Two-message quantum interactive proofs are in PSPACE

Rahul Jain* Sarvagya Upadhyay† John Watrous†

May 7, 2009

Abstract

We prove that QIP(2), the class of problems having two-message quantum interactive proof systems, is a subset of PSPACE. This relationship is obtained by means of an efficient parallel algorithm, based on the multiplicative weights update method, for approximately solving a certain class of semidefinite programs.

1 Introduction

Since their introduction roughly 25 years ago [Bab85, GMR85], interactive proof systems have become a fundamental notion in the theory of computational complexity. The expressive power of one of the most basic variant of the interactive proof system model, wherein a polynomial-time probabilistic verifier interacts with a computationally unbounded prover for a polynomial number of rounds, is characterized [LFKN92, Sha92] by the well-known relationship

$$\mathrm{IP} = \mathrm{PSPACE}.$$

Many variants of interactive proof systems have been studied, including public-coin interactive proof systems (or Arthur–Merlin games) [Bab85, BM88, GS89], zero-knowledge interactive proofs [GMR89, GMW91] and multi-prover interactive proofs [BOGKW88].

This paper is concerned primarily with quantum interactive proof systems, which are defined in a similar way to ordinary interactive proof systems except that the prover and verifier may perform quantum computations. Like their classical analogues, several variants of quantum interactive proof systems have been studied, including ordinary quantum interactive proofs [Wat03, KW00], public-coin quantum interactive proofs [MW05], zero-knowledge quantum interactive proofs [Wat02, Wat06, Kob08, HKSZ08], and multi-prover quantum interactive proofs [KM03, KKMV08]. The complexity class QIP of problems having quantum interactive proof systems is known [KW00] to satisfy

$$\mathrm{PSPACE} \subseteq \mathrm{QIP} \subseteq \mathrm{EXP}.$$

The containment $\mathrm{QIP} \subseteq \mathrm{EXP}$ follows from the existence of polynomial-time algorithms for approximately solving semidefinite programs [GLS93]. (Somewhat simpler proofs of the containment $\mathrm{QIP} \subseteq \mathrm{EXP}$ follow from the results of [BATS09, Wat09b], but these proofs still require efficient algorithms for solving convex/semidefinite programs.)

Quantum interactive proof systems have an interesting property that classical interactive proof systems are conjectured not to hold, which is that they can be parallelized to a constant number

*Department of Computer Science and Centre for Quantum Technologies, National University of Singapore.
†School of Computer Science and Institute for Quantum Computing, University of Waterloo.

of rounds of interaction [KW00]. More precisely, it holds that $\mathrm{QIP}(3) = \mathrm{QIP}$, where in general $\mathrm{QIP}(m)$ denotes the class of problems having quantum interactive proof systems in which m messages are exchanged between the prover and verifier (with the prover always sending the last message). This leaves four basic classes that are defined naturally by quantum interactive proof systems: $\mathrm{QIP}(0) = \mathrm{BQP}$, $\mathrm{QIP}(1) = \mathrm{QMA}$, $\mathrm{QIP}(2)$, and $\mathrm{QIP}(3) = \mathrm{QIP}$. Of these classes, $\mathrm{QIP}(2)$ seems to be the most mysterious. It is known that $\oplus\mathrm{MIP}^* \subseteq \mathrm{QIP}(2)$ [Weh06] and $\mathrm{QSZK} \subseteq \mathrm{QIP}(2)$ [Wat02, Wat06]. Here, $\oplus\mathrm{MIP}^*$ denotes the class of problems having one-round two-prover classical interactive proof systems in which the provers share quantum entanglement, answer one bit each, and the verifier accepts or rejects based on the parity of these bits; and QSZK denotes the class of problems having statistical zero-knowledge quantum interactive proof systems. No upper bound other than the trivial containment $\mathrm{QIP}(2) \subseteq \mathrm{QIP}$, which implies $\mathrm{QIP}(2) \subseteq \mathrm{EXP}$, was previously known.

In this paper we prove the new containment:

$$\mathrm{QIP}(2) \subseteq \mathrm{PSPACE}.$$

Similar to $\mathrm{QIP} \subseteq \mathrm{EXP}$, this containment is proved using semidefinite programming; but this time the containment is achieved by using an NC algorithm rather than a sequential polynomial-time algorithm. Our algorithm is based on the *multiplicative weights update* method, which was developed by several researchers and is described in the survey [AHK05] and in the PhD thesis of Kale [Kal07]. In particular, our algorithm is based on a general method that was independently discovered by Arora and Kale [AK07] and Warmuth and Kuzmin [WK06]. The key aspect of this approach that makes it useful for proving $\mathrm{QIP}(2) \subseteq \mathrm{PSPACE}$ is its parallelizability: it is an iterative method in which each iteration is easily parallelized, and is such that only a very small number of iterations is needed for an approximation that is accurate enough for our needs. A related approach was used by two of us [JW09] to prove the containment of a different quantum complexity class (called $\mathrm{QRG}(1)$) in PSPACE, but the specific technical details of the simulations are rather different.

The rest of this paper has the following structure. We begin with Section 2, which includes a brief discussion of background information needed for the rest of the paper, including linear algebra notation and parallel algorithms for matrix computations. Section 3 introduces two message quantum interactive proof systems and establishes a simple fact concerning their robustness with respect to error bounds. In Section 4 we present a semidefinite programming formulation of the maximum probability with which a verifier in a two-message quantum interactive proof system can be made to accept, and the actual simulation of $\mathrm{QIP}(2)$ in PSPACE is split into the three sections that follow: Section 5 presents an overview of the simulation, while Sections 6 and 7 describe in more detail its two most technical parts. The paper concludes with Section 8.

2 Preliminaries

2.1 Linear algebra notation and terminology

For complex vector spaces of the form $\mathcal{X} = \mathbb{C}^N$ and $\mathcal{Y} = \mathbb{C}^M$, we write $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ to denote the space of linear operators mapping $\mathcal{X}$ to $\mathcal{Y}$, which is identified with the set of $M \times N$ complex matrices in the usual way. An inner product on $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ is defined as $\langle A, B \rangle = \mathrm{Tr}(A^* B)$ for all $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, where A^* denotes the adjoint (or conjugate transpose) of A. The notation $\mathrm{L}(\mathcal{X})$ is shorthand for $\mathrm{L}(\mathcal{X}, \mathcal{X})$, and the identity operator on $\mathcal{X}$ is denoted $\mathbb{1}_{\mathcal{X}}$ (or just $\mathbb{1}$ when $\mathcal{X}$ is understood).

The following special types of operators are relevant to the paper:

1. An operator $A \in \mathrm{L}(\mathcal{X})$ is *Hermitian* if $A = A^*$. We write $\lambda(A) = (\lambda_1(A), \ldots, \lambda_N(A))$ to denote the vector of eigenvalues of A, sorted from largest to smallest: $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_N(A)$.
2. An operator $P \in \mathrm{L}(\mathcal{X})$ is *positive semidefinite* if it is Hermitian and all of its eigenvalues are nonnegative. The set of such operators is denoted $\mathrm{Pos}(\mathcal{X})$. The notation $P \geq 0$ also indicates that P is positive semidefinite, and more generally the notations $A \leq B$ and $B \geq A$ indicate that $B - A \geq 0$ for Hermitian operators A and B.
3. A positive semidefinite operator $\Pi \in \mathrm{Pos}(\mathcal{X})$ is a *projection* if all of its eigenvalues are either 0 or 1. (Sometimes such an operator is called an *orthogonal projection*, but we have no need to discuss the more general sort of projection.)
4. An operator $\rho \in \mathrm{L}(\mathcal{X})$ is a *density operator* if it is both positive semidefinite and has trace equal to 1. The set of such operators is denoted $\mathrm{D}(\mathcal{X})$.
5. An operator $U \in \mathrm{L}(\mathcal{X})$ is *unitary* if $U^*U = \mathbb{1}_{\mathcal{X}}$. The set of such operators is denoted $\mathrm{U}(\mathcal{X})$.

Three operator norms are discussed in this paper: the *trace norm*, *Frobenius norm*, and *spectral norm*, defined as

$$\|A\|_1 = \mathrm{Tr}\sqrt{A^*A}, \quad \|A\|_2 = \sqrt{\langle A, A\rangle}, \quad \text{and} \quad \|A\| = \max\{\|Au\| : u \in \mathcal{X}, \|u\| = 1\}$$

respectively, for each $A \in \mathrm{L}(\mathcal{X})$. Alternately, these norms are given by the 1, 2 and ∞ norms of the vector of singular values of A. For every operator A it holds that $\|A\| \leq \|A\|_2 \leq \|A\|_1$. We also use the inequalities $|\langle A, B\rangle| \leq \|A\| \, \|B\|_1$, $\|AB\| \leq \|A\| \, \|B\|$, and $\|AB\|_1 \leq \|A\| \, \|B\|_1$ a couple of times in the paper.

The *fidelity function* is defined as

$$\mathrm{F}(P, Q) = \left\| \sqrt{P}\sqrt{Q} \right\|_1$$

for positive semidefinite operators P and Q of equal dimension.

A *super-operator* is a linear mapping of the form $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$, for spaces of the form $\mathcal{X} = \mathbb{C}^N$ and $\mathcal{Y} = \mathbb{C}^M$. The identity super-operator on $\mathrm{L}(\mathcal{X})$ is denoted $\mathbb{1}_{\mathrm{L}(\mathcal{X})}$. The adjoint super-operator to Φ is the unique super-operator $\Phi^* : \mathrm{L}(\mathcal{Y}) \to \mathrm{L}(\mathcal{X})$ for which $\langle Y, \Phi(X)\rangle = \langle \Phi^*(Y), X\rangle$ for all $X \in \mathrm{L}(\mathcal{X})$ and $Y \in \mathrm{L}(\mathcal{Y})$.

