2012 IEEE 27th Conference on Computational Complexity

Parallel approximation of min-max problems with applications to classical and quantum zero-sum games

Gus Gutoski* and Xiaodi Wu†
*Institute for Quantum Computing and School of Computer Science
University of Waterloo, Waterloo, Ontario, Canada
†Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, Michigan, USA

***Abstract*—This paper presents an efficient parallel algorithm for a new class of min-max problems based on the matrix multiplicative weights update method. Our algorithm can be used to find near-optimal strategies for competitive two-player classical or quantum games in which a referee exchanges any number of messages with one player followed by any number of additional messages with the other. This algorithm considerably extends the class of games which admit parallel solutions, demonstrating for the first time the existence of a parallel algorithm for a game in which one player reacts adaptively to the other.**

As a consequence, we prove that several competing-provers complexity classes collapse to PSPACE such as QRG(2), SQG and two new classes called DIP and DQIP. A special case of our result is a parallel approximation scheme for a new class of semidefinite programs whose feasible region consists of lists of semidefinite matrices that satisfy a "transcript-like" consistency condition. Applied to this special case, our algorithm yields a direct polynomial-space simulation of multi-message quantum interactive proofs resulting in a first-principles proof of QIP=PSPACE.

I. Introduction

A. Results

Parallel approximation of semidefinite programs and min-max problems: This paper presents an efficient parallel algorithm for a new class of min-max problems with applications to classical and quantum zero-sum games and interactive proofs. A special case of our result is a parallel approximation scheme for semidefinite programs (SDPs) of the form

$$\begin{aligned} \text{minimize} \quad & \operatorname{Tr}(X_k P) && (1)\\ \text{subject to} \quad & \operatorname{Tr}_{n_i}(X_i) = \Phi_{i-1}(X_{i-1}) \text{ for } i = 2, \dots, k \\ & \operatorname{Tr}_{n_1}(X_1) = Q \\ & 0 \preceq X_i \in \mathbb{M}_{m_i n_i} \text{ for } i = 1, \dots, k \end{aligned}$$

Here $\mathbb{M}_d$ denotes the space of all $d \times d$ complex matrices and Tr_n is the *partial trace*—the unique linear map from matrices to matrices satisfying

$$\operatorname{Tr}_n : \mathbb{M}_{mn} \to \mathbb{M}_m : A \otimes B \mapsto \operatorname{Tr}(B) A$$

for every choice of $A \in \mathbb{M}_m$ and $B \in \mathbb{M}_n$. An SDP (1) is specified by arbitrary choices of positive semidefinite matrices $P \in \mathbb{M}_{m_k n_k}$ and $Q \in \mathbb{M}_{m_1}$ and completely positive and trace-preserving linear maps $\Phi_i : \mathbb{M}_{m_i n_i} \to \mathbb{M}_{m_{i+1}}$. The run time of our parallel algorithm scales linearly with the largest eigenvalue of P and with the trace of Q, so it is only efficient when these quantities are bounded by a fixed polynomial in the logarithm of the bit length of the input $Q, R, \Phi_1, \dots, \Phi_{k-1}$. (In keeping with convention, one can think of these quantities as the *width* of the SDPs we consider. Our algorithm is efficient only for *width-bounded* SDPs.)

It has long since been known that the problem of approximating the optimal value of an arbitrary SDP is logspace-hard for P [1], [2], so there cannot be a parallel approximation scheme for *all* SDPs unless $\mathrm{NC} = \mathrm{P}$. However, the precise extent to which SDPs admit parallel solutions is not known. This special case of our result adds considerably to the set of such SDPs, subsuming all prior work in the area at the time it was made public. (Since that time parallel approximation schemes have been found for some SDPs of unbounded width that are not covered by our scheme [3].) Our result is stated in full generality as follows.

Theorem 1. *Let* $\mathbf{A}$ *denote the feasible region of the width-bounded SDP* (1) *and let* $\mathbf{P}$ *be a compact convex set of positive semidefinite operators. There is an efficient parallel oracle-algorithm for finding approximate solutions to the min-max problem*

$$\min_{(X_1,\dots,X_k)\in\mathbf{A}} \ \max_{P\in\mathbf{P}} \operatorname{Tr}(X_k P) \qquad (2)$$

with an oracle for optimization over the set $\mathbf{P}$. *The SDP* (1) *is recovered from the above min-max problem* (2) *in the special case where* $\mathbf{P} = \{P\}$ *is a singleton set.*

We also describe parallel implementations of this oracle for certain sets $\mathbf{P}$, yielding an unconditionally efficient parallel approximation algorithm for the min-max problem (2) for those choices of $\mathbf{P}$.

Applications to zero-sum games: This algorithm can be used to find near-optimal strategies for a new class of competitive two-player games that are moderated by a referee and obey the following protocol:

1) The referee exchanges several messages only with Alice.

1093-0159/12 $26.00 © 2012 IEEE
DOI 10.1109/CCC.2012.12

2) After processing this interaction with Alice, the referee exchanges several additional messages only with Bob.
3) After further processing, the referee declares a winner.

Indeed, our algorithm applies even to *quantum* games, in which the referee and players are free to exchange and process quantum information. Due to the similarity with the oft-studied interactive proof model of computation, games of this form shall be called *double interactive proofs*: the referee in such a game executes a standard interactive proof with Alice followed by a second interactive proof with Bob. This protocol is depicted in Figure 1.

If the referee is specified succinctly by circuits rather than in explicit matrix form then our parallel algorithm can be used to find near-optimal strategies in polynomial space.[1] This algorithm is optimal in that it is PSPACE-hard even to distinguish games that Alice can win with near certainty from games that Bob can win with near certainty. This strong form of PSPACE-hardness holds even in the special case of *two-turn games*, in which the referee exchanges only two messages *synchronously* with each player [5].

Ordinary interactive proofs could also be cast as a special type of game in which the referee completely ignores Bob. Taking this view, the celebrated proof of IP $=$ PSPACE [6], [7] implies a similar hardness result: it is PSPACE-hard to distinguish interactive proofs that Alice can win with certainty from those which she can win with only exponentially small probability.

Prior to the present work polynomial-space algorithms were known only for two-turn classical games and for quantum interactive proofs. The algorithm for two-turn games is due to Feige and Kilian [5]. Algorithms for quantum interactive proofs are presented in proofs of QIP $=$ PSPACE [8], [9].

Our result unifies and subsumes both of these algorithms. It also demonstrates for the first time the existence of a parallel algorithm for a game in which one player reacts adaptively to the other, including as a special case the (non-adaptive) class of two-turn quantum games.

Applications to complexity theory: In complexity theory, our result implies the collapse to PSPACE of several classical and quantum interactive proof classes. Letting DIP and DQIP denote the competing-provers complexity classes associated with classical and quantum double interactive proofs, respectively, we have

$$\mathrm{DQIP} = \mathrm{DIP} = \mathrm{PSPACE}.$$

In contrast to the classical case, the competing-provers complexity class $\mathrm{QRG}(2)$ associated with two-turn quantum games was not known to be a subset of PSPACE prior to the present work. A special case of our result yields the equality

$$\mathrm{QRG}(2) = \mathrm{PSPACE}$$

thus solving an open problem of Ref. [8]. Of course, every other complexity class whose protocol can be cast as a double interactive proof also collapses to PSPACE, such as SQG [10].

Our results also illustrate a difference in the effect of public randomness between *single*-prover interactive proofs and *competing*-prover interactive proofs. Any classical interactive proof with single prover can be simulated by another *public coin* interactive proof where the verifier's messages to the prover consist entirely of uniformly random bits and the verifier uses no other randomness [11]. Extending the notion of public coin interaction to competitive games, it is easy to see that any game with a public-coin referee can be simulated by a double interactive proof.[2] Letting RG denote the complexity class of decision problems that admit competing-prover interactive proofs, we therefore have that the public-coin version of RG is a subset of DIP, which we now know is equal to PSPACE. Thus, by contrast to the single-prover case where we have $\text{public-coin-IP} = \mathrm{IP}$, in the competing-prover case we have $\text{public-coin-RG} \neq \mathrm{RG}$ unless PSPACE $=$ EXP.

In the special case of the SDP (1) our algorithm yields a direct polynomial-space simulation of multi-message quantum interactive proofs, resulting in a first-principles proof of QIP $=$ PSPACE. By contrast, all other known proofs [8], [9] rely upon the fact that the verifier can be assumed to exchange only three messages with the prover [12]. The original proof of Jain *et al.* also relies on the additional fact that verifier's only message to the prover can be just a single classical coin flip [13].

Due to space constraints a proper review of relevant literature and some technical detail must be deferred to the full version of this paper.

B. Techniques

Our algorithm is an example of the *matrix multiplicative weights update method (MMW)* as presented in Refs. [14], [15], [16]. We also draw upon the valuable experience of recent applications of this method to parallel algorithms for quantum complexity classes [17], [18], [8], [9].

