text{no}}$ and $x \in A_{\text{no}}$ implies $(R_0, R_1) \in (\alpha, \beta)\text{-QSD}_{\text{yes}}$. ∎

6 Consequences

Based on Theorems 10, 11, 13, and 14, the following corollaries are immediate.

Corollary 16 *$QSZK_{HV}$ is closed under complement.*

Corollary 17 *For any language in $QSZK_{HV}$ there exists a two-message honest verifier quantum statistical zero-knowledge proof system with exponentially small completeness and soundness error.*

Corollary 18 *For any language in $QSZK_{HV}$ there exists a two-message honest verifier quantum statistical zero-knowledge proof system with exponentially small completeness error and soundness error exponentially close to 1/2 in which the prover's message to the verifier consists of a single bit.*

Corollary 19 *For any language in $QSZK_{HV}$ there exists a three-message public coin honest verifier quantum statistical zero-knowledge proof system with exponentially small completeness error and soundness error exponentially close to 1/2 in which the verifier's message consists of a single coin-flip.*

Finally, the fact that (α, β)-QSD can be shown to be in PSPACE yields the following corollary.

Corollary 20 *$QSZK_{HV} \subseteq PSPACE$.*

In order to prove this Corollary, let us consider the following problem.

Trace Norm Approximation (TNA)

Input: An $n \times n$ matrix X (with entries having rational real and imaginary parts) and an accuracy parameter 1^k.

Output: A nonnegative rational number r satisfying $|\, r - \|X\|_{\mathrm{tr}} \,| \; < \; 2^{-k}$.

Proposition 21 *TNA $\in$ NC.*

Proof. [Sketch] Consider the following algorithm.

1. Compute $Y = XX^\dagger$.
2. Compute the characteristic polynomial of Y (the coefficients will be real since Y is necessarily Hermitian).
3. Calculate the n roots $\lambda_1, \ldots, \lambda_n$ of the characteristic polynomial of Y to $O(k + \log n)$ bits of precision.
4. Compute $r = \frac{1}{2}\sum_{j=1}^{n} \sqrt{\lambda_j}$, where each square root is approximated to $O(k + \log n)$ bits of precision, and output r.

The output r is an approximation to one-half the trace of $\sqrt{XX^\dagger}$, which is $\|X\|_{\mathrm{tr}}$. The approximation is correct to $O(k)$ bits of precision as required. Each step can be performed in NC; simple arithmetic operations and multiplication of matrices are well-known to be in NC, the fact that the characteristic polynomial can be computed in NC was shown by Csanky [12], and polynomial root approximation was shown to be in NC by Neff [30]. ■

Proof of Corollary 20. [Sketch] By Theorem 14 it suffices to show that (α, β)-QSD is in PSPACE. Recall that for any function $s(n) \geq \log n$, $\mathrm{NC}(2^s)$ denotes the class of languages computable by space $O(s)$-uniform boolean circuits having size $2^{O(s)}$ and depth $s^{O(1)}$ [10]. The class $\mathrm{NC}(2^s)$ is contained in $\mathrm{DSPACE}(s^{O(1)})$ [9]. Thus, it will suffice to prove that (α, β)-QSD is contained in $\mathrm{NC}(2^n)$.

Let (Q_0, Q_1) be an input pair of quantum circuits specifying density matrices (ρ_0, ρ_1) on k qubits, and let n be the length of the description of the pair (Q_0, Q_1). Obviously we may assume $k \leq n$, the number of qubits m on which Q_0 and Q_1 act satisfies $m \leq n$, and each of Q_0 and Q_1 contains at most n gates. We assume Q_0 and Q_1 are composed of gates that can be described by unitary matrices having entries with rational real and imaginary parts (see Section 2.2). Thus, ρ_0 and ρ_1 correspond to $N \times N$ matrices where $N \leq 2^n$, and for each entry of ρ_0 and ρ_1 the numerators and denominators of the real and imaginary parts are $O(n)$-bit integers.