The following special types of super-operators are relevant to the paper.

1. $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ is *completely positive* if it holds that $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})})(P) \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{W})$ for every choice of $\mathcal{W} = \mathbb{C}^k$ and $P \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{W})$.
2. $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ is *trace-preserving* if $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$ for every $X \in \mathrm{L}(\mathcal{X})$.
3. $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ is a *quantum operation* (also called an *admissible super-operator* or a *quantum channel*) if it is both completely positive and trace-preserving.

2.2 Remarks on NC and parallel matrix computations

To prove that QIP(2) is contained in PSPACE, we will make use of various facts concerning parallel computation. First, let us recall the definition of two complexity classes based on bounded-depth circuit families: NC and NC(*poly*). The class NC contains all functions that can be computed by logarithmic-space uniform Boolean circuits of polylogarthmic depth, while the class NC(*poly*)

contains all functions that can be computed by polynomial-space uniform families of Boolean circuits having polynomial-depth. By restricting these classes to predicates we obtain classes of languages (or more generally promise problems).

There are two main facts about these classes that we will need. The first fact, which follows from a result of Borodin [Bor77], is that for languages (or promise problems) we have $\mathrm{NC}(\mathit{poly}) \subseteq \mathrm{PSPACE}$. (In fact it holds that $\mathrm{NC}(\mathit{poly}) = \mathrm{PSPACE}$, but we only need a containment in one direction.) The second fact is that functions in these classes compose nicely. In particular, if $F : \{0,1\}^* \to \{0,1\}^*$ is a function in $\mathrm{NC}(\mathit{poly})$ and $G : \{0,1\}^* \to \{0,1\}^*$ is a function in NC, then the composition $G \circ F$ is also in $\mathrm{NC}(\mathit{poly})$. This follows from the most obvious way of composing the families of circuits that compute F and G, along with the observation that $|F(x)|$ can be at most exponential in $|x|$.

Finally, we will make use of the fact that many computations involving matrices can be performed by NC algorithms. We may restrict our attention to matrix computations on matrices whose entries have rational real and imaginary parts. Numbers of this form, $\alpha = (a/b) + i(c/d)$ for integers a, b, c, and d, are sometimes referred to as Gaussian rationals. We assume any number of this form is encoded as a 4-tuple (a, b, c, d) using binary notation, so that the *length* of α is understood to be the total number of bits needed for such an encoding.

It is known that elementary matrix operations, such as additions, multiplications, and inversions can be performed in NC. (The survey [Gat93], for instance, describes NC algorithms for these tasks.) We will also make use of the fact that matrix exponentials and spectral decompositions can be approximated to high precision in NC. In more precise terms, we have that the following problems are in NC:

Matrix exponentials

Input: An $n \times n$ matrix M, a positive rational number ε, and an integer k expressed in unary notation (i.e., 1^k), such that $\|M\| \leq k$.

Output: An $n \times n$ matrix X such that $\|\exp(M) - X\| < \varepsilon$.

Spectral decompositions

Input: An $n \times n$ Hermitian matrix H and a positive rational number ε.

Output: An $n \times n$ unitary matrix U and an $n \times n$ real diagonal matrix Λ such that

$$\|M - U\Lambda U^*\| < \varepsilon.$$

Singular-value decompositions

Input: An $n \times m$ matrix M and a positive rational number ε.

Output: An $n \times r$ matrix U with orthonormal columns, an $m \times r$ matrix V with orthonormal columns, and an $r \times r$ diagonal matrix Σ with positive diagonal entries such that

$$\|M - U\Sigma V^*\| < \varepsilon.$$

Note that in these problems, the description of ε has roughly $\log(1/\varepsilon)$ bits, which means that highly accurate approximations are possible in NC. The fact that matrix exponentials can be approximated in NC as claimed follows by truncating the series

$$\exp(M) = \mathbb{1} + M + M^2/2 + M^3/6 + \cdots$$

to a number of terms polynomial in k and $\log(1/\varepsilon)$. (This is not a very practial way to compute matrix exponentials, but it establishes the fact we need.) The fact that spectral and singular value decompositions can be approximated in NC follows from a composition of known facts: in NC one can compute characteristic polynomials and null spaces of matrices, perform orthogonalizations of vectors, and approximate roots of integer polynomials to high precision [Csa76, BGH82, BCP83, BOFKT86, Gat93, Nef94].

3 Two-message quantum interactive proof systems

The purpose of this section is to introduce the class QIP(2), including its definition and a simple proof that it is robust with respect to error bounds. For a general discussion of quantum interactive proof systems, as opposed to the somewhat simplified case in which only two messages are exchanged, the reader is referred to [KW00] and [Wat09a].

3.1 Definition of two-message quantum interactive proofs

To define the class QIP(2), we begin by defining a *two-message quantum verifier* V as a classical polynomial-time algorithm that, on each input string x, outputs the description of two quantum circuits: U_x and V_x. The circuit U_x describes the verifier's initial preparation of a state, part of which is sent to the prover, while the circuit V_x describes the verifier's actions upon receiving a response from the prover. For the sake of simplicity, and without loss of generality, we assume that for every input string x, the circuits U_x and V_x are both composed of gates from some finite, universal set of unitary quantum gates whose entries have rational real and imaginary parts. The number of qubits on which the circuits U_x and V_x act is assumed to be equal to $2p(n)$, where $n = |x|$ and p is some polynomial-bounded function. The first $p(n)$ qubits represent the communication channel between the prover and verifier, while the remaining $p(n)$ qubits serve as the private memory of the verifier. (It is not really necessary that the number of message qubits and private qubits agree, but it causes no change in the computational power of the model.)

A two-message quantum prover P is simply a collection of quantum operations (or, equivalently, completely positive and trace preserving super-operators) $\{\Psi_x : x \in \{0,1\}^*\}$. Such a prover is compatible with a given verifier V if each operation Φ_x acts on $p(n)$ qubits for the function p mentioned above.

An interaction between a two-message verifier V and a compatible prover P on an input x proceeds as follows:

1. $2p(n)$ qubits are initialized in the $|0\rangle$ state.
2. The circuit U_x is applied to all of the qubits.
3. The prover's operation Ψ_x is applied to the first $p(n)$ qubits.
4. The circuit V_x is applied to all of the qubits.
5. The first qubit is measured in the standard basis, with the outcome 1 indicating *acceptance* and 0 indicating *rejection*.

Figure 1 illustrates such an interaction.

Now, a promise problem $A = (A_{\text{yes}}, A_{\text{no}})$ is in QIP(2) if and only if there exists a two-message verifier V with the following completeness and soundness properties:

1. (Completeness) If $x \in A_{\text{yes}}$, then there exists a prover P that causes V to accept x with probability at least $2/3$.

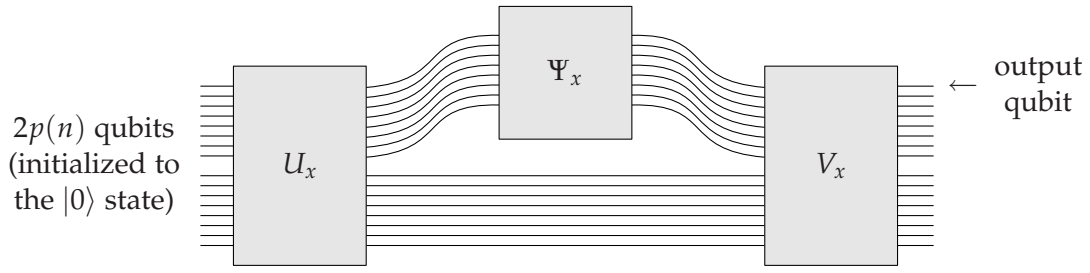

Figure 1: An interaction between a verifier V and a prover P on an input x. The verifier's actions are determined by the circuits U_x and V_x acting on $2p(n)$ qubits, while the prover's action corresponds to the quantum operation Ψ_x on just the first $p(n)$ qubits.

2. (Soundness) If $x \in A_{\text{no}}$, then every prover P that is compatible with V causes V to accept x with probability at most $1/3$.

3.2 Robustness of QIP(2) with respect to error bounds

It was proved in [KW00] that quantum interactive proof systems with negligible completeness error are amenable to parallel repetition. This allows for an exponential reduction in error for quantum interactive proof systems with three or more messages, because such proof systems can be transformed to have perfect completeness by a different method. However, this method does not work for two-message quantum interactive proof systems, because the perfect-completeness transformation requires the addition of messages. So, we will require a different method of error reduction.

Assume that A is a promise problem in QIP(2) and that (V, P) is a two-message quantum interactive proof system for A with completeness and soundness probabilities bounded by a and b, where $a - b \geq 1/q$ for some polynomial-bounded function q. We will define a new verifier V' that has completeness probability at least $1 - 2^{-r}$ and soundness probability at most 2^{-r}, for any choice of a polynomial-bounded function r. A description of V' follows.

1. Let $s = 2rq$ and let $t = 8rq^2s$. Run st independent, parallel executions of the protocol for V, one for each pair (i, j) with $i \in \{1, \ldots, s\}$ and $j \in \{1, \ldots, t\}$. Measure the output qubit for each execution, and let the result of the measurement for execution (i, j) be $y_{i,j} \in \{0, 1\}$.
2. For each $i = 1, \ldots, s$, set
$$z_i = \begin{cases} 1 & \text{if } \sum_{j=1}^{t} y_{i,j} \geq t \cdot \frac{a+b}{2} \\ 0 & \text{otherwise} \end{cases}$$
3. Accept if $\bigwedge_{i=1}^{s} z_i = 1$, reject otherwise.