One of the novelties of our algorithm is that we apply the MMW method *twice* in a two-level recursive fashion. At the top level the MMW is used to iteratively converge toward

[1] There is a standard method by which parallel algorithms on succinct inputs are simulated in polynomial-space: first, inflate the succinct specification of the referee into exponential-size matrix form. Then run the parallel algorithm on this inflated input, so that the algorithm may be viewed as a polynomial-depth circuit with exponential-size input (the inflated referee) and exponential-size output (optimal strategies for the players). Such a circuit can be simulated in the usual way by a Turing machine with a polynomial-size read/write work tape and an unbounded write-only output tape [4].

[2] *Proof sketch:* As the referee's questions to a player are uniformly random, they cannot depend on prior responses from the other player and can therefore be reordered so that all messages with one player are exchanged before any messages are exchanged with the other.

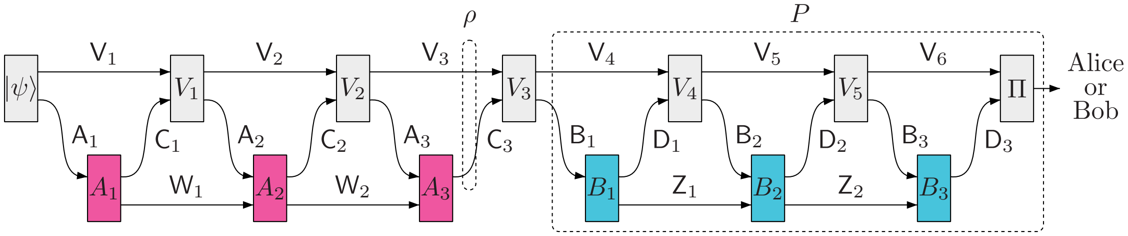

Figure 1. An illustration of a double quantum interactive proof in which the referee $R = (|\psi\rangle, V_1, \ldots, V_5, \Pi)$ exchanges $a = 3$ rounds of messages with Alice followed by $b = 3$ rounds of messages with Bob before performing the measurement $\{\Pi, I - \Pi\}$ and announcing a winner. Any choice of (A_1, A_2, A_3) and (B_1, B_2, B_3) induces a state ρ and a measurement operator P as indicated. Bob's winning probability is given by $\langle \rho, P\rangle = \mathrm{Tr}(\rho P)$.

an optimal strategy for Alice; at the bottom level the MMW is used again to solve an SDP for "best responses" for Bob to a given strategy for Alice.

Double quantum interactive proofs admit a representation for strategies that is amenable to the MMW. In Kitaev's *transcript* representation [19] the actions of a player are represented by a list $\rho_1, \ldots, \rho_n$ of density matrices that satisfy a special consistency condition. Intuitively, these density matrices correspond to "snapshots" of the state of the referee's qubits at various times during the interaction. (See Figure 2.)

Whereas the MMW in its unaltered form can be used to solve min-max problems over the domain of density operators, we introduce a new extension to this method for min-max problems over the domain of transcripts—a domain consisting of lists of *multiple* operators, each drawn from a *strict subset* of the density operators. The high-level approach of our method is as follows:

1) **Extend the domain from a single density matrix to a list of n density matrices.** This step is straightforward: the MMW can be applied without complication to all n density matrices at the same time. (Equivalently, n density matrices may be viewed as a single, larger, block-diagonal density matrix.)
2) **Restrict the domain to a strict subset of density matrices.** This step is more difficult. It is accomplished by relaxing the game so as to allow *all* density matrices, with an additional *penalty term* to remove incentive for the players to use inconsistent transcripts.
3) **Round strategies in the relaxed game to strategies in the original game.** For this step one must prove a "rounding" theorem (Theorem 5), which establishes that near-optimal, fully admissible strategies can be obtained from near-optimal strategies in the unrestricted domain with penalty term.

Comparison of methods for semidefinite programming: In their proof of QIP = PSPACE, Jain *et al.* employ the MMW to solve a special SDP for quantum interactive proofs by making direct use of the *primal-dual* approach described in Kale's thesis [16]. By contrast, we do not use this primal-dual approach for solving SDPs. Instead we use the MMW to solve a min-max problem as suggested by the algorithmic proof (also presented in Kale's thesis) of a min-max theorem for a simple class of zero-sum quantum games. By introducing a penalty term for inadmissible strategies we are able to extend this algorithm to a much richer class of games beyond the one-turn games considered by Kale. We wish to stress that our parallel algorithm for SDPs arises as a *special case* of a more general min-max algorithm, whereas the primal-dual approach for SDPs does not generalize to min-max problems in any obvious way.

The Bures metric: Finally, it is noteworthy that the proof of our rounding theorem (Theorem 5) contains an interesting and nontrivial application of the Bures metric, which is a distance measure for quantum states that is defined in terms of the more familiar fidelity function.

Properties of the trace norm, which captures the physical distinguishability of quantum states, are sufficient for most needs in quantum information. When some property of the fidelity is also required one uses the Fuchs-van de Graaf inequalities to convert between the trace norm and fidelity [20].

However, every such conversion incurs a quadratic slackening of relevant accuracy parameters. Our study calls for repeated conversions, which would incur an unacceptable exponential slackening if done naively via Fuchs-van de Graaf. Instead, we make only a *single* conversion between the trace norm and the Bures metric and then repeatedly exploit the simultaneous properties of (i) the triangle inequality, (ii) contractivity under quantum channels, and (iii) preservation of subsystem fidelity.

Although conversion inequalities between the trace norm and Bures metric are implied by Fuchs-van de Graaf, to our knowledge explicit conversion inequalities have not yet appeared in published literature. The required inequalities are derived in the present paper (Proposition 3).

II. Preliminaries

Hereafter we must assume familiarity with standard concepts from quantum information [21], [22] and parallel algorithms for various matrix manipulation tasks [23], [8]. See Table I for a summary of our notation and terminology.

Table I
NOTATION AND TERMINOLOGY

$\mathcal{X}, \mathcal{Y}, \mathcal{XY}$	Finite-dimensional complex vector spaces. $\mathcal{XY}$ is shorthand for $\mathcal{X} \otimes \mathcal{Y}$.
$\mathbf{L}(\mathcal{X})$	The (complex) space of all linear operators $A : \mathcal{X} \to \mathcal{X}$.
$I_{\mathcal{X}}$	The identity operator acting on $\mathcal{X}$.
$\mathbf{Pos}(\mathcal{X})$	The convex cone of positive semidefinite operators within $\mathbf{L}(\mathcal{X})$.
$\mathbf{Dens}(\mathcal{X})$	The compact convex set of all density operators within $\mathbf{Pos}(\mathcal{X})$.
$\mathbf{Meas}(\mathcal{X})$	The compact convex set of all measurement operators within $\mathbf{Pos}(\mathcal{X})$. A *measurement operator* is a positive semidefinite operator M with $M \preceq I_{\mathcal{X}}$.
A^*	The adjoint of an operator $A : \mathcal{X} \to \mathcal{Y}$, which has the form $A^* : \mathcal{Y} \to \mathcal{X}$.
$\langle A, B \rangle$	The standard inner product between $A, B : \mathcal{X} \to \mathcal{Y}$. Defined by $\langle A, B \rangle = \mathrm{Tr}(A^* B)$.

A. Preservation of subsystem fidelity

We employ the following property of the fidelity function.

Proposition 2 (Preservation of subsystem fidelity—see Ref. [18, Lemma 7.2])**.** *Let $\sigma, \sigma' \in \mathbf{Dens}(\mathcal{V})$ and $\rho \in \mathbf{Dens}(\mathcal{AV})$ be density operators with $\mathrm{Tr}_{\mathcal{A}}(\rho) = \sigma$. There exists a density operator $\rho' \in \mathbf{Dens}(\mathcal{AV})$ with $\mathrm{Tr}_{\mathcal{A}}(\rho') = \sigma'$ and $F(\rho, \rho') = F(\sigma, \sigma')$. Moreover ρ' can be computed efficiently in parallel given σ, σ', ρ.*

A formal construction of such a ρ' appears in Ref. [18]. Since their construction consists entirely of elementary matrix operations, it admits an efficient parallel implementation.

B. The Bures angle

The *Bures angle* or simply the *angle* $A(\rho, \xi)$ between quantum states ρ, ξ is defined by $A(\rho, \xi) \stackrel{\text{def}}{=} \arccos F(\rho, \xi)$. The angle is a metric on quantum states, meaning that it is nonnegative, equals zero only when $\rho = \xi$, and obeys the triangle inequality [21]. Moreover, the angle is *contractive*, so that $A(\Phi(\rho), \Phi(\xi)) \leq A(\rho, \xi)$ for any quantum channel Φ. The Fuchs-van de Graaf Inequalities establish a relationship between the fidelity and trace norm [20]. The inequalities are $1 - F(\rho, \xi) \leq \frac{1}{2} \| \rho - \xi \|_{\mathrm{Tr}} \leq \sqrt{1 - F(\rho, \xi)^2}$. These inequalities can be used to derive a relationship between $A(\rho, \xi)$ and $\| \rho - \xi \|_{\mathrm{Tr}}$. For example,

Proposition 3 (Relationship between trace norm and Bures angle)**.** *For all density matrices ρ, ξ it holds that*

$$\frac{1}{2} \| \rho - \xi \|_{\mathrm{Tr}} \leq A(\rho, \xi) \leq \sqrt{\frac{\pi}{2} \| \rho - \xi \|_{\mathrm{Tr}}}.$$

Proof: The upper bound follows immediately from Fuchs-van de Graaf:

$$\frac{1}{2} \| \rho - \xi \|_{\mathrm{Tr}} \leq \sqrt{1 - \cos A(\rho, \xi)^2} = \sin A(\rho, \xi) \leq A(\rho, \xi)$$

where we used the identity $\sin x \leq x$ for all $x \geq 0$.