For each $i = 0, 1$ it is possible to compute $|\psi_i\rangle = Q_i|0^m\rangle$ (expressed as a 2^m-dimensional vector with rational real and imaginary parts) in $\mathrm{NC}(2^n)$, simply by computing the product of the matrices corresponding to each individual gate. (In fact, there are better ways to do this from a complexity-theoretic standpoint [15], but this method is sufficient for our needs.) Once these vectors are computed, it is possible to compute $\rho_0 - \rho_1$ in $\mathrm{NC}(2^n)$ by constructing $|\psi_0\rangle\langle\psi_0|$ and $|\psi_1\rangle\langle\psi_1|$, performing the partial trace on the non-output qubits for each matrix (which involves computing a sum of at most 2^n matrices, each of which is obtained by multiplying $|\psi_i\rangle\langle\psi_i|$ on the left and on the right by a $2^k \times 2^m$ or $2^m \times 2^k$ matrix, respectively, as in the definition of the partial trace), and then computing the difference of the resulting matrices. Once we have $\rho_0 - \rho_1$, we may use the method described in Proposition 21 to compute $\|\rho_0 - \rho_1\|_{\mathrm{tr}}$ in $\mathrm{NC}(2^n)$ (which is NC with respect to the size of $\rho_0 - \rho_1$). Since it is only required that the cases $\|\rho_0 - \rho_1\|_{\mathrm{tr}} \leq \alpha$ and $\|\rho_0 - \rho_1\|_{\mathrm{tr}} \geq \beta$ be discriminated, $\|\rho_0 - \rho_1\|_{\mathrm{tr}}$ need in fact only be computed to $O(1)$ bits of precision. This completes the proof. ■

7 Conclusion

In this paper we have given a definition for honest verifier quantum statistical zero-knowledge, and studied the resulting complexity class $\mathrm{QSZK_{HV}}$. Figure 6 shows some relationships among this class, other classes based on quantum interactive proofs, and a few other well-studied classes.

We conclude by mentioning a few open questions concerning quantum zero-knowledge.

- What are other possible definitions for quantum zero-knowledge, and how do they compare to our definition? In particular, how does our definition for honest verifier quantum statistical zero-knowledge compare to possible definitions for (not necessarily honest verifier) quantum statistical zero-knowledge?

- What further relations among $\mathrm{QSZK_{HV}}$ and other complexity classes can be shown? Almost certainly the upper-bound of PSPACE on $\mathrm{QSZK_{HV}}$ can be improved. Is it possible that $\mathrm{NP} \subseteq \mathrm{QSZK_{HV}}$, or do unexpected consequences result from such an assumption?

- The Quantum State Distinguishability problem is natural from the perspective of quantum information processing, but is rather unnatural outside of this scope. Are there more natural problems that are candidates for problems in $\mathrm{QSZK_{HV}}$ but not in SZK?

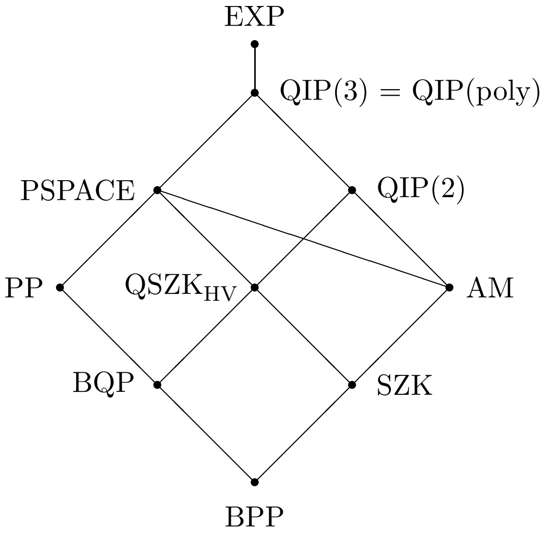

Figure 6: Diagram of complexity classes

Acknowledgments

I thank Gilles Brassard and Claude Crépeau for several helpful discussions on quantum zero-knowledge. This work was supported by Canada's NSERC and the Canada Research Chairs program.

References

[1] S. Aaronson. Quantum lower bound for the collision problem. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, 2002.

[2] L. Adleman, J. DeMarrais, and M. Huang. Quantum computability. *SIAM Journal on Computing*, 26(5):1524–1540, 1997.

[3] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, pages 20–30, 1998.

[4] W. Aiello and J. Håstad. Statistical zero-knowledge languages can be recognized in two rounds. *Journal of Computer and System Sciences*, 42(3):327–345, 1991.

[5] A. Ambainis. Quantum lower bounds by quantum arguments. *Journal of Computer and System Sciences*, 64:750–767, 2002.

[6] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM*, 48(4):778–797, 2001.

[7] C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, 1997.

[8] C. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing*, pages 175–179, 1984.