Now let us consider the maximum probability with which V' can be made to accept. Suppose first that an input $x \in A_{\text{yes}}$ is fixed, so that V is made to accept with probability at least a by the prover P. Our goal is to define a prover P' that causes V' to accept with probability at least $1 - 2^{-r}$. This is easily done by defining P' so that it runs st independent simulations of P. Let $Y_{i,j}$ and Z_i, for $i \in \{1, \ldots, s\}$ and $j \in \{1, \ldots, t\}$, be Boolean-valued random variables corresponding to the values $y_{i,j}$ and z_i when V' interacts with the prover P' just described. Given that P' simulates

st independent copies of P, we have that the random variables $Y_{i,j}$ are independent and satisfy $\mathrm{E}[Y_{i,j}] \geq a$ for each pair (i, j). By the Chernoff Bound, we therefore have

$$\Pr[Z_i = 0] = \Pr\left[Y_{i,1} + \cdots + Y_{i,t} < at\left(1 - \frac{a-b}{2a}\right)\right] \leq \exp\left(-\frac{t}{8a}(a-b)^2\right) \leq e^{-rs},$$

and thus the probability of rejection is at most $se^{-rs} < 2^{-r}$.

Suppose on the other hand that an input $x \in A_{\mathrm{no}}$ is fixed, so that no prover P can convince V to accept with probability greater than b. Fix an arbitrary prover P', and as before let $Y_{i,j}$ and Z_i be Boolean-valued random variables corresponding to $y_{i,j}$ and z_i. There may not be independence among these random variables, as P' may not treat the parallel executions independently. We do know, however, that $\mathrm{E}[Y_{i,j}] \leq b$ for every i, j, given that the maximum acceptance probability of V is b. By Markov's Inequality we therefore have

$$\Pr[Z_i = 1] = \Pr\left[Y_{i,1} + \cdots + Y_{i,t} \geq \frac{t(a+b)}{2}\right] \leq 2\frac{\mathrm{E}[Y_{i,1} + \cdots + Y_{i,t}]}{t(a+b)} < 1 - \frac{1}{2q}.$$

We may view $z_1, \ldots, z_s$ as being the outcomes of s parallel executions of a quantum interactive proof system that accepts with probability at most $1 - \frac{1}{2q}$. The verifier V' accepts if and only if all of these executions accept, and so by the result on parallel repetition proved in [KW00] we may conclude that the probability of acceptance of V' is at most

$$\left(1 - \frac{1}{2q}\right)^s < \exp\left(-\frac{s}{2q}\right) < 2^{-r}.$$

Thus, the verifier V' has been shown to have completeness and soundness probabilities as required, completing the proof.

4 Maximum acceptance probability as a semidefinite program

The maximum acceptance probability of a verifier in a quantum interactive proof system can be phrased as semidefinite programming problem [KW00]. For this paper a semidefinite programming formulation based on ones described in [GW07, Wat09b] will be used.

Suppose V is a two-message verifier, and that an input string x of length n is being considered. Let us also define

$$|\psi\rangle = U_x\,|0^{2p(n)}\rangle \qquad \text{and} \qquad \Pi = V_x^*(|1\rangle\langle 1| \otimes \mathbb{1})V_x.$$

In words, $|\psi\rangle$ denotes the quantum state initially prepared by the verifier, the first half of which is sent to the prover; and Π denotes the projection operator, corresponding to the *accept* outcome of the measurement that the verifier effectively performs after receiving the prover's message.

For convenience, we will assign distinct names to the complex vector spaces that arise from an interaction between V on input x and a compatible prover operation Ψ. Specifically, let $\mathcal{X}$ denote the space corresponding to the verifier's message to the prover, let $\mathcal{Y}$ denote the space corresponding to the prover's response, and let $\mathcal{Z}$ denote the space corresponding to the verifier's private qubits. Thus, it holds that $|\psi\rangle \in \mathcal{X} \otimes \mathcal{Z}$ and Π is a projection on $\mathcal{Y} \otimes \mathcal{Z}$. When the prover applies the operation $\Psi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$, the verifier accepts with probability

$$\Big\langle \Pi, (\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(|\psi\rangle\langle\psi|) \Big\rangle. \tag{1}$$

To express the maximum probability for V to accept, over all choices of an operation Ψ, as a semidefinite program, it is helpful to recall the *Choi-Jamiołkowski representation* of super-operators. Let us take $\{|0\rangle, \ldots, |N-1\rangle\}$ to be the standard basis of $\mathcal{X}$. Then the Choi-Jamiołkowski representation of $\Psi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ is the operator $J(\Psi) \in \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$ defined by

$$J(\Psi) = \sum_{0 \leq i,j \leq N-1} \Psi(|i\rangle\langle j|) \otimes |i\rangle\langle j| .$$

It holds that Ψ is completely positive if and only if $J(\Psi)$ is positive semidefinite [Jam72, Cho75], and Ψ is trace-preserving if and only if $\mathrm{Tr}_{\mathcal{Y}}(J(\Psi)) = \mathbb{1}_{\mathcal{X}}$.

Now let us write

$$|\psi\rangle = \sum_{0 \leq i \leq N-1} |i\rangle |\psi_i\rangle$$

for vectors $|\psi_0\rangle, \ldots, |\psi_{N-1}\rangle \in \mathcal{Z}$, and define $B \in \mathrm{L}(\mathcal{X}, \mathcal{Z})$ as

$$B = \sum_{0 \leq i \leq N-1} |\psi_i\rangle \langle i| .$$

Then it is clear that

$$(\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(|\psi\rangle\langle\psi|) = (\mathbb{1}_{\mathcal{Y}} \otimes B) J(\Psi) (\mathbb{1}_{\mathcal{Y}} \otimes B^*).$$

We therefore find that the probability of acceptance (1) may alternately be written

$$\langle \Pi, (\mathbb{1}_{\mathcal{Y}} \otimes B) J(\Psi) (\mathbb{1}_{\mathcal{Y}} \otimes B^*) \rangle = \langle (\mathbb{1}_{\mathcal{Y}} \otimes B^*) \Pi (\mathbb{1}_{\mathcal{Y}} \otimes B), J(\Psi) \rangle = \langle Q, J(\Psi) \rangle$$

for $Q = (\mathbb{1}_{\mathcal{Y}} \otimes B^*) \Pi (\mathbb{1}_{\mathcal{Y}} \otimes B)$. We call Q the *interactive measurement operator* that is determined by V on input x. It is clear that the interactive measurement operator Q is positive semidefinite, and moreover that $Q \leq \mathbb{1}_{\mathcal{Y}} \otimes \xi$ for the density operator $\xi = B^* B$.

Now, let us define

$$\mu(Q) = \max_{\Psi} \langle Q, J(\Psi) \rangle ,$$

where the maximum is over all valid quantum operations of the form $\Psi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$. The quantity $\mu(Q)$ will be called the *maximum acceptance probability* of Q, as this value is precisely the maximum acceptance probability of the verifier V on input x, whose description alone has led us to the definition of Q. As stated above, when $\Psi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ ranges over the set of all valid quantum operations, $J(\Psi)$ ranges over the set of positive semidefinite operators satisfying the linear constraint $\mathrm{Tr}_{\mathcal{Y}}(J(\Psi)) = \mathbb{1}_{\mathcal{X}}$. This implies that the quantity $\mu(Q)$ is represented by a semidefinite program:

$$\begin{array}{ll} \text{maximize:} & \langle Q, X \rangle \\ \text{subject to:} & \mathrm{Tr}_{\mathcal{Y}}(X) \leq \mathbb{1}_{\mathcal{X}}, \\ & X \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X}) . \end{array}$$

The feasible region of this semidefinite program is well-bounded (in the sense of [GLS93]), and therefore its optimal value $\mu(Q)$ can be approximated to high precision in time polynomial in the size of Q (which is exponential in $|x|$). This fact does not help us to prove $\mathrm{QIP}(2) \subseteq \mathrm{PSPACE}$, however. As is described in the next section, we will need an NC algorithm rather than just a polynomial-time algorithm to draw this conclusion.

It will be necessary for us to rephrase the semidefinite program above, and to explicitly state its dual program. As will be discussed shortly, we will only need to consider this formulation for

invertible interactive measurement operators, so Q is hereafter assumed to be invertible. Define a super-operator $\Phi : \mathrm{L}(\mathcal{Y} \otimes \mathcal{X}) \to \mathrm{L}(\mathcal{X})$ as

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Y}} \left(Q^{-1/2} X Q^{-1/2} \right).$$

The adjoint super-operator $\Phi^* : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$ to Φ is given by

$$\Phi^*(Y) = Q^{-1/2} (\mathbb{1}_{\mathcal{Y}} \otimes Y) Q^{-1/2}.$$

The value $\mu(Q)$ is then seen to be the optimal value of the following semidefinite program:

Primal problem		Dual problem	
maximize:	$\mathrm{Tr}(X)$	minimize:	$\mathrm{Tr}(Y)$
subject to:	$\Phi(X) \leq \mathbb{1}_{\mathcal{X}}$,	subject to:	$\Phi^*(Y) \geq \mathbb{1}_{\mathcal{Y} \otimes \mathcal{X}}$,
	$X \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$.		$Y \in \mathrm{Pos}(\mathcal{X})$.

Strong duality follows from strict feasibility, which is easily verified, and so the optimal primal and dual solutions are given by $\mu(Q)$.

5 Overview of the simulation

We will now explain, in high-level terms, our simulation of QIP(2) in PSPACE. To prove that QIP(2) $\subseteq$ PSPACE, it will suffice to prove QIP(2) $\subseteq$ NC(*poly*). This will be facilitated by the fact, discussed in Section 2.2, that many computations involving matrices, including elementary operations such as addition, multiplication, and inversion, as well as approximations of spectral decompositions, singular-value decompositions, and matrix exponentials, can be performed in NC.