To obtain the lower bound we employ the identity $\cos x \leq 1 - x^2/\pi$ for $x \in [0, \pi/2]$, which can be verified using basic calculus. Then we have

$$\frac{1}{2} \| \rho - \xi \|_{\mathrm{Tr}} \geq 1 - \cos A(\rho, \xi) \geq \frac{A(\rho, \xi)^2}{\pi}$$

from which the proposition follows. ■

III. DOUBLE QUANTUM INTERACTIVE PROOFS

A *double quantum interactive proof* is completely specified by a *referee*, which consists of a tuple $R = (|\psi\rangle, V_1, \ldots, V_{a+b-1}, \Pi)$ of objects where

1) $|\psi\rangle \in \mathcal{A}_1 \mathcal{V}_1$ is a pure state.
2) $V_1, \ldots, V_{a+b-1}$ are unitary operators

$$\begin{aligned} V_i &: \mathcal{C}_i \mathcal{V}_i \to \mathcal{A}_{i+1} \mathcal{V}_{i+1} && (1 \leq i \leq a-1) \\ V_a &: \mathcal{C}_a \mathcal{V}_a \to \mathcal{B}_1 \mathcal{V}_{a+1} && \\ V_{a+i} &: \mathcal{D}_i \mathcal{V}_{a+i} \to \mathcal{B}_{i+1} \mathcal{V}_{a+i+1} && (1 \leq i \leq b-1) \end{aligned}$$

3) $\Pi \in \mathbf{Meas}(\mathcal{D}_b \mathcal{V}_{a+b})$ is a projection.

Each of the spaces $\mathcal{A}_i, \mathcal{B}_i, \mathcal{C}_i, \mathcal{D}_i, \mathcal{V}_i$ corresponds to a register $\mathsf{A}_i, \mathsf{B}_i, \mathsf{C}_i, \mathsf{D}_i, \mathsf{V}_i$. The registers A_i are question registers from the referee to Alice and the registers C_i are answer registers from Alice to the referee. Similarly, the registers B_i are question registers from the referee to Bob and the registers D_i are answer registers from Bob to the referee. The registers V_i are private memory registers for the referee.

The actions of the players during each round of interaction are specified by unitary operators acting upon the question register and a private memory register for that player. In particular, Alice's actions in round $i = 1, \ldots, a$ are specified by unitary operators $A_i : \mathcal{A}_i \mathcal{W}_{i-1} \to \mathcal{C}_i \mathcal{W}_i$ where the spaces $\mathcal{W}_0, \ldots, \mathcal{W}_a$ correspond to the private memory registers $\mathsf{W}_0, \ldots, \mathsf{W}_a$ for Alice. Similarly, Bob's actions in round $a + i$ for $i = 1, \ldots, b$ are specified by unitary operators $B_i : \mathcal{B}_i \mathcal{Z}_{i-1} \to \mathcal{D}_i \mathcal{Z}_i$ where the spaces $\mathcal{Z}_0, \ldots, \mathcal{Z}_b$ correspond to the private memory registers $\mathsf{Z}_0, \ldots, \mathsf{Z}_b$ for Bob.

The game proceeds as suggested by Figure 1 and is described as follows:

1) The referee prepares the registers $(\mathsf{A}_1, \mathsf{V}_1)$ in the pure state $|\psi\rangle$. The players' private registers $\mathsf{W}_0, \mathsf{Z}_0$ are both initialized to the pure state $|0\rangle$.
2) For $i = 1, \ldots, a$:
    a) The register A_i is sent to Alice, who applies A_i to the registers $(\mathsf{A}_i, \mathsf{W}_{i-1})$ to obtain registers $(\mathsf{C}_i, \mathsf{W}_i)$.
    b) The register C_i is then returned to the referee, who applies V_i to the registers $(\mathsf{C}_i, \mathsf{V}_i)$ to obtain registers $(\mathsf{A}_{i+1}, \mathsf{V}_{i+1})$. (In the case $i = a$ this final pair of registers is instead labelled $(\mathsf{B}_1, \mathsf{V}_{a+1})$.)
3) For $i = 1, \ldots, b$:
    a) The register B_i is sent to Bob, who applies B_i to the registers $(\mathsf{B}_i, \mathsf{Z}_{i-1})$ to obtain registers $(\mathsf{D}_i, \mathsf{Z}_i)$.
    b) The register D_i is then returned to the referee. If $i = b$ then go to step 4. Otherwise, the referee applies

V_{a+i} to the registers $(\mathsf{D}_i, \mathsf{V}_i)$ to obtain registers $(\mathsf{B}_{i+1}, \mathsf{V}_{i+1})$.

4) The referee applies the binary-valued measurement $\{\Pi, I - \Pi\}$ on the registers $(\mathsf{D}_b, \mathsf{V}_{a+b})$ with the outcome associated with Π indicating victory for Bob.

Basic quantum formalism tells us that if Alice and Bob act according to $\vec{A} = (A_1, \ldots, A_a)$ and $\vec{B} = (B_1, \ldots, B_b)$, respectively, then the probability with which Bob is declared the winner is given by

$$\begin{aligned}&\Pr[\text{Bob wins} \mid \vec{A}, \vec{B}]\\ &= \left\| \Pi B_b V_{a+b-1} \cdots B_1 V_a A_a V_{a-1} \cdots V_1 A_1 |\psi\rangle \right\|^2 . \end{aligned} \tag{3}$$

Of course, Bob wishes to maximize this quantity while Alice wishes to minimize this quantity. It follows immediately from the min-max theorem for zero-sum quantum games [24] that every double quantum interactive proof with referee R has a *value* $\lambda(R)$ given by

$$\begin{aligned}\lambda(R) &= \min_{\vec{A}} \max_{\vec{B}} \Pr[\text{Bob wins} \mid \vec{A}, \vec{B}] \qquad (4)\\ &= \max_{\vec{B}} \min_{\vec{A}} \Pr[\text{Bob wins} \mid \vec{A}, \vec{B}]\end{aligned}$$

where the minima are over all private spaces $\mathcal{W}_0, \ldots, \mathcal{W}_a$ for Alice and all unitaries $A_1, \ldots, A_a$ and the maxima are over all private spaces $\mathcal{Z}_0, \ldots, \mathcal{Z}_b$ for Bob and all unitaries $B_1, \ldots, B_b$. In particular, for every double quantum interactive proof with referee R there exist *optimal actions* $\vec{A^\star} = (A_1^\star, \ldots, A_a^\star)$ for Alice and $\vec{B^\star} = (B_1^\star, \ldots, B_b^\star)$ for Bob such that

$$\begin{aligned}&\Pr[\text{Bob wins} \mid \vec{A^\star}, \vec{B}] \leq \lambda(R) \qquad \text{for all } \vec{B},\\ &\Pr[\text{Bob wins} \mid \vec{A}, \vec{B^\star}] \geq \lambda(R) \qquad \text{for all } \vec{A}.\end{aligned}$$

From an operational perspective, the min-max expression (4) for $\lambda(R)$ in terms of tuples of unitaries $\vec{A}$ and $\vec{B}$ is natural and intuitive. However, this expression does not lend itself well to the MMW, which is designed to solve min-max problems over domains of *density operators*—not tuples of *unitaries*. To address this problem we derive an alternate expression for $\lambda(R)$ that is more amenable to the MMW.