[9] A. Borodin. On relating time and space to size and depth. *SIAM Journal on Computing*, 6:733–744, 1977.

[10] A. Borodin, S. Cook, and N. Pippenger. Parallel computation for well-endowed rings and space-bounded probabilistic machines. *Information and Control*, 58:113–136, 1983.

[11] K. Cheung and M. Mosca. Decomposing finite Abelian groups. *Quantum Information and Computation*, 1(3):26–32, 2001.

[12] L. Csanky. Fast parallel matrix inversion algorithms. *SIAM Journal on Computing*, 5(4):618–623, 1976.

[13] S. Even, A. Selman, and Y. Yacobi. The complexity of promise problems with applications to public-key cryptography. *Information and Control*, 61(2):159–173, 1984.

[14] L. Fortnow. The complexity of perfect zero-knowledge. In S. Micali, editor, *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 327–343. JAI Press, Greenwich, 1989.

[15] L. Fortnow and J. Rogers. Complexity limitations on quantum computation. *Journal of Computer and System Sciences*, 59(2):240–252, 1999.

[16] O. Goldreich. Zero-knowledge twenty years after its invention. Electronic Colloquium on Computational Complexity (http://www.eccc.uni-trier.de/eccc/), Report No. 63, 2002.

[17] O. Goldreich, S. Micali, and A. Wigderson. Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. *Journal of the ACM*, 38(1):691–729, 1991.

[18] O. Goldreich, A. Sahai, and S. Vadhan. Honest verifier statistical zero knowledge equals general statistical zero knowledge. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 23–26, 1998.

[19] O. Goldreich and S. Vadhan. Comparing entropies in statistical zero-knowledge with applications to the structure of SZK. In *Proceedings of the 14th Annual IEEE Conference on Computational Complexity*, pages 54–73, 1999.

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

[21] J. van de Graaf. *Towards a formal definition of security for quantum protocols*. PhD thesis, Université de Montréal, 1997.

[22] S. Hallgren. Polynomial-time quantum algorithms for Pell's equation and the principal ideal problem. In *Proceedings of the 34th ACM Symposium on Theory of Computing*, 2002.

[23] G. Ivanyos, F. Magniez, and M. Santha. Efficient algorithms for some instances of the non-Abelian hidden subgroup problem. In *Thirteenth ACM Symposium on Parallel Algorithms and Architectures*, 2001.

[24] A. Kitaev. Quantum measurements and the Abelian stabilizer problem. Available as arXiv.org e-Print quant-ph/9511026, 1995.

[25] A. Kitaev, A. Shen, and M. Vyalyi. *Classical and Quantum Computation*, volume 47 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.

[26] A. Kitaev and J. Watrous. Parallelization, amplification, and exponential time simulation of quantum interactive proof system. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, pages 608–617, 2000.

[27] D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78(17):3414–3417, 1997.

[28] M. Mosca. *Quantum Computer Algorithms*. PhD thesis, University of Oxford, 1999.

[29] A. Nayak and P. Shor. Bit-commitment based coin flipping. Available as arXiv.org e-Print quant-ph/0206123, 2002.

[30] C. A. Neff. Specified precision polynomial root isolation is in NC. *Journal of Computer and System Sciences*, 48(3):429–463, 1994.

[31] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[32] T. Okamoto. On relationships between statistical zero-knowledge proofs. *Journal of Computer and System Sciences*, 60(1):47–108, 2000.

[33] R. Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, pages 358–376, 1999.

[34] A. Sahai and S. Vadhan. A complete promise problem for statistical zero-knowledge. In *Proceedings of the 38th Annual IEEE Symposium on the Foundations of Computer Science*, pages 448–457, 1997. Full version available at http://www.eecs.harvard.edu/~salil/research.html.

[35] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[36] P. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters*, 85(2):441–444, 2000.

[37] R. Spekkens and T. Rudolph. Degrees of concealment and bindingness in quantum bit-commitment protocols. *Physical Review A*, 65: article 123410, 2002.

[38] S. Vadhan. *A Study of Statistical Zero-Knowledge Proofs*. PhD thesis, Massachusetts Institute of Technology, August 1999.

[39] J. Watrous. PSPACE has constant-round quantum interactive proof systems. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 112–119, 1999.

[40] J. Watrous. Quantum algorithms for solvable groups. In *Proceedings of the 33rd ACM Symposium on Theory of Computing*, pages 60–67, 2001.