For the remainder of this paper, assume that $A = (A_{\mathrm{yes}}, A_{\mathrm{no}})$ is an arbitrary promise problem in QIP(2), and let V be a two-message verifier for A that has exponentially small completeness and soundness error. The goal of the simulation is to determine whether or not V can be made to accept a given input string x with high probability. The variable n will always denote the input length $n = |x|$, and $p(n)$ will denote the number of qubits exchanged by the verifier and prover on each of the two messages as discussed in Section 3.

There are three main steps of the simulation:

1. Compute from x an explicit description of $|\psi\rangle$ and Π.
2. Process the description of the vector $|\psi\rangle$ and the projection Π into a well-conditioned interactive measurement operator Q and positive rational numbers γ and ε satisfying

   $$\begin{aligned} x \in A_{\mathrm{yes}} &\;\Rightarrow\; \mu(Q) \geq (1 + 4\varepsilon)\gamma, \\ x \in A_{\mathrm{no}} &\;\Rightarrow\; \mu(Q) \leq (1 - 4\varepsilon)\gamma. \end{aligned}$$

   For some polynomial q it will hold that $\kappa(Q) \leq q(n)$, where $\kappa(Q) = \|Q\| \, \|Q^{-1}\|$ denotes the condition number of Q. Moreover it will hold that $1/q(n) \leq \varepsilon$ and $1/q(n) \leq \gamma$.
3. Use a parallel algorithm, based on the multiplicative weights update method, to test whether $\mu(Q)$ is larger or smaller than γ.

The first step is easily performed in NC($poly$), using an exact computation. In particular, one may simply compute products of the matrices that describe the individual gates of the verifier's circuits. Given that this step is straightforward, we will not comment on it further. The second and third steps are more complicated, and are described separately in Sections 6 and 7 below. Both correspond to NC computations (where the input size is exponential in n), and by composing these computations with the first step just described, we will obtain that A is in NC($poly$), and therefore QIP(2) $\subseteq$ PSPACE.

6 Preparing a well-conditioned interactive measurement operator

After the first step of the simulation, we have a unit vector $|\psi\rangle$ and a projection operator Π. Let us write $M = 2^{p(n)}$ to denote the dimension of both of the message spaces and the verifier's private work space defined by V on input x, and let us also define $\mathcal{X}_0 = \mathbb{C}^M$, $\mathcal{Y} = \mathbb{C}^M$, and $\mathcal{Z}_0 = \mathbb{C}^M$. We view that the space $\mathcal{X}_0$ corresponds to the verifier's message to the prover, that $\mathcal{Y}$ corresponds to the prover's message back to the verifier, and that $\mathcal{Z}_0$ represents the verifier's private workspace; and thus $|\psi\rangle \in \mathcal{X}_0 \otimes \mathcal{Z}_0$ and $\Pi \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{Z}_0)$. The reason for subscripting $\mathcal{X}_0$ and $\mathcal{Z}_0$ with 0 is that we are viewing these as initial choices of spaces. In the processing of $|\psi\rangle$ and Π, we will define an interactive measurement operator Q over spaces $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ where $\mathcal{X} = \mathbb{C}^N$ and $\mathcal{Z} = \mathbb{C}^N$ for some choice of a positive integer $N \leq M$.

Along the same lines as was discussed in Section 3, we may define an interactive measurement operator $R \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X}_0)$ as $R = (\mathbb{1}_{\mathcal{Y}} \otimes B^*)\Pi(\mathbb{1}_{\mathcal{Y}} \otimes B)$, for

$$|\psi\rangle = \sum_{0 \leq i \leq M-1} |i\rangle\,|\psi_i\rangle \qquad \text{and} \qquad B = \sum_{0 \leq i \leq M-1} |\psi_i\rangle\,\langle i|\,.$$

The quantity $\mu(R)$ is precisely the maximum acceptance probability of V on input x, but nothing can be said about the condition number of R (which may not even be invertible).

Our goal is to compute a new measurement operator $Q \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$, where $\mathcal{X} = \mathbb{C}^N$ for some choice of $N \leq M$, along with positive rational numbers γ and ε, such that the following properties hold for some polynomial q:

1. The interactive measurement operator Q is well-conditioned: $\kappa(Q) \leq q(n)$.
2. The values γ and ε are non-negligible: $1/q(n) \leq \varepsilon$ and $1/q(n) \leq \gamma$.
3. The value $\mu(Q)$ satisfies the properties

$$\begin{aligned} x \in A_{\mathrm{yes}} &\;\Rightarrow\; \mu(Q) \geq (1 + 4\varepsilon)\gamma, \\ x \in A_{\mathrm{no}} &\;\Rightarrow\; \mu(Q) \leq (1 - 4\varepsilon)\gamma. \end{aligned} \tag{2}$$

The first step in this process is to replace $|\psi\rangle$ by a more uniform vector $|\phi\rangle \in \mathcal{X}_0 \otimes \mathcal{Z}_0$ that is "similar enough" to $|\psi\rangle$ in the sense to be described. We will, in particular, take $|\phi\rangle$ to be maximally entangled over certain subspaces of $\mathcal{X}_0$ and $\mathcal{Z}_0$. This is done by performing the following operations:

1. Let
$$|\psi\rangle = \sum_{0 \leq j \leq M-1} \sqrt{\lambda_j}\,|x_j\rangle\,|z_j\rangle$$
be a Schmidt decomposition of $|\psi\rangle$.

2. For each positive integer i, define the interval $I_i = \left(2^{-i}, 2^{-i+1}\right]$, and define

$$\Sigma_i = \left\{ j \in \{0, \ldots, M-1\} \,:\, \lambda_j \in I_i \right\}.$$

3. Let $k = p(n) + 1$ and choose $i \in \{1, \ldots, k\}$ so that

$$\sum_{j \in \Sigma_i} \lambda_j \geq \frac{1}{2k}.$$

   The fact that such an i exists is proved below, and hereafter we write $\Sigma = \Sigma_i$ for this choice of i.

4. Define

$$|\phi\rangle = \frac{1}{\sqrt{|\Sigma|}} \sum_{j \in \Sigma} |x_j\rangle\, |z_j\rangle\,.$$

Now, consider the interactive measurement operator $S \in \mathrm{Pos}\left(\mathcal{Y} \otimes \mathcal{X}_0\right)$ that is obtained by replacing $|\psi\rangle$ with $|\phi\rangle$ (with Π unchanged). In other words, S is defined by the same process as R (which is determined by $|\psi\rangle$ and Π as described above), and satisfies the equation

$$\langle S, J(\Psi)\rangle = \Big\langle \Pi, (\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)})(|\phi\rangle\langle\phi|) \Big\rangle$$

for every super-operator $\Psi : \mathrm{L}\left(\mathcal{X}_0\right) \to \mathrm{L}\left(\mathcal{Y}\right)$. We will prove that

$$\mu(R) - \left(1 - \frac{1}{8k}\right) \leq \mu(S) \leq 4k\,\mu(R). \tag{3}$$

These are fairly loose bounds—but for the two extremes where $\mu(R)$ is exponentially close to 0 or 1, the corresponding values for $\mu(S)$ will be separated by the reciprocal of a polynomial, which is good enough for our needs.

First, note that

$$\sum_{i>k} \sum_{j \in \Sigma_i} \lambda_j \leq M 2^{-k} \leq \frac{1}{2}$$

and therefore

$$\sum_{i=1}^{k} \sum_{j \in \Sigma_i} \lambda_j \geq \frac{1}{2}.$$

Thus, there must exist a suitable choice of i in step 3 so that

$$\sum_{j \in \Sigma_i} \lambda_j \geq \frac{1}{2k},$$

as claimed. A lower bound on the size of $\Sigma = \Sigma_i$ may be obtained by noting that

$$\sum_{j \in \Sigma} \lambda_j \leq 2^{-i+1}\,|\Sigma|$$

and therefore

$$|\Sigma| \geq \frac{2^i}{4k}.$$

Now, the inner product of $|\psi\rangle$ and $|\phi\rangle$ is easily bounded from below as

$$\langle\phi|\psi\rangle = \sum_{j\in\Sigma} \sqrt{\frac{\lambda_j}{|\Sigma|}} \geq |\Sigma| \sqrt{\frac{2^{-i}}{|\Sigma|}} = \sqrt{|\Sigma|\, 2^{-i}} \geq \frac{1}{\sqrt{4k}},$$

and therefore

$$\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_1 = 2\sqrt{1 - |\langle\phi|\psi\rangle|^2} \leq 2\sqrt{1 - \frac{1}{4k}} \leq 2\left(1 - \frac{1}{8k}\right).$$

For every choice of an admissible super-operator $\Psi : \mathrm{L}(\mathcal{X}_0) \to \mathrm{L}(\mathcal{Y})$ it holds that

$$0 \leq \left(\Psi^* \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)}\right)(\Pi) \leq \mathbb{1}_{\mathcal{X}_0 \otimes \mathcal{Z}_0},$$

and by combining this observation with the fact that $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$ is traceless we obtain

$$\langle R - S, J(\Psi)\rangle = \left\langle \left(\Psi^* \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)}\right)(\Pi), |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \right\rangle \leq \frac{1}{2} \| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_1 \leq 1 - \frac{1}{8k}.$$

Therefore

$$\mu(R) - \mu(S) \leq 1 - \frac{1}{8k},$$

which establishes the lower bound on $\mu(S)$ claimed in (3) above.