To this end, for any $\vec{A}$ and $\vec{B}$ let ρ be the reduced state of the registers $(\mathsf{C}_a, \mathsf{V}_a)$ immediately after Alice's final unitary is applied and let P be the measurement operator on $(\mathsf{C}_a, \mathsf{V}_a)$ obtained by bundling the referee-Bob interaction into a single measurement operator as suggested by Figure 1. The expression (3) for Bob's probability of victory can be rewritten in terms of ρ, P as $\Pr[\text{Bob wins} \mid \vec{A}, \vec{B}] = \langle \rho, P \rangle$. Similarly, the expression (4) for $\lambda(R)$ can be rewritten[3] as

$$\lambda(R) = \min_{\rho \in \mathbf{A}(R)} \max_{P \in \mathbf{P}(R)} \langle \rho, P \rangle = \max_{P \in \mathbf{P}(R)} \min_{\rho \in \mathbf{A}(R)} \langle \rho, P \rangle$$

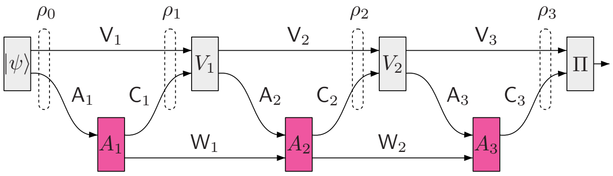

Figure 2. The states ρ_1, ρ_2, ρ_3 are a transcript of the referee's conversation with Alice. It follows easily from the unitary equivalence of purifications that a triple (ρ_1, ρ_2, ρ_3) is a valid transcript if and only if it obeys the recursive relation $\mathrm{Tr}_{\mathcal{C}_i}(\rho_i) = \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*)$ for $i = 1, 2, 3$ where $V_0 = I$.

where the sets $\mathbf{A}(R) \subset \mathbf{Dens}(\mathcal{C}_a\mathcal{V}_a)$ and $\mathbf{P}(R) \subset \mathbf{Meas}(\mathcal{C}_a\mathcal{V}_a)$ are given by

$$\begin{aligned}\mathbf{A}(R) = \{&\mathrm{Tr}_{\mathcal{W}_a}(|\phi\rangle\langle\phi|) : |\phi\rangle = A_a V_{a-1} \cdots V_1 A_1 |\psi\rangle \qquad (5)\\ &\text{for some } \vec{A}\},\\ \mathbf{P}(R) = \{&U^* \Pi U : U = B_b V_{a+b-1} \cdots B_1 V_a \qquad (6)\\ &\text{for some } \vec{B}\}.\end{aligned}$$

At this point, we have rewritten $\lambda(R)$ so that the set of all possible actions available to Alice has been identified with a subset $\mathbf{A}(R)$ of density operators, as desired. (Bob's actions will be addressed later.) However, the MMW is designed to solve min-max problems whose domain is the *entire* set of density operators. In the next section we present a new adaptation of the MMW that applies to min-max problems on strict subsets of density operators. We will see that this adaptation yields a parallel algorithm for the above formulation of $\lambda(R)$.

IV. Rounding theorem for a relaxed min-max problem

In this section we define a new min-max expression $\mu_\varepsilon(R)$ that approximates the desired quantity $\lambda(R)$ in the limit as ε approaches zero. The new expression is a relaxation of $\lambda(R)$ that is more amenable to the MMW. We prove a "rounding theorem" by which near-optimal points for $\lambda(R)$ are efficiently obtained from near-optimal points for $\mu_\varepsilon(R)$.

A. *Consistency conditions for Alice*

The set $\mathbf{A}(R)$ of density operators that represent admissible actions for Alice as defined in (5) is unwieldy. In order to optimize over this set we begin by writing it not in terms of unitaries $A_1, \ldots, A_a$ but in terms of states $\rho_1, \ldots, \rho_a$ that represent a *transcript* of the referee's conversation with Alice. Such a transcript is depicted in Figure 2. It is straightforward to use the unitary equivalence of purifications to characterize those density matrices which constitute valid transcripts. This characterization was first noted by Kitaev [19] and a formal proof can be found in Ref. [27].

Proposition 4 (Kitaev's consistency conditions—see Ref. [27])**.** *Let $R = (|\psi\rangle, V_1, \ldots, V_{a+b-1}, \Pi)$ be a referee*

[3] The fact that the ordering of minimization and maximization is immaterial follows from min-max theorem for zero-sum quantum games [24]. Or more directly, it follows from the well-known extensions of von Neumann's Min-Max Theorem [25], [26] given the fact that $\mathbf{A}(R), \mathbf{P}(R)$ are convex compact sets and $\langle \rho, P \rangle$ is a bilinear form over the two sets.

and let $\mathbf{A}(R)$ be the set of admissible states for Alice as defined in Eq. (5)*. A given state ρ is an element of $\mathbf{A}(R)$ if and only if there exist $\rho_i \in \mathbf{Dens}(\mathcal{C}_i\mathcal{V}_i)$ with $\rho_a = \rho$ and*

$$\mathrm{Tr}_{\mathcal{C}_i}(\rho_i) = \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*) \quad \textit{for } i = 1,\dots,a$$

where we have written $V_0 = I$ and $\rho_0 = |\psi\rangle\langle\psi|$ for convenience.

Any states $\rho_1,\dots,\rho_a$ obeying the consistency condition of Proposition 4 are said to be *consistent with R*. It therefore follows from Proposition 4 that the value $\lambda(R)$ of the game may be written[4]

$$\lambda(R) = \min_{\substack{(\rho_1,\dots,\rho_a) \\ \text{consistent with } R}} \max_{P\in\mathbf{P}(R)} \langle \rho_a, P\rangle . \tag{7}$$

B. A relaxed min-max problem and a rounding theorem

Define the relaxation $\mu_\varepsilon(R)$ of $\lambda(R)$ by

$$\begin{aligned}\mu_\varepsilon(R) \stackrel{\text{def}}{=} &\min_{(\rho_1,\dots,\rho_a)} \max_{\substack{P\in\mathbf{P}(R)\\(\Pi_1,\dots,\Pi_a)}} \langle\rho_a,P\rangle \\ &+ \frac{a}{\varepsilon}\sum_{i=1}^{a}\left\langle \mathrm{Tr}_{\mathcal{C}_i}(\rho_i) - \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*), \Pi_i\right\rangle \\ = &\min_{(\rho_1,\dots,\rho_a)} \max_{P\in\mathbf{P}(R)} \langle\rho_a,P\rangle \\ &+ \frac{a}{\varepsilon}\sum_{i=1}^{a}\frac{1}{2}\left\| \mathrm{Tr}_{\mathcal{C}_i}(\rho_i) - \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*)\right\|_{\mathrm{Tr}}\end{aligned}$$

Here the minimum is taken over all density operators $\rho_i \in \mathbf{Dens}(\mathcal{C}_i\mathcal{V}_i)$ and the maximum over all $P\in\mathbf{P}(R)$ and over all measurement operators $\Pi_i \in \mathbf{Meas}(\mathcal{V}_i)$.[5] The second equality follows immediately from the identity

$$\frac{1}{2}\|\rho-\xi\|_{\mathrm{Tr}} = \max_{0\preceq\Pi\preceq I}\langle\rho-\xi,\Pi\rangle \tag{8}$$

which holds for all density operators ρ,ξ.

Notice that the minimum in the definition of $\mu_\varepsilon(R)$ is taken over *all* density operators, not just those consistent with R. Each term in the summation serves to penalize any violation of consistency in the choice of $\rho_1,\dots,\rho_a$ by adding the magnitude of that violation to Bob's probability of victory. The a/ε factor amplifies the penalty so as to remove incentive for Alice to select an inconsistent course of action. Indeed, it is clear that $\lim_{\varepsilon\to0}\mu_\varepsilon(R) = \lambda(R)$. The following "rounding" theorem establishes a specific rate of convergence for this limit and a means by which near-optimal points for $\lambda(R)$ are efficiently computed from near-optimal points for $\mu_\varepsilon(R)$.

Theorem 5 (Rounding theorem)**.** *For any $\varepsilon > 0$ it holds that $\lambda(R) \geq \mu_\varepsilon(R) > \lambda(R) - \varepsilon$.*

Proof: The first inequality is easy: let $(\rho_1,\dots,\rho_a)$ be optimal for $\lambda(R)$ and let $(P,\Pi_1,\dots,\Pi_a)$ be optimal for $\mu_\varepsilon(R)$. Then we have $\lambda(R)\geq\langle\rho_a,P\rangle$, which equals

$$\langle\rho_a,P\rangle + \frac{a}{\varepsilon}\sum_{i=1}^{a}\left\langle \mathrm{Tr}_{\mathcal{C}_i}(\rho_i) - \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*), \Pi_i\right\rangle$$

because $\rho_1,\dots,\rho_a$ are consistent with R, so each term in the sum is zero. This quantity is at least $\mu_\varepsilon(R)$ because P is optimal for $\mu_\varepsilon(R)$.

The second inequality is more difficult. We invoke the following lemma, the proof of which appears in Section IV-C.

Lemma 6. *Fix any $\varepsilon > 0$, any referee $R = (|\psi\rangle, V_1,\dots,V_{a+b-1},\Pi)$, and any states $\rho_1,\dots,\rho_a$ where each ρ_i is an element of $\mathbf{Dens}(\mathcal{C}_i\mathcal{V}_i)$. There exist states $\rho_1',\dots,\rho_a'$ consistent with R such that $\frac{1}{2}\|\rho_a-\rho_a'\|_{\mathrm{Tr}}$ is less than*

$$\varepsilon + \frac{a}{\varepsilon}\sum_{i=1}^{a}\frac{1}{2}\|\mathrm{Tr}_{\mathcal{C}_i}(\rho_i) - \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*)\|_{\mathrm{Tr}}.$$

Moreover, $\rho_1',\dots,\rho_a'$ can be computed efficiently in parallel given $\rho_1,\dots,\rho_a$.