To establish the upper bound on $\mu(S)$, which is the second inequality in (3), let us first choose an admissible super-operator Ψ so that

$$\mu(S) = \left\langle \Pi, \left(\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)}\right)(|\phi\rangle\langle\phi|) \right\rangle.$$

Now, observe that

$$\frac{1}{4k}\frac{1}{|\Sigma|} \leq 2^{-i} < \lambda_j$$

for each $j \in \Sigma$, and thus

$$\frac{1}{4k} \operatorname{Tr}_{\mathcal{X}_0}(|\phi\rangle\langle\phi|) \leq \operatorname{Tr}_{\mathcal{X}_0}(|\psi\rangle\langle\psi|).$$

It is therefore possible to choose a density operator $\xi \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Z}_0)$ such that

$$\left(1 - \frac{1}{4k}\right) \operatorname{Tr}_{\mathcal{Y}}(\xi) = \operatorname{Tr}_{\mathcal{X}_0}(|\psi\rangle\langle\psi|) - \frac{1}{4k} \operatorname{Tr}_{\mathcal{X}_0}(|\phi\rangle\langle\phi|).$$

Because Ψ is admissible, we may therefore conclude that

$$\operatorname{Tr}_{\mathcal{X}_0}(|\psi\rangle\langle\psi|) = \operatorname{Tr}_{\mathcal{Y}}\left(\frac{1}{4k}\left(\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)}\right)(|\phi\rangle\langle\phi|) + \left(1 - \frac{1}{4k}\right)\xi\right),$$

and so there must exist an admissible super-operator $\Xi : \mathrm{L}(\mathcal{X}_0) \to \mathrm{L}(\mathcal{Y})$ so that

$$\left(\Xi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)}\right)(|\psi\rangle\langle\psi|) = \frac{1}{4k}\left(\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)}\right)(|\phi\rangle\langle\phi|) + \left(1 - \frac{1}{4k}\right)\xi.$$

Consequently we have

$$\mu(R) \geq \langle R, J(\Xi)\rangle = \left\langle \Pi, \left(\Xi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)}\right)(|\psi\rangle\langle\psi|) \right\rangle \geq \frac{1}{4k} \left\langle \Pi, \left(\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_0)}\right)(|\phi\rangle\langle\phi|) \right\rangle = \frac{1}{4k}\mu(S)$$

as required.

Having obtained a uniform vector $|\phi\rangle$ that gives (when combined with Π) an interactive measurement operator S satisfying (3), we must make a couple of additional modifications to be sure that a well-conditioned interactive measurement operator has been obtained.

First, we will replace $\mathcal{X}_0$ and $\mathcal{Z}_0$ with the spaces $\mathcal{X} = \mathbb{C}^N$ and $\mathcal{Z} = \mathbb{C}^N$ for $N = |\Sigma|$. Let $X \in \mathrm{U}(\mathcal{X}, \mathcal{X}_0)$ and $Z \in \mathrm{U}(\mathcal{Z}, \mathcal{Z}_0)$ be linear isometries defined as

$$X = \sum_{j=1}^{N} |x_j\rangle\langle j| \qquad \text{and} \qquad Z = \sum_{j=1}^{N} |z_j\rangle\langle j|,$$

and define

$$|\tau\rangle = (X^* \otimes Z^*)\,|\phi\rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} |j\rangle\,|j\rangle \qquad \text{and} \qquad P = (\mathbb{1}_{\mathcal{Y}} \otimes Z^*)\,\Pi\,(\mathbb{1}_{\mathcal{Y}} \otimes Z).$$

It is clear that $|\tau\rangle$ is a unit vector and P is an ordinary measurement operator. (It might not be that P is a projection operator, but it is positive semidefinite and satisfies $P \leq \mathbb{1}_{\mathcal{Y}\otimes\mathcal{Z}}$.) Finally, let $Q \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$ be the interactive measurement operator defined by the vector $|\tau\rangle$ and the ordinary measurement operator

$$\left(1 - \frac{1}{64k}\right) P + \frac{1}{64k}\mathbb{1}_{\mathcal{Y}\otimes\mathcal{Z}}.$$

It holds that

$$\mu(Q) = \left(1 - \frac{1}{64k}\right)\mu(S) + \frac{1}{64k},$$

and therefore

$$\mu(R) - \left(1 - \frac{1}{8k}\right) \leq \mu(Q) \leq 4k\mu(R) + \frac{1}{64k}. \tag{4}$$

Now let us verify that Q has the properties we require of it. First let us consider the condition number $\kappa(Q)$. It is easily shown that

$$Q \geq \frac{1}{64kN}\mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}} \qquad \text{and} \qquad Q \leq \frac{1}{N}\mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}},$$

and therefore $\kappa(Q) \leq 64k$. It remains to define nonnegligible values γ and ε, and to consider their relationship to $\mu(Q)$ in the two cases: $x \in A_{\text{yes}}$ and $x \in A_{\text{no}}$.

We have assumed that the original interactive proof system has exponentially small completeness and soundness errors, and therefore we may assume that for sufficiently large n we have

$$x \in A_{\text{no}} \Rightarrow \mu(R) \leq \frac{1}{256\,k^2} \qquad \text{and} \qquad x \in A_{\text{yes}} \Rightarrow \mu(R) \geq 1 - \frac{1}{32k}.$$

(Alternately we may assume these bounds hold for all n by hard-coding small inputs x into the verifier.) Thus, by the bounds (4) above, we have

$$x \in A_{\text{no}} \Rightarrow \mu(Q) \leq \frac{1}{32k} \qquad \text{and} \qquad x \in A_{\text{yes}} \Rightarrow \mu(Q) \geq \frac{1}{16k}.$$

By taking

$$\gamma = \frac{3}{64k} = \frac{3}{64(p(n)+1)}, \quad \varepsilon = \frac{1}{12}, \quad \text{and} \quad q(n) \geq 64k = 64(p(n)+1),$$

we therefore have the properties required.

1. Let
$$\delta = \frac{\varepsilon^2}{8\,\kappa(Q)^2} \qquad \text{and} \qquad T = \left\lceil \frac{24\,\ln(NM)}{\varepsilon^3\gamma^3\delta} \right\rceil.$$
2. Let $W_0 = \mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}}$ and let $\rho_0 = W_0/\operatorname{Tr}(W_0)$.
3. For $t = 0, \ldots, T-1$ do:
   a. Compute a spectral decomposition of $\Phi(\rho_t)$:
   $$\Phi(\rho_t) = \sum_{j=1}^{N} \lambda_j \left|x_j\right\rangle\!\left\langle x_j\right|.$$
   b. Let $S = \{j \in \{1,\ldots,N\} \,:\, \gamma\lambda_j > 1\}$ and let $s = \sum_{j\in S}\lambda_j$.
   c. If $s \leq \delta \left\|Q^{-1}\right\|$, then halt and *accept*.
   d. Let
   $$Y_t = \frac{1}{s}\sum_{j\in S} \left|x_j\right\rangle\!\left\langle x_j\right|.$$
   e. Let
   $$W_{t+1} = \exp\left(-\frac{\varepsilon\gamma\delta}{2}\,\Phi^*(Y_0 + \cdots + Y_t)\right),$$
   and let $\rho_{t+1} = W_{t+1}/\operatorname{Tr}(W_{t+1})$.
4. Halt and *reject*.

Figure 2: Algorithm to test if $\mu(Q) \geq \gamma$.

7 Verifying maximum acceptance probability

We now describe and analyze a parallel algorithm, based on the multiplicative weights update method, to distinguish the two cases (2) above. The algorithm operates as described in Figure 2, and the super-operators Φ and Φ^* are as defined in Section 4.

7.1 Lemmas used in the analysis

We will need a few basic facts and three lemmas to analyze the algorithm described in Figure 2. We begin by noting two facts concerning matrix exponentials. First, the *Golden-Thompson Inequality* (see Section IX.3 of [Bha97]) states that, for any two Hermitian matrices X and Y of equal dimension, we have
$$\operatorname{Tr}\left(e^{X+Y}\right) \leq \operatorname{Tr}\left(e^X e^Y\right).$$
Second is the following inequality concerning the matrix exponential of positive semidefinite matrices.

Lemma 7.1. *Let P be an operator satisfying $0 \leq P \leq \mathbb{1}$. Then for every real number $\eta > 0$, it holds that*
$$\exp(-\eta P) \leq \mathbb{1} - \eta\exp(-\eta)P.$$

Proof. It is sufficient to prove the inequality for P replaced by a scalar $\lambda \in [0,1]$, for then the operator inequality follows by considering a spectral decomposition of P. If $\lambda = 0$ then the inequality is trivial, so assume $\lambda > 0$. By the Mean Value Theorem there exists a value $\lambda_0 \in (0,\lambda)$ such that

$$\frac{\exp(-\eta\lambda) - 1}{\lambda} = -\eta \exp(-\eta\lambda_0) \leq -\eta \exp(-\eta),$$

which yields the inequality. □

For the proofs of the remaining two lemmas, some additional notation is used. Suppose that $\mathcal{X} = \mathbb{C}^N$ and $\mathcal{Y} = \mathbb{C}^M$, and assume that the standard bases of these spaces are $\{|0\rangle, \ldots, |N-1\rangle\}$ and $\{|0\rangle, \ldots, |M-1\rangle\}$, respectively. Then we define a linear mapping

$$\mathrm{vec} : \mathrm{L}(\mathcal{X}, \mathcal{Y}) \to \mathcal{Y} \otimes \mathcal{X}$$

by taking $\mathrm{vec}(|i\rangle\langle j|) = |i\rangle |j\rangle$ for each choice of $i \in \{0, \ldots, M-1\}$ and $j \in \{0, \ldots, N-1\}$.

Lemma 7.2. *Let $\mathcal{X} = \mathbb{C}^N$ and $\mathcal{Y} = \mathbb{C}^M$ for positive integers N and M. Let $P_0, P_1 \in \mathrm{Pos}(\mathcal{X})$ and let $R_0 \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ satisfy $\mathrm{Tr}_{\mathcal{Y}}(R_0) = P_0$. Then there exists an operator $R_1 \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ such that $\mathrm{Tr}_{\mathcal{Y}}(R_1) = P_1$ and $\mathrm{F}(R_0, R_1) = \mathrm{F}(P_0, P_1)$.*

Proof. By the monotonicity of the fidelity function, it must hold that $\mathrm{F}(R_0, R_1) \leq \mathrm{F}(P_0, P_1)$ for every choice of R_1 satisfying $\mathrm{Tr}_{\mathcal{Y}}(R_1) = P_1$. It therefore suffices to show that equality can be achieved.

Choose a unitary operator $V \in \mathrm{U}(\mathcal{X})$ for which $\sqrt{P_0}\sqrt{P_1}V$ is positive semidefinite. For such a V it holds that $\mathrm{F}(P_0, P_1) = \mathrm{Tr}(\sqrt{P_0}\sqrt{P_1}V)$. Now let $\mathcal{W} = \mathbb{C}^{NM}$ and let $|u_0\rangle \in \mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{W}$ be a purification of R_0. Given that $|u_0\rangle$ also purifies P_0, it must take the form

$$|u_0\rangle = \mathrm{vec}\left(\sqrt{P_0}U^*\right)$$

for some choice of a linear isometry $U \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{W})$. Finally, let

$$R_1 = \mathrm{Tr}_{\mathcal{W}}\left(\mathrm{vec}\left(\sqrt{P_1}VU^*\right)\mathrm{vec}\left(\sqrt{P_1}VU^*\right)^*\right).$$

It holds that

$$\mathrm{F}(R_0, R_1) \geq \left|\left\langle \sqrt{P_0}U^*, \sqrt{P_1}VU^*\right\rangle\right| = \mathrm{Tr}\left(\sqrt{P_0}\sqrt{P_1}V\right) = \mathrm{F}(P_0, P_1)$$

as required. □

Remark 7.3. Lemma 7.2 is a fairly straightforward extension of Uhlmann's Theorem [Uhl76]. (See also pages 410–411 of [NC00].)

Lemma 7.4. *Let $R_0, R_1 \in \mathrm{Pos}(\mathcal{X})$ for $\mathcal{X} = \mathbb{C}^N$. Then*

$$\|R_0 - R_1\|_1 \leq \sqrt{2\,\mathrm{Tr}(R_0)^2 + 2\,\mathrm{Tr}(R_1)^2 - 4\,\mathrm{F}(R_0, R_1)^2}.$$

Proof. Choose $V \in \mathrm{U}(\mathcal{X})$ so that $\sqrt{R_0}\sqrt{R_1}V$ is positive semidefinite, and therefore $\mathrm{F}(R_0, R_1) = \mathrm{Tr}(\sqrt{R_0}\sqrt{R_1}V)$. We have that

$$\mathrm{vec}\left(\sqrt{R_0}\right) \in \mathcal{X} \otimes \mathcal{X} \qquad \text{and} \qquad \mathrm{vec}\left(\sqrt{R_1}V\right) \in \mathcal{X} \otimes \mathcal{X}$$

purify R_0 and R_1, respectively, so by the monotonicity of the trace norm it holds that

$$\|R_0 - R_1\|_1 \leq \left\| \operatorname{vec}\left(\sqrt{R_0}\right) \operatorname{vec}\left(\sqrt{R_0}\right)^* - \operatorname{vec}\left(\sqrt{R_1}V\right) \operatorname{vec}\left(\sqrt{R_1}V\right)^* \right\|_1.$$

Using the inequality

$$\|A\|_1 \leq \sqrt{\operatorname{rank}(A)}\, \|A\|_2,$$

which follows easily from the expressions of the trace and Frobenius norms in terms of singular values, along with the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\|R_0 - R_1\|_1 &\leq \sqrt{2} \left\| \operatorname{vec}\left(\sqrt{R_0}\right) \operatorname{vec}\left(\sqrt{R_0}\right)^* - \operatorname{vec}\left(\sqrt{R_1}V\right) \operatorname{vec}\left(\sqrt{R_1}V\right)^* \right\|_2 \\
&= \sqrt{2\operatorname{Tr}(R_0)^2 + 2\operatorname{Tr}(R_1)^2 - 4\operatorname{Tr}\left(\sqrt{R_0}\sqrt{R_1}V\right)^2} \\
&= \sqrt{2\operatorname{Tr}(R_0)^2 + 2\operatorname{Tr}(R_1)^2 - 4\operatorname{F}(R_0, R_1)^2}
\end{aligned}$$

as required. □

Remark 7.5. When $R_0 = \rho_0$ and $R_1 = \rho_1$ for density operators ρ_0 and ρ_1, we obtain the familiar inequality

$$\|\rho_0 - \rho_1\|_1 \leq 2\sqrt{1 - \operatorname{F}(\rho_0, \rho_1)^2},$$

or equivalently

$$\operatorname{F}(\rho_0, \rho_1) \leq \sqrt{1 - \frac{1}{4}\|\rho_0 - \rho_1\|_1^2},$$

which is one of the Fuchs-van de Graaf inequalities [FvdG99].

7.2 Analysis of the algorithm (ignoring precision)

Our algorithm cannot be implemented exactly using bounded-depth Boolean circuits: the spectral decompositions and matrix exponentials can only be approximated. However, for the sake of exposition, the issue of precision will be completely ignored in this section; meaning that we will imagine that all of the operations can be performed exactly. In the subsection that follows this one, the actual precision requirements of the algorithm are considered. As is shown there, it turns out that the algorithm is not particularly sensitive to errors, and in fact it is possible to perform all of the required computations in parallel with exponentially greater accuracy than would be required for the correctness of the algorithm.

Let us consider first the case that the algorithm accepts. Let $\rho = \rho_t$ for the iteration t of the loop in step 3 in which acceptance occurs. To prove that the algorithm has answered correctly, we will construct an operator $X \in \operatorname{Pos}(\mathcal{Y} \otimes \mathcal{X})$ such that $\Phi(X) \leq \mathbb{1}_{\mathcal{Y}}$ and $\operatorname{Tr}(X) \geq (1-\varepsilon)\gamma$; and therefore $\mu(Q) \geq (1-\varepsilon)\gamma$. By the conditions (2) on Q, this implies that $\mu(Q) \geq (1+4\varepsilon)\gamma$, and therefore $x \in A_{\text{yes}}$, as required.

The operator X is defined as follows. First, let $R_0 = Q^{-1/2}\rho Q^{-1/2}$, let $P_0 = \operatorname{Tr}_{\mathcal{Y}}(R_0) = \Phi(\rho)$, and let

$$P_1 = \frac{1}{\gamma}\sum_{j \in S} |x_j\rangle\langle x_j| + \sum_{j \notin S} \lambda_j |x_j\rangle\langle x_j|.$$

By Lemma 7.2 there must exist $R_1 \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$ such that $\mathrm{Tr}_{\mathcal{Y}}(R_1) = P_1$ and $\mathrm{F}(R_0, R_1) = \mathrm{F}(P_0, P_1)$. We then take

$$X = \gamma \sqrt{Q} R_1 \sqrt{Q}.$$

It holds that $X \geq 0$ and $\Phi(X) \leq \mathbb{1}_{\mathcal{X}}$. To establish a lower bound on $\mathrm{Tr}(X)$, we first note that

$$1 - \mathrm{Tr}\left(\sqrt{Q} R_1 \sqrt{Q}\right) = \langle Q, R_0 - R_1 \rangle \leq \|Q\| \, \|R_0 - R_1\|_1 .$$

By Lemma 7.4, we conclude that

$$\|R_0 - R_1\|_1 \leq \sqrt{2\,\mathrm{Tr}(P_0)^2 + 2\,\mathrm{Tr}(P_1)^2 - 4\,\mathrm{F}(P_0, P_1)^2}.$$

It holds that $\mathrm{Tr}(P_1) \leq \mathrm{Tr}(P_0)$, and we also have

$$\mathrm{F}(P_0, P_1) \geq \sum_{j \notin S} \lambda_j = \mathrm{Tr}(P_0) - s \geq \mathrm{Tr}(P_0) - \delta \|Q^{-1}\|.$$

Therefore, given that $\mathrm{Tr}(P_0) \leq \|Q^{-1}\|$, we have $\|R_0 - R_1\|_1 \leq \sqrt{8\delta}\|Q^{-1}\|$, and so

$$1 - \mathrm{Tr}\left(\sqrt{Q} R_1 \sqrt{Q}\right) \leq \sqrt{8\delta}\|Q^{-1}\| \, \|Q\| = \sqrt{8\delta}\kappa(Q) = \varepsilon.$$

It follows that

$$\mathrm{Tr}(X) = \gamma \,\mathrm{Tr}\left(\sqrt{Q} R_1 \sqrt{Q}\right) \geq (1 - \varepsilon)\gamma$$

as required.

Now let us consider the case that the algorithm rejects. Along similar lines to the previous case, we will construct an operator $Y \in \mathrm{Pos}(\mathcal{X})$ such that $\Phi^*(Y) \geq \mathbb{1}_{\mathcal{Y} \otimes \mathcal{X}}$ and $\mathrm{Tr}(Y) \leq (1 + \varepsilon)\gamma$. By the conditions (2) on Q this implies that $\mu(Q) \leq (1 - 4\varepsilon)\gamma$, and therefore $x \in A_{\mathrm{no}}$. In particular, we may take

$$Y = \frac{1 + \varepsilon}{T}(Y_0 + \cdots + Y_{T-1}).$$

Each operator Y_t satisfies

$$\mathrm{Tr}(Y_t) = \frac{|S|}{s} < \frac{1}{s} \sum_{j \in S} \gamma \lambda_j = \gamma,$$

and therefore $\mathrm{Tr}(Y) < (1 + \varepsilon)\gamma$. Each Y_t is also clearly positive semidefinite, so it remains to prove that $\Phi^*(Y) \geq \mathbb{1}_{\mathcal{Y} \otimes \mathcal{X}}$. To this end we will first establish two conditions on each operator Y_t. First, we have

$$\langle \rho_t, \Phi^*(Y_t) \rangle = \langle \Phi(\rho_t), Y_t \rangle = \frac{1}{s} \sum_{j \in S} \lambda_j = 1. \tag{5}$$

Second, given that $s > \delta \|Q^{-1}\|$ for the case at hand, we have

$$\|\Phi^*(Y_t)\| = \|Q^{-1/2}(\mathbb{1}_{\mathcal{Y}} \otimes Y_t)Q^{-1/2}\| \leq \|Q^{-1}\| \, \|Y_t\| = \frac{\|Q^{-1}\|}{s} < \frac{1}{\delta},$$

and therefore $\|\delta \Phi^*(Y_t)\| < 1$.