Let $(\rho_1,\dots,\rho_a)$ be optimal for $\mu_\varepsilon(R)$, let $\rho_1',\dots,\rho_a'$ be the density operators obtained by invoking Lemma 6, and let $P\in\mathbf{P}(R)$ be optimal for $\lambda(R)$. Because $(\rho_1,\dots,\rho_a)$ is optimal for $\mu_\varepsilon(R)$ we have that $\mu_\varepsilon(R)$ is at least

$$\langle\rho_a,P\rangle + \frac{a}{\varepsilon}\sum_{i=1}^{a}\frac{1}{2}\left\| \mathrm{Tr}_{\mathcal{C}_i}(\rho_i) - \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*)\right\|_{\mathrm{Tr}} \tag{9}$$

Employing the identity (8), the quantity $\langle\rho_a,P\rangle$ becomes

$$\langle\rho_a',P\rangle + \langle\rho_a-\rho_a',P\rangle \geq \langle\rho_a',P\rangle - \frac{1}{2}\|\rho_a-\rho_a'\|_{\mathrm{Tr}}.$$

Substituting the bound on $\frac{1}{2}\|\rho_a-\rho_a'\|_{\mathrm{Tr}}$ from Lemma 6, we see that the summation in (9) is canceled, so that this quantity is larger than $\langle\rho_a',P\rangle-\varepsilon$. The result follows from the fact that P is optimal for $\lambda(R)$. ■

Corollary 5.1 (Rounding theorem—construction of near-optimal strategies)**.** *The following hold for any referee R and any $\varepsilon,\delta > 0$:*

1) *If $(\rho_1,\dots,\rho_a)$ is δ-optimal for $\mu_\varepsilon(R)$ then there exist density operators $(\rho_1',\dots,\rho_a')$ consistent with R that can be computed efficiently in parallel such that $\rho_a'\in\mathbf{A}(R)$ is $(\delta+\varepsilon)$-optimal for $\lambda(R)$.*
2) *If $(P,\Pi_1,\dots,\Pi_a)$ is δ-optimal for $\mu_\varepsilon(R)$ then $P\in\mathbf{P}(R)$ is also $(\delta+\varepsilon)$-optimal for $\lambda(R)$.*

Proof of item 1: Let $(\rho_1,\dots,\rho_a)$ be δ-optimal for $\mu_\varepsilon(R)$, let $\rho_1',\dots,\rho_a'$ be the density operators obtained by

[4] Again, the ordering of the minimum and maximum is immaterial because the two sets are compact and convex and the objective function is bilinear.

[5] As usual, the ordering of the minimum and maximum is immaterial. In this case, the additional variables $(\Pi_1,\dots,\Pi_a)$ come from a compact convex set and the new sum in the objective function is also a bilinear form.

invoking Lemma 6, and let $P \in \mathbf{P}(R)$. We must show that $\langle \rho'_a, P\rangle \leq \lambda(R) + \varepsilon + \delta$. To this end, (8) implies

$$\langle \rho'_a, P\rangle \leq \langle \rho_a, P\rangle + \frac{1}{2}\left\|\rho_a - \rho'_a\right\|_{\mathrm{Tr}}$$

Substituting the bound on $\frac{1}{2}\left\|\rho_a - \rho'_a\right\|_{\mathrm{Tr}}$ from Lemma 6, we see that this quantity is at most

$$\langle \rho_a, P\rangle + \varepsilon + \frac{a}{\varepsilon}\sum_{i=1}^{a}\frac{1}{2}\|\mathrm{Tr}_{\mathcal{C}_i}(\rho_i) - \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*)\|_{\mathrm{Tr}}$$

As $(\rho_1, \ldots, \rho_a)$ is δ-optimal for $\mu_\varepsilon(R)$ this quantity is at most $\mu_\varepsilon(R) + \varepsilon + \delta$, which is at most $\lambda(R) + \varepsilon + \delta$ because $\mu_\varepsilon(R) \leq \lambda(R)$. ■

Proof of item 2: Let $(P, \Pi_1, \ldots, \Pi_a)$ be δ-optimal for $\mu_\varepsilon(R)$ and let $(\rho_1, \ldots, \rho_a)$ be consistent with R. We must show $\langle \rho_a, P\rangle \geq \lambda(R) - \varepsilon - \delta$. To this end note that $\langle \rho_a, P\rangle$ equals

$$\langle \rho_a, P\rangle + \frac{a}{\varepsilon}\sum_{i=1}^{a}\left\langle \mathrm{Tr}_{\mathcal{C}_i}(\rho_i) - \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*), \Pi_i\right\rangle$$

because $(\rho_1, \ldots, \rho_a)$ are consistent with R, so each term in the sum is zero. This quantity is at least $\mu_\varepsilon(R) - \delta$ because $(P, \Pi_1, \ldots, \Pi_a)$ is δ-optimal for $\mu_\varepsilon(R)$. The result follows from the fact that $\mu_\varepsilon(R) > \lambda(R) - \varepsilon$. ■

C. Rounding lemma for obtaining consistent states

In this subsection we prove Lemma 6, which appeared in the proof of Theorem 5. Given any states $(\rho_1, \ldots, \rho_a)$ this lemma asserts that these states can be "rounded" to valid transcript states $(\rho'_1, \ldots, \rho'_a)$ in such a way that the distance between the final states ρ_a and ρ'_a is bounded by a function of the extent to which $(\rho_1, \ldots, \rho_a)$ violate the consistency condition of Proposition 4. The proof of this lemma is interesting because it provides a nontrivial application of the Bures angle. Let us first re-state Lemma 6 in terms of the Bures angle.

Lemma 7 (Rounding lemma)**.** *Fix any referee $R = (|\psi\rangle, V_1, \ldots, V_{a+b-1}, \Pi)$ and any states $\rho_1, \ldots, \rho_a$ where each ρ_i is an element of $\mathbf{Dens}(\mathcal{C}_i\mathcal{V}_i)$. There exist $\rho'_1, \ldots, \rho'_a$ consistent with R such that*

$$A(\rho_a, \rho'_a) \leq \sum_{i=1}^{a} A\left(\mathrm{Tr}_{\mathcal{C}_i}(\rho_i), \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*)\right).$$

Moreover, $\rho'_1, \ldots, \rho'_a$ can be computed efficiently in parallel given $\rho_1, \ldots, \rho_a$.

Proof: Define $\rho'_1, \ldots, \rho'_a$ recursively as follows. Let $\rho'_0 = \rho_0$. For each $i = 1, \ldots, a$ by the preservation of subsystem fidelity (Proposition 2) there exists ρ'_i (which can be computed efficiently in parallel) with $\mathrm{Tr}_{\mathcal{C}_i}(\rho'_i) = \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho'_{i-1}V_{i-1}^*)$ and

$$A(\rho_i, \rho'_i) = A\left(\mathrm{Tr}_{\mathcal{C}_i}(\rho_i), \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho'_{i-1}V_{i-1}^*)\right).$$

By the triangle inequality this quantity is at most

$$\begin{aligned}&A\left(\mathrm{Tr}_{\mathcal{C}_i}(\rho_i), \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*)\right)\\ &\quad + A\left(\mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}V_{i-1}^*), \mathrm{Tr}_{\mathcal{A}_i}(V_{i-1}\rho'_{i-1}V_{i-1}^*)\right).\end{aligned}$$

By contractivity of the Bures angle under channels, the summand on the right is at most $A(\rho_{i-1}, \rho'_{i-1})$. The lemma now follows inductively from the fact that $A(\rho_0, \rho'_0) = 0$. ■

It is easy to recover Lemma 6 from Lemma 7: it follows immediately from Lemma 7 and Proposition 3 (Relationship between trace norm and Bures angle) that

$$\frac{1}{2}\|\rho_a - \rho'_a\|_{\mathrm{Tr}} \leq \sum_{i=0}^{a-1}\sqrt{\frac{\pi}{2}\|\mathrm{Tr}_{\mathcal{C}}(\rho_{i+1}) - \mathrm{Tr}_{\mathcal{C}}(V_i\rho_iV_i^*)\|_{\mathrm{Tr}}}.$$

Lemma 6 then follows from the fact that $\sqrt{\frac{\pi}{2}x} < \frac{1}{2\delta}x + \delta$ for all $x \geq 0$ and all $\delta > 0$.

V. The MMW oracle-algorithm for double quantum interactive proofs

In this section we describe an efficient parallel oracle-algorithm that approximates $\lambda(R)$ to arbitrary precision. It is a simple matter to modify this algorithm so as to also produce unitaries $A_1, \ldots, A_a$ for Alice and $B_1, \ldots, B_b$ for Bob (*i.e.* strategies for both players) that are arbitrarily close to optimal. Details are provided in the full version of this paper.

A. Formal statement of the problem, review of the MMW

Precise statements of the problem solved by our algorithm and the oracle it requires are given below. For matrix inputs, each entry is written explicitly. The real and complex parts of all numbers are written as rational numbers in binary.

Problem 1 (Approximation of $\lambda(R)$)**.**

Input: A referee R and an accuracy parameter $\delta > 0$.

Oracle: Weak optimization for $\mathbf{P}(R)$. (See Problem 2.)

Output: A number $\tilde{\lambda}$ with $|\tilde{\lambda} - \lambda(R)| < \delta$.

Problem 2 (Weak optimization for $\mathbf{P}(R)$)**.**

Input: A referee R, a state $\rho \in \mathbf{Dens}(\mathcal{C}_a\mathcal{V}_a)$, and an accuracy parameter $\delta > 0$.

Output: A measurement operator $\tilde{P} \in \mathbf{P}(R)$ such that $\langle \rho, \tilde{P}\rangle \geq \langle \rho, P\rangle - \delta$ for every $P \in \mathbf{P}(R)$.

The precise formulation of the MMW used in this paper is stated below as Theorem 8. Our statement of this theorem is somewhat nonstandard: the result is usually presented in the form of an algorithm, whereas our presentation is purely mathematical. However, a cursory examination of the literature—say, Kale's thesis [16, Chapter 3]—reveals that our mathematical formulation is equivalent to the more conventional algorithmic form.

Theorem 8 (Multiplicative weights update method—see Ref. [16, Theorem 10]). *Fix $\gamma \in (0,1/2)$ and $\alpha > 0$. Let $M^{(1)},\dots,M^{(T)}$ be arbitrary $D \times D$ "loss" matrices with $0 \preceq M^{(t)} \preceq \alpha I$. Let $W^{(1)},\dots,W^{(T)}$ be $D \times D$ "weight" matrices given by $W^{(1)} = I$,*

$$W^{(t+1)} = \exp\left(-\gamma\left(M^{(1)} + \cdots + M^{(t)}\right)\right).$$

Let $\rho^{(1)},\dots,\rho^{(T)}$ be density operators obtained by normalizing each $W^{(1)},\dots,W^{(T)}$ so that $\rho^{(t)} = W^{(t)}/\operatorname{Tr}(W^{(t)})$. For all density operators ρ it holds that

$$\frac{1}{T}\sum_{t=1}^{T}\left\langle \rho^{(t)}, M^{(t)}\right\rangle \le \left\langle \rho, \frac{1}{T}\sum_{t=1}^{T} M^{(t)}\right\rangle + \alpha\left(\gamma + \frac{\ln D}{\gamma T}\right).$$

Note that Theorem 8 holds for *all* choices of loss matrices $M^{(1)},\dots,M^{(T)}$, including those for which each $M^{(t)}$ is chosen adversarially based upon $W^{(1)},\dots,W^{(t)}$. This adaptive selection of loss matrices is typical in implementations of the MMW.

B. Statement and analysis of the MMW oracle-algorithm

Let $\varepsilon > 0$ and consider the linear mapping

$$\begin{aligned} f_{R,\varepsilon} : (\rho_1,\dots,\rho_a) \mapsto \Big(&\rho_a,\\ &\frac{a}{\varepsilon}\left[\operatorname{Tr}_{\mathcal{C}_a}(\rho_a) - \operatorname{Tr}_{\mathcal{A}_a}(V_{a-1}\rho_{a-1}V_{a-1}^*)\right],\dots,\\ &\frac{a}{\varepsilon}\left[\operatorname{Tr}_{\mathcal{C}_2}(\rho_2) - \operatorname{Tr}_{\mathcal{A}_2}(V_1\rho_1V_1^*)\right],\\ &\frac{a}{\varepsilon}\left[\operatorname{Tr}_{\mathcal{C}_1}(\rho_1) - \operatorname{Tr}(\rho_1)\operatorname{Tr}_{\mathcal{A}_1}(|\psi\rangle\langle\psi|)\right]\Big) \end{aligned}$$

It is clear that $\mu_\varepsilon(R)$ equals

$$\min_{(\rho_1,\dots,\rho_a)} \max_{\substack{P\in\mathbf{P}(R)\\(\Pi_1,\dots,\Pi_a)}} \left\langle f_{R,\varepsilon}(\rho_1,\dots,\rho_a), (P,\Pi_a,\dots,\Pi_1)\right\rangle.$$

It is tedious but straightforward to compute the adjoint map:

$$\begin{aligned} f_{R,\varepsilon}^* : (P,\Pi_a,\dots,\Pi_1) \mapsto \Big(&P + \frac{a}{\varepsilon}\Pi_a \otimes I_{\mathcal{C}_a},\\ &\frac{a}{\varepsilon}\left[\Pi_{a-1}\otimes I_{\mathcal{C}_{a-1}} - V_{a-1}^*(\Pi_a \otimes I_{\mathcal{A}_a})V_{a-1}\right],\dots,\\ &\frac{a}{\varepsilon}\left[\Pi_2 \otimes I_{\mathcal{C}_2} - V_2^*(\Pi_3\otimes I_{\mathcal{A}_3})V_2\right],\\ &\frac{a}{\varepsilon}\left[\Pi_1\otimes I_{\mathcal{C}_1} - V_1^*(\Pi_2\otimes I_{\mathcal{A}_2})V_1 - \langle\psi|\Pi_1|\psi\rangle I_{\mathcal{C}_1\mathcal{V}_1}\right]\Big) \end{aligned}$$

The statement of our MMW algorithm in Figure 3 employs this formula for the adjoint. Hereafter we write $D = \max_i\{\dim(\mathcal{C}_i\mathcal{V}_i)\}$.

1) Let $\varepsilon = \delta/2$, let $\gamma = \frac{\varepsilon\delta}{16a^2}$, and let $T = \left\lceil \frac{\ln D}{\gamma^2}\right\rceil$. Let $W_i^{(1)} = I_{\mathcal{C}_i\mathcal{V}_i}$ for each $i = 1,\dots,a$.
2) Repeat for each $t = 1,\dots,T$:
    a) For $i = 1,\dots,a$: Compute the updated density operators $\rho_i^{(t)} = W_i^{(t)}/\operatorname{Tr}(W_i^{(t)})$.
    b) For $i = 1,\dots,a$: Compute the projection $\Pi_i^{(t)}$ onto the positive eigenspace of
    $$\operatorname{Tr}_{\mathcal{C}_i}(\rho_i^{(t)}) - \operatorname{Tr}_{\mathcal{A}_i}(V_{i-1}\rho_{i-1}^{(t)}V_{i-1}^*).$$
    c) Use the oracle to obtain a $\delta/2$-optimal solution $P^{(t)}$ to the Weak optimization problem for $\mathbf{P}(R)$ (Problem 2) on input $\rho_a^{(t)}$.
    d) Compute the loss matrices
    $$\left(M_a^{(t)},\dots,M_1^{(t)}\right) = \frac{\varepsilon}{4a^2}\left[f_{R,\varepsilon}^*(P^{(t)},\Pi_a^{(t)},\dots,\Pi_1^{(t)}) + \frac{2a}{\varepsilon}(I,\dots,I)\right]$$
    so that each loss matrix $M_i^{(t)}$ satisfies $0 \preceq M_i^{(t)} \preceq \frac{1}{a}I$.
    e) Update each weight matrix according to the standard MMW update rule:
    $$W_i^{(t+1)} = \exp\left(-\gamma\left(M_i^{(1)} + \cdots + M_i^{(t)}\right)\right).$$
3) Return
$$\frac{1}{T}\sum_{t=1}^{T}\left\langle f_{R,\varepsilon}(\rho_1^{(t)},\dots,\rho_a^{(t)}), (P^{(t)},\Pi_a^{(t)},\dots,\Pi_1^{(t)})\right\rangle$$
as the δ-approximation $\tilde{\lambda}$ to $\lambda(R)$.
4) If optimal strategies are desired then compute
$$(\rho_1,\dots,\rho_a) = \frac{1}{T}\sum_{t=1}^{T}(\rho_1^{(t)},\dots,\rho_a^{(t)})$$
$$(P,\Pi_a,\dots,\Pi_1) = \frac{1}{T}\sum_{t=1}^{T}(P^{(t)},\Pi_a^{(t)},\dots,\Pi_1^{(t)}),$$
both of which are δ-optimal for $\mu_\varepsilon(R)$. Compute $(\rho_1',\dots,\rho_a')$ from $(\rho_1,\dots,\rho_a)$ as described in item 1 of Corollary 5.1. Return ρ_a' and P, both of which are $3\delta/2$-optimal for $\lambda(R)$.

Figure 3. An efficient parallel oracle-algorithm for approximating $\lambda(R)$ (Problem 1).