Now, for the sake of clarity, let us write $\eta = \varepsilon\gamma/2$. Note that, for $0 \leq t \leq T - 1$, it holds that

$$\begin{aligned}
\mathrm{Tr}(W_{t+1}) &= \mathrm{Tr}\left[\exp(-\eta\delta\Phi^*(Y_0 + \cdots + Y_t))\right] \\
&\leq \mathrm{Tr}\left[\exp(-\eta\delta\Phi^*(Y_0 + \cdots + Y_{t-1}))\exp(-\eta\delta\Phi^*(Y_t))\right] \\
&= \mathrm{Tr}\left[W_t \exp(-\eta\delta\Phi^*(Y_t))\right],
\end{aligned}$$

where we have used the Golden–Thompson Inequality. Given that $\|\delta\Phi^*(Y_t)\| \leq 1$ we have by Lemma 7.1 that

$$\exp(-\eta\delta\Phi^*(Y_t)) \leq \mathbb{1} - \eta\delta\exp(-\eta)\Phi^*(Y_t),$$

and therefore

$$\mathrm{Tr}(W_{t+1}) \leq \mathrm{Tr}(W_t)(1 - \eta\delta\exp(-\eta)\langle\rho_t, \Phi^*(Y_t)\rangle) \leq \mathrm{Tr}(W_t)\exp(-\eta\delta\exp(-\eta)).$$

(Here we have used the inequality $\exp(-\alpha) \geq 1 - \alpha$, which holds for all real numbers α, as well as the fact that $\mathrm{Tr}(AB) \leq \mathrm{Tr}(AC)$ whenever $A \geq 0$ and $B \leq C$). Repeating this argument, and substituting $\mathrm{Tr}(W_0) = NM$, we have

$$\mathrm{Tr}(W_T) \leq NM\exp(-\eta\delta T\exp(-\eta)).$$

On the other hand, it is clear that

$$\mathrm{Tr}(W_T) = \mathrm{Tr}\left[\exp(-\eta\delta\Phi^*(Y_0 + \cdots + Y_{T-1}))\right] \geq \exp(-\eta\delta\lambda_{NM}(\Phi^*(Y_0 + \cdots + Y_{T-1}))),$$

and therefore

$$\lambda_{NM}\left(\Phi^*\left(\frac{Y_0 + \cdots + Y_{T-1}}{T}\right)\right) \geq \exp(-\eta) - \frac{\ln(NM)}{\eta\delta T}.$$

Substituting the specified value of T, and using the fact that $\exp(-\eta) - \frac{\eta^2}{3} \geq 1 - \eta$ (which holds for any $\eta \in [0, 1]$), we have

$$\lambda_{NM}\left(\Phi^*\left(\frac{Y_0 + \cdots + Y_{T-1}}{T}\right)\right) \geq 1 - \eta = 1 - \frac{\varepsilon\gamma}{2}.$$

Therefore

$$\lambda_{NM}(\Phi^*(Y)) \geq (1 + \varepsilon)\left(1 - \frac{\varepsilon\gamma}{2}\right) \geq 1,$$

and so $\Phi^*(Y) \geq \mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}}$ as required.

We therefore have that the algorithm works correctly. It remains only to observe that it can be implemented in NC (meaning that it results in an NC($poly$) computation when composed with the first two steps of the simulation). Some of the details required to argue this can be found below; but at a high level one sees that each iteration of the loop in step 3 can be performed with high precision in NC, and the total number of iterations required is polynomial in n (and therefore polylogarithmic in the size of Q).

7.3 Precision requirements for the simulation

We now discuss the precision that is required for the simulation to yield a correct answer. It will turn out that the simulation is not particularly sensitive to errors, and one could in fact afford to take exponentially more precision than is required and still be within the class NC($poly$).

The first step of the simulation, in which an explicit description of $|\psi\rangle$ and Π is obtained, can be performed exactly in NC($poly$) as has already been observed. So, let us move on to the second step, in which $|\psi\rangle$ and Π are processed to obtain a well-conditioned interactive measurement operator Q. This step requires the approximation of one singular value decomposition (to approximate the Schmidt decomposition of $|\psi\rangle$), along with a few other operations that can be performed exactly or with high precision in NC.

For the moment let us denote by $\widetilde{Q}$ the actual operator that is computed by an NC implementation of this step, as opposed to the true operator Q that would be output by an idealized, exact

complex number algorithm. By computing the singular value decomposition to high precision, it is possible to take such an approximation so that

$$\left\| Q - \widetilde{Q} \right\| < 2^{-2^{poly(n)}},$$

where *poly* denotes any polynomial of our choice. It is not difficult to prove that the quantity $|\mu(Q) - \mu(\widetilde{Q})|$ is upper-bounded by N times $\|Q - \widetilde{Q}\|$, and therefore we may take $\widetilde{Q}$ so that

$$\left| \mu(Q) - \mu(\widetilde{Q}) \right| < 2^{-2^{poly(n)}},$$

again for *poly* denoting any polynomial of our choice. We do not need this much precision: we only need a $1/poly(n)$ separation between the values of $\mu(\widetilde{Q})$ for the cases $x \in A_{\text{yes}}$ and $x \in A_{\text{no}}$, which requires that $\|Q - \widetilde{Q}\|$ is exponentially (rather than double-exponentially) small in n. So, to be concrete, we may decide to take sufficient precision so that

$$\left\| Q - \widetilde{Q} \right\| < \varepsilon\gamma/N,$$

and therefore

$$\left| \mu(Q) - \mu(\widetilde{Q}) \right| < \varepsilon\gamma.$$

Thus,

$$\begin{aligned} x \in A_{\text{yes}} \quad &\Rightarrow \quad \mu(\widetilde{Q}) \geq (1 + 3\varepsilon)\gamma \\ x \in A_{\text{no}} \quad &\Rightarrow \quad \mu(\widetilde{Q}) \leq (1 - 3\varepsilon)\gamma. \end{aligned} \tag{6}$$

Hereafter we will return to writing Q rather than $\widetilde{Q}$, with the understanding that Q now represents an approximation that is stored by our algorithm. In addition, we will assume that $\sqrt{Q}$, and therefore $Q^{-1/2}$ as well, has Gaussian rational entries and is known precisely. This assumption is easily met by replacing Q with the square of a high precision approximation to $\sqrt{Q}$. The point of this assumption is that we avoid having to consider an additional error term every time $\sqrt{Q}$ or $Q^{-1/2}$ is involved in any computation. (There is no reason beyond simplifying the analysis to make this assumption.)

Now suppose that the algorithm described above is performed with limited precision. Consider first the case that the algorithm accepts, and let $\rho = \rho_t$ denote the density operator that is stored by the algorithm on the iteration t in which acceptance occurs. Note that it is not necessary to view that ρ is an approximation of something else: the simple fact that ρ causes acceptance will allow us to conclude that $x \in A_{\text{yes}}$ in a similar way to the error-free analysis.

Specifically, we will construct an operator $X \in \text{Pos}(\mathcal{Y} \otimes \mathcal{X})$ such that $\Phi(X) \leq \mathbb{1}_{\mathcal{Y}}$ and $\text{Tr}(X) \geq (1 - 2\varepsilon)\gamma$. As before we will let $R_0 = Q^{-1/2}\rho Q^{-1/2}$ and $P_0 = \text{Tr}_{\mathcal{Y}}(R_0)$. This time, we must consider that the spectral decomposition is approximate. Let us write

$$\widetilde{P}_0 = \sum_{j=1}^{N} \lambda_j |x_j\rangle\langle x_j|$$

to denote the approximate value of the spectral decomposition, so that $\|P_0 - \widetilde{P}_0\|$ represents the error in this approximation, and let us assume that sufficient accuracy is taken so that

$$\left\| P_0 - \widetilde{P}_0 \right\| < \frac{\delta}{4N \left\| Q^{-1} \right\|}. \tag{7}$$

As above, this is significantly less accuracy than is available—for we could take

$$\left\| P_0 - \widetilde{P}_0 \right\| < 2^{-2^{poly(n)}}$$

if it were advantageous. Continuing on as before, let

$$P_1 = \frac{1}{\gamma} \sum_{j \in S} |x_j\rangle\langle x_j| + \sum_{j \notin S} \lambda_j |x_j\rangle\langle x_j|,$$

let $R_1 \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$ to be an extension of P_1 for which $\mathrm{F}(R_0, R_1) = \mathrm{F}(P_0, P_1)$, and let

$$X = \gamma \sqrt{Q} R_1 \sqrt{Q}.$$

We have that $X \geq 0$ and $\Phi(X) \leq \mathbb{1}_{\mathcal{X}}$ as before, and to establish a lower bound on $\mathrm{Tr}(X)$ we again use the fact that

$$1 - \mathrm{Tr}\left(\sqrt{Q} R_1 \sqrt{Q}\right) \leq \|Q\| \, \|R_0 - R_1\|_1,$$

as well as the bound

$$\|R_0 - R_1\|_1 \leq \sqrt{2\,\mathrm{Tr}(P_0)^2 + 2\,\mathrm{Tr}(P_1)^2 - 4\,\mathrm{F}(P_0, P_1)^2}.$$

Based on the bound (7), it follows that

$$\mathrm{Tr}(P_1)^2 \leq \mathrm{Tr}(P_0)^2 + \delta$$

and

$$\mathrm{F}(P_0, P_1)^2 \geq \left(\mathrm{Tr}(P_0) - \delta \left\| Q^{-1} \right\|\right)^2 + \delta,$$

and therefore

$$\|R_0 - R_1\|_1 \leq \sqrt{14\delta} \|Q^{-1}\|.$$

It follows that

$$1 - \mathrm{Tr}\left(\sqrt{Q} R_1 \sqrt{Q}\right) < 2\varepsilon,$$

and therefore

$$\mathrm{Tr}(X) = \gamma \,\mathrm{Tr}\left(\sqrt{Q} R_1 \sqrt{Q}\right) \geq (1 - 2\varepsilon)\gamma$$

as required.