Proposition 9. *The oracle-algorithm presented in Figure 3 approximates $\lambda(R)$ to precision δ (Problem 1). Assuming unit cost for the oracle, this algorithm can be implemented in parallel with run time bounded by a polynomial in $a + b$, $1/\delta$, and $\log D$.*

Proof: First, we note the fact that each loss matrix $M_i^{(t)}$ satisfies $0 \preceq M_i^{(t)} \preceq \frac{1}{a}I$ follows immediately from its definition in step 2d and the observation that the adjoint mapping $f^*_{R,\varepsilon}$ satisfies

$$\begin{aligned}\left(0, -\frac{a}{\varepsilon}I, \dots, -\frac{a}{\varepsilon}I, -2\frac{a}{\varepsilon}I\right) &\preceq f^*_{R,\varepsilon}(P, \Pi_a, \dots, \Pi_1)\\ &\preceq \left(\left(1+\frac{a}{\varepsilon}\right)I, \frac{a}{\varepsilon}I, \dots, \frac{a}{\varepsilon}I\right).\end{aligned}$$

For each $i = 1, \dots, a$ it is clear that the construction of the density operators $\rho_i^{(t)}$ in terms of the loss matrices $M_i^{(t)}$ presented in Figure 3 obeys the condition of Theorem 8. It therefore follows that for any $\rho_i^\star \in \mathbf{Dens}(\mathcal{C}_i\mathcal{V}_i)$ we have

$$\frac{1}{T}\sum_{t=1}^{T}\left\langle \rho_i^{(t)}, M_i^{(t)}\right\rangle \leq \left\langle \rho_i^\star, \frac{1}{T}\sum_{t=1}^{T} M_i^{(t)}\right\rangle + \frac{1}{a}\left(\gamma + \frac{\ln D}{\gamma T}\right).$$

Summing these inequalities over all i we find that for any density operators $(\rho_1^\star, \dots, \rho_a^\star)$ it holds that

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\left\langle \left(\rho_1^{(t)}, \dots, \rho_a^{(t)}\right), \left(M_1^{(t)}, \dots, M_a^{(t)}\right)\right\rangle \leq\\ &\left\langle (\rho_1^\star, \dots, \rho_a^\star), \frac{1}{T}\sum_{t=1}^{T}\left(M_1^{(t)}, \dots, M_a^{(t)}\right)\right\rangle + \left(\gamma + \frac{\ln D}{\gamma T}\right).\end{aligned}$$

Substituting the definition of the loss matrices $M_i^{(t)}$ from step 2d and simplifying, we have that $\tilde{\lambda}$ equals

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\left\langle (\rho_1^{(t)}, \dots, \rho_a^{(t)}), f^*_{R,\varepsilon}(P^{(t)}, \Pi_a^{(t)}, \dots, \Pi_1^{(t)})\right\rangle \qquad (10)\\ &\leq \left\langle (\rho_1^\star, \dots, \rho_a^\star), \frac{1}{T}\sum_{t=1}^{T} f^*_{R,\varepsilon}\left(P^{(t)}, \Pi_a^{(t)}, \dots, \Pi_1^{(t)}\right)\right\rangle\\ &\quad + \underbrace{\frac{4a^2}{\varepsilon}\left(\gamma + \frac{\ln D}{\gamma T}\right)}_{\text{error term}}.\end{aligned}$$

Substituting the choice of γ, T from step 1 we see that the error term on the right side is at most $\delta/2$. Since this inequality holds for any choice of $(\rho_1^\star, \dots, \rho_a^\star)$ it certainly holds for the optimal choice, from which it follows that the right side is at most $\mu_\varepsilon(R) + \delta/2$. By construction each $(P^{(t)}, \Pi_a^{(t)}, \dots, \Pi_1^{(t)})$ is a $\delta/2$-best response to $(\rho_1^{(t)}, \dots, \rho_a^{(t)})$ so it must be that the left side of this inequality is at least $\mu_\varepsilon(R) - \delta/2$. It then follows from Theorem 5 (Rounding theorem) and the choice $\varepsilon = \delta/2$ that $|\tilde{\lambda} - \lambda(R)| < \delta$ as desired.

Next we argue that the density operator ρ_a' returned in step 4 is $3\delta/2$-optimal for $\lambda(R)$. By item 1 of Corollary 5.1 it suffices to argue that $(\rho_1, \dots, \rho_a)$ are δ-optimal for $\mu_\varepsilon(R)$. To this end, choose any $(P^\star, \Pi_1^\star, \dots, \Pi_a^\star)$. Since each $(P^{(t)}, \Pi_a^{(t)}, \dots, \Pi_1^{(t)})$ is a $\delta/2$-best response to $(\rho_1^{(t)}, \dots, \rho_a^{(t)})$ it holds that the inner product

$$\left\langle (\rho_1^{(t)}, \dots, \rho_a^{(t)}), f^*_{R,\varepsilon}(P^{(t)}, \Pi_a^{(t)}, \dots, \Pi_1^{(t)})\right\rangle$$

can increase by no more than $\delta/2$ when $(P^\star, \Pi_1^\star, \dots, \Pi_a^\star)$ is substituted for $(P^{(t)}, \Pi_a^{(t)}, \dots, \Pi_1^{(t)})$. It then follows from (10) that

$$\begin{aligned}&\left\langle \frac{1}{T}\sum_{t=1}^{T}\left(\rho_1^{(t)}, \dots, \rho_a^{(t)}\right), f^*_{R,\varepsilon}\left(P^\star, \Pi_a^\star, \dots, \Pi_1^\star\right)\right\rangle\\ &\leq \tilde{\lambda} + \delta/2 \leq \mu_\varepsilon(R) + \delta\end{aligned}$$

and hence $(\rho_1, \dots, \rho_a)$ is δ-optimal for $\mu_\varepsilon(R)$ as desired.

Next we argue that the operator P returned in step 4 is $3\delta/2$-optimal for $\lambda(R)$. By item 2 of Corollary 5.1 it suffices to argue that $(P, \Pi_a, \dots, \Pi_1)$ are δ-optimal for $\mu_\varepsilon(R)$. To this end, choose any $(\rho_1^\star, \dots, \rho_a^\star)$. It follows from (10) that

$$\left\langle (\rho_1^\star, \dots, \rho_a^\star), f^*_{R,\varepsilon}\left(P, \Pi_a, \dots, \Pi_1\right)\right\rangle \geq \tilde{\lambda} - \delta/2 \geq \mu_\varepsilon(R) - \delta$$

and hence $(P, \Pi_a, \dots, \Pi_1)$ is δ-optimal for $\mu_\varepsilon(R)$ as desired.

The efficiency of this algorithm is not difficult to argue. Each individual step consists only of matrix operations that are known to admit an efficient parallel implementation. Efficiency then follows from the observation that the number T of iterations is polynomial in $a + b$, $1/\delta$, and $\log D$. ■

C. Special case: semidefinite programs on consistent density operators, a direct simulation of QIP

Consider a special case of the problem of approximating $\lambda(R)$ (Problem 1) in which $b = 0$. Since there is no interaction with Bob, this scenario corresponds to an ordinary, single-prover quantum interactive proof. In this case, $\mathbf{P}(R) = \{\Pi\}$ is a singleton set and the expression (7) for $\lambda(R)$ simplifies to

$$\lambda(R) = \min_{\substack{(\rho_1, \dots, \rho_a)\\ \text{consistent with } R}} \langle \rho_a, \Pi\rangle, \qquad (11)$$

which is a semidefinite program whose feasible region consists of density operators $\rho_1, \dots, \rho_a$ consistent with R. Since $\mathbf{P}(R)$ is a singleton set, the oracle for weak optimization over $\mathbf{P}(R)$ is trivial to implement so it follows immediately from Proposition 9 that the algorithm presented in Figure 3 can be used to solve SDPs of this form efficiently in parallel and thus prove QIP = PSPACE via direct simulation of a multi-message quantum interactive proof.

Later we will see that the oracle for weak optimization for $\mathbf{P}(R)$ (Problem 2) required for general instances of $\lambda(R)$ can be reduced to an instance of this SDP special case of Problem 1 plus some post-processing.

An arbitrary SDP having the form (1) from the beginning of this paper can easily be transformed into an equivalent SDP of the above form (11) (and therefore solved by the algorithm in Figure 3). Such a transformation also establishes Theorem 1 from the beginning of this paper.

VI. Implementation of the oracle

In Section V we presented a parallel oracle-algorithm (Figure 3) for the problem of approximating $\lambda(R)$ (Problem 1) and proved its correctness and efficiency (Proposition 9). In order to complete the description of our algorithm for double quantum interactive proofs it remains only to describe the implementation of the oracle for weak optimization for $\mathbf{P}(R)$ (Problem 2). In this section we establish the following.

Proposition 10. *The weak optimization problem (Problem 2) for the set $\mathbf{P}(R)$ specified in Eq. (6) admits a parallel algorithm with run time bounded by a polynomial in $a+b$, $1/\delta$, and $\log D$.*

It follows that the algorithm of Figure 3 is an unconditionally efficient parallel algorithm for approximating $\lambda(R)$ (Problem 1).

As mentioned earlier, this instance of Problem 2 will be rephrased as a new instance of Problem 1 (plus some post-processing) so that the algorithm of Section V can be reused in the implementation of our oracle. Incidentally, we shall see that this new instance of Problem 1 has the special SDP form described in Section V-C.

Choose any state ρ and suppose that a (possibly cheating) Alice was somehow able to make it so that the state of the registers $(\mathsf{C}_a, \mathsf{V}_a)$ after the interaction with Alice is in state ρ. Let W_a be a register large enough to admit a purification of ρ and let $|\varphi\rangle \in \mathcal{W}_a\mathcal{C}_a\mathcal{V}_a$ be any such purification. If Bob acts according to $(B_1, \dots, B_b)$ then (similar to Eq. (3)) his probability of victory is

$$\left\| \Pi B_b V_{a+b-1} B_{b-1} \cdots B_1 V_a |\varphi\rangle \right\|^2 .$$

Notice that this quantity also represents the probability of victory in a different, one-player game with a referee R' whose initial state is $V_a|\varphi\rangle$. (Formally, the referee R' exchanges b rounds of messages with one of the players and zero messages with the other.) The unitaries $B_1, \dots, B_b$ could specify actions for either Alice or Bob—a choice that depends only upon how we label the components of the referee R'.