Now let us consider the case that the algorithm rejects, and in particular let us focus on the operators $Y_0, \ldots, Y_{T-1}$ that are computed over the course of the algorithm. As for the case of acceptance, these operators are not viewed as approximations to anything: the fact that these operators exist and cause rejection in the algorithm is enough to conclude that $x \in A_{\mathrm{no}}$. Let us continue to write

$$W_{t+1} = \exp\left(-\frac{\varepsilon\gamma\delta}{2} \Phi^*(Y_0 + \cdots + Y_t)\right)$$

and $\rho_t = W_t / \mathrm{Tr}(W_t)$ for each $t = 0, \ldots, T-1$; but we must keep in mind that the algorithm only computes approximations of these operators. The algorithm must also approximate the spectral decomposition of each $\Phi(\rho_t)$, where the source of errors in this case comes from both from the spectral decomposition computation and the fact that ρ_t is approximated.

Now, in the error free analysis, the conditions

$$\mathrm{Tr}(Y_t) \leq \gamma, \qquad \langle \rho_t, \Phi^*(Y_t) \rangle = 1, \qquad \text{and} \qquad \| \delta\Phi^*(Y_t) \| \leq 1$$

were proved, and these conditions allowed us to conclude that Y satisfies

$$\Phi^*(Y) \geq \mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}} \qquad \text{and} \qquad \mathrm{Tr}(Y) \leq (1+\varepsilon)\gamma.$$

If we follow precisely the same proof, but with the condition $\langle \rho_t, \Phi^*(Y_t) \rangle = 1$ replaced by

$$\langle \rho_t, \Phi^*(Y_t) \rangle \geq 1 - \alpha$$

for some choice of $\alpha > 0$, we once again find immediately that $\mathrm{Tr}(Y) \leq (1+\varepsilon)\gamma$. This time we have

$$\lambda_{NM}\left(\Phi^*\left(\frac{Y_0 + \cdots + Y_{T-1}}{T}\right)\right) \geq (1-\alpha)\exp(-\eta) - \frac{\ln(NM)}{\eta\delta T},$$

but under the assumption $\alpha < \eta^2/12$, say, it follows again that $\Phi^*(Y) \geq \mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}}$.

Thus, given that the conditions $\mathrm{Tr}(Y_t) \leq \gamma$ and $\| \delta\Phi^*(Y_t) \| \leq 1$ follow from an inspection of the algorithm as before, it suffices to compute the matrix exponentials and spectral decompositions with sufficient accuracy that $\langle \rho_t, \Phi^*(Y_t) \rangle > 1 - \eta^2/12$. This is easily done: as the argument of the matrix exponentials have norm bounded by T, one is able to compute both the matrix exponentials and the spectral decompositions with exponentially greater accuracy in NC than is required.

8 Conclusion

We have proved that QIP(2) $\subseteq$ PSPACE using a semidefinite programming formulation of the maximum acceptance probability of two-message quantum interactive proof systems, along with the multiplicative weights update method for verifying these values. An obvious question remains: can this method be extended, or some other method devised, to prove QIP = PSPACE?

Acknowledgments

The research presented in this paper was supported by Canada's NSERC, MITACS, the Canadian Institute for Advanced Research, and the US Army Research Office and National Security Agency; and was partially done while Rahul Jain was a member of the Institute for Quantum Computing and the School of Computer Science at the University of Waterloo.

References

[AHK05] S. Arora, E. Hazan, and S. Kale. Fast algorithms for approximate semidefinite programming using the multiplicative weights update method. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 339–348, 2005.

[AK07] S. Arora and S. Kale. A combinatorial, primal-dual approach to semidefinite programs. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, pages 227–236, 2007.

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

[BATS09] A. Ben-Aroya and A. Ta-Shma. On the complexity of approximating the diamond norm. Available as arXiv.org e-Print 0902.3397, 2009.

[BCP83] 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.

[BGH82] A. Borodin, J. von zur Gathen, and J. Hopcroft. Fast parallel matrix and GCD computations. In *Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science*, pages 65–71, 1982.

[Bha97] R. Bhatia. *Matrix Analysis*. Springer, 1997.

[BM88] L. Babai and S. 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.

[BOFKT86] M. Ben-Or, E. Feig, D. Kozen, and P. Tiwari. A fast parallel algorithm for determining all roots of a polynomial with real roots. In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing*, pages 340–349, 1986.

[BOGKW88] M. Ben-Or, S. Goldwasser, J. Kilian, and A. Wigderson. Multi-prover interactive proofs: how to remove intractability assumptions. In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, pages 113–131, 1988.

[Bor77] A. Borodin. On relating time and space to size and depth. *SIAM Journal on Computing*, 6:733–744, 1977.

[Cho75] M.-D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra and Its Applications*, 10(3):285–290, 1975.

[Csa76] L. Csanky. Fast parallel matrix inversion algorithms. *SIAM Journal on Computing*, 5(4):618–623, 1976.

[FvdG99] C. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999.

[Gat93] J. von zur Gathen. Parallel linear algebra. In J. Reif, editor, *Synthesis of Parallel Algorithms*, chapter 13. Morgan Kaufmann Publishers, Inc., 1993.

[GLS93] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer–Verlag, second corrected edition, 1993.

[GMR85] S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proof systems. In *Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, pages 291–304, 1985.

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

[GMW91] 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.

[GS89] S. Goldwasser and M. Sipser. Private coins versus public coins in interactive proof systems. In S. Micali, editor, *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 73–90. JAI Press, 1989.

[GW07] G. Gutoski and J. Watrous. Toward a general theory of quantum games. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 565–574, 2007.

[HKSZ08] S. Hallgren, A. Kolla, P. Sen, and S. Zhang. Making classical honest verifier zero knowledge protocols secure against quantum attacks. In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming*, volume 5126 of *Lecture Notes in Computer Science*, pages 592–603. Springer, 2008.

[Jam72] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[JW09] R. Jain and J. Watrous. Parallel approximation of non-interactive zero-sum quantum games. In *Proceedings of the 24th IEEE Conference on Computational Complexity*, 2009. To appear.

[Kal07] S. Kale. *Efficient algorithms using the multiplicative weights update method*. PhD thesis, Princeton University, 2007.

[KKMV08] J. Kempe, H. Kobayashi, K. Matsumoto, and T. Vidick. Using entanglement in quantum multi-prover interactive proofs. In *Proceedings of the 23rd Annual Conference on Computational Complexity*, 2008.

[KM03] H. Kobayashi and K. Matsumoto. Quantum multi-prover interactive proof systems with limited prior entanglement. *Journal of Computer and System Sciences*, 66(3), 2003.

[Kob08] H. Kobayashi. General properties of quantum zero-knowledge proofs. In *Proceedings of the Fifth IACR Theory of Cryptography Conference*, volume 4948 of *Lecture Notes in Computer Science*, pages 107–124. Springer, 2008.

[KW00] A. Kitaev and J. Watrous. Parallelization, amplification, and exponential time simulation of quantum interactive proof system. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pages 608–617, 2000.

[LFKN92] C. Lund, L. Fortnow, H. Karloff, and N. Nisan. Algebraic methods for interactive proof systems. *Journal of the ACM*, 39(4):859–868, 1992.

[MW05] C. Marriott and J. Watrous. Quantum Arthur-Merlin games. *Computational Complexity*, 14(2):122–152, 2005.

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Nef94] C. A. Neff. Specified precision polynomial root isolation is in NC. *Journal of Computer and System Sciences*, 48(3):429–463, 1994.

[Sha92] A. Shamir. IP = PSPACE. *Journal of the ACM*, 39(4):869–877, 1992.

[Uhl76] A. Uhlmann. The "transition probability" in the state space of a $*$-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[Wat02] J. 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, 2002.

[Wat03] J. Watrous. PSPACE has constant-round quantum interactive proof systems. *Theoretical Computer Science*, 292(3):575–588, 2003.

[Wat06] J. Watrous. Zero-knowledge against quantum attacks. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 296–305, 2006.

[Wat09a] J. Watrous. Quantum computational complexity. In *Encyclopedia of Complexity and System Science*. Springer, 2009. Available as arXiv.org e-Print 0804.3401.

[Wat09b] J. Watrous. Semidefinite programs for completely bounded norms. Available as arXiv.org e-Print 0901.4709, 2009.

[Weh06] S. Wehner. Entanglement in interactive proof systems with binary answers. In *Proceedings of the 23rd Annual Symposium on Theoretical Aspects of Computer Science*, volume 3884 of *Lecture Notes in Computer Science*, pages 162–171. Springer, 2006.

[WK06] M. Warmuth and D. Kuzmin. Online variance minimization. In *Proceedings of the 19th Annual Conference on Learning Theory*, volume 4005 of *Lecture Notes in Computer Science*, pages 514–528. Springer, 2006.