Since our goal is to reduce Problem 2 to an instance of the SDP special case of Problem 1 (which is a minimization problem), it befits us to view $B_1, \dots, B_b$ as actions for Alice in the game with referee R'. Let us write $R' = (V_a|\varphi\rangle, V'_1, \dots, V'_{b-1}, \Pi')$ where each $V'_i = V_{a+i} \otimes I_{\mathcal{W}_a}$ and $\Pi' = (I - \Pi) \otimes I_{\mathcal{W}_a}$. Each of the private memory registers V'_i of the new referee R' is identified with the registers $(\mathsf{V}_{a+i}, \mathsf{W}_a)$ and each of the message registers $\mathsf{A}'_i, \mathsf{C}'_i$ of the new referee is identified with $\mathsf{B}_i, \mathsf{D}_i$, respectively.

In this case, the set $\mathbf{P}(R') = \{\Pi'\}$ is a singleton set. Each choice of unitaries $B_1, \dots, B_b$ induces both a measurement operator $P \in \mathbf{P}(R)$ and a state $\xi \in \mathbf{A}(R')$ with

$$\langle \rho, P \rangle = \left\| \Pi B_b V_{a+b-1} B_{b-1} \cdots B_1 V_a |\varphi\rangle \right\|^2 = 1 - \langle \xi, \Pi' \rangle$$

and therefore

$$\max_{P \in \mathbf{P}(R)} \langle \rho, P \rangle = 1 - \lambda(R') = 1 - \min_{\xi \in \mathbf{A}(R')} \langle \xi, \Pi' \rangle .$$

Moreover, if $P \in \mathbf{P}(R)$ achieves the maximum on the left side then the unitaries $B_1, \dots, B_b$ that induce P also induce a state $\xi \in \mathbf{A}(R')$ that achieves the minimum on the right side. As the right side is an instance of the SDP special case of Problem 1 a solution to Problem 2 presents itself:

1) Use the algorithm of Figure 3 to find $\xi \in \mathbf{A}(R')$ minimizing $\langle \xi, \Pi' \rangle$.
2) Find the unitaries $B_1, \dots, B_b$ that induce ξ. These unitaries also induce a measurement operator $P \in \mathbf{P}(R)$ maximizing $\langle \rho, P \rangle$. Compute P using $B_1, \dots, B_b$ via standard matrix multiplication.

We already saw in Section V how the algorithm of Figure 3 can be used to accomplish step 1. In the remainder of this section we fill in the details for step 2. Recall that the algorithm of Figure 3 finds near-optimal density operators $\xi_1, \dots, \xi_b$ consistent with the referee $R' = (V_a|\varphi\rangle, V'_1, \dots, V'_{b-1}, \Pi')$, meaning that $\mathrm{Tr}_{\mathcal{C}'_i}(\xi_i) = \mathrm{Tr}_{\mathcal{A}'_i}(V'_{i-1}\xi_{i-1}V'^{*}_{i-1})$ for each $i = 1, \dots, b$ where $V'_0 = V_a \otimes I_{\mathcal{W}_a}$ and $\xi_0 = |\varphi\rangle\langle\varphi|$ for convenience. The following algorithm finds the unitaries $B_1, \dots, B_b$.

1) Let $\mathcal{W}'_1, \dots, \mathcal{W}'_b$ be spaces large enough to admit purifications of $\xi_1, \dots, \xi_b$ and write $|\alpha_0\rangle = |\varphi\rangle|0_{\mathcal{W}'_0}\rangle$ for some space $\mathcal{W}'_0$.
2) For each $i = 1, \dots, b$:
   a) Compute a purification $|\alpha_i\rangle \in \mathcal{C}'_i\mathcal{V}'_i\mathcal{W}'_i$ of ξ_i.
   b) Compute a unitary $B_i : \mathcal{A}'_i\mathcal{W}'_{i-1} \to \mathcal{C}'_i\mathcal{W}'_i$ that maps $V'_{i-1}|\alpha_{i-1}\rangle$ to $|\alpha_i\rangle$.
3) Return the desired unitaries $(B_1, \dots, B_b)$.

Correctness of this construction is straightforward (though notationally cumbersome). As in Proposition 9, each individual step consists only of matrix operations that are known to admit an efficient parallel implementation, from which it follows that the entire construction is efficient.

VII. Containment of DQIP inside PSPACE

A decision problem L is said to admit a double quantum interactive proof with *completeness* $c(|x|)$ and *soundness* $s(|x|)$ if there exists a polynomial-time uniform quantum referee R_x such that $x \in L \implies \lambda(R_x) \geq c(|x|)$ and $x \notin L \implies \lambda(R_x) \leq s(|x|)$. The complexity class DQIP consists of all decision problems L that admit double quantum interactive proofs for which there exists a polynomial-bounded function $p(|x|)$ such that $c - s \geq 1/p$. The class DIP is defined similarly except that the referee is classical. By definition it holds that DIP $\subseteq$ DQIP.

Proposition 11. DQIP $\subseteq$ PSPACE.

Proof sketch: Let $L \in$ DQIP and let R_x be a referee witnessing this fact. Membership of an arbitrary instance x

of L can be decided by computing a sufficiently accurate approximation of $\lambda(R_x)$ via an efficient parallel implementation of the algorithm of Figure 3 on an exponential-size explicit description of R_x. This process can be simulated in polynomial space by standard methods (see footnote 1), from which it follows that $L \in$ PSPACE and hence DQIP $\subseteq$ PSPACE. ■

The characterization DQIP = DIP = PSPACE now follows immediately from the well known fact that IP = PSPACE and from the trivial containment IP $\subseteq$ DIP.

Acknowledgements

The authors are grateful to Tsuyoshi Ito, Rahul Jain, Zhengfeng Ji, Yaoyun Shi, Sarvagya Upadhyay, and John Watrous for helpful discussions. Particularly, the alternative formulation of the strategies by density operators and measurements is inspired during the discussion with John Watrous. XW also wants to thank the hospitality and invaluable guidance of John Watrous when he was visiting the Institute for Quantum Computing, University of Waterloo. The research was partially conducted during this visit and was supported by the Canadian Institute for Advanced Research (CIFAR). XW's research is also supported by NSF grant 1017335. GG's research is supported by the Government of Canada through Industry Canada, the Province of Ontario through the Ministry of Research and Innovation, NSERC, DTO-ARO, CIFAR, and QuantumWorks.

References

[1] M. Serna, "Approximating linear programming is log-space complete for P," *Information Processing Letters*, vol. 37, no. 4, pp. 233–236, 1991.

[2] N. Megiddo, "A note on approximate linear programming," *Information Processing Letters*, vol. 42, no. 1, p. 53, 1992.

[3] R. Jain and P. Yao, "A parallel approximation algorithm for positive semidefinite programming." in *FOCS*, 2011, pp. 463–471, arXiv:1104.2502v1 [cs.CC].

[4] A. Borodin, "On relating time and space to size and depth," *SIAM Journal on Computing*, vol. 6, no. 4, pp. 733–744, 1977.

[5] U. Feige and J. Kilian, "Making games short," in *STOC*, 1997, pp. 506–516.

[6] C. Lund, L. Fortnow, H. Karloff, and N. Nisan, "Algebraic methods for interactive proof systems," *Journal of the ACM*, vol. 39, no. 4, pp. 859–868, 1992.

[7] A. Shamir, "IP = PSPACE," *Journal of the ACM*, vol. 39, no. 4, pp. 869–877, 1992.

[8] R. Jain, Z. Ji, S. Upadhyay, and J. Watrous, "QIP=PSPACE," in *STOC*, 2010, pp. 573–582, arXiv:0907.4737v2 [quant-ph].

[9] X. Wu, "Equilibrium value method for the proof of QIP=PSPACE," 2010, arXiv:1004.0264v2 [quant-ph].

[10] G. Gutoski and J. Watrous, "Quantum interactive proofs with competing provers," in *STACS*, 2005, pp. 605–616, arXiv:cs/0412102v1 [cs.CC].

[11] S. Goldwasser and M. Sipser, "Private coins versus public coins in interactive proof systems," in *Randomness and Computation*, ser. Advances in Computing Research, S. Micali, Ed. JAI Press, 1989, vol. 5, pp. 73–90.

[12] A. Kitaev and J. Watrous, "Parallelization, amplification, and exponential time simulation of quantum interactive proof system," in *STOC*, 2000, pp. 608–617.

[13] C. Marriott and J. Watrous, "Quantum Arthur-Merlin games," *Computational Complexity*, vol. 14, no. 2, pp. 122–152, 2005, arXiv:cs/0506068v1 [cs.CC].

[14] S. Arora, E. Hazan, and S. Kale, "The multiplicative weights update method: a meta algorithm and applications," 2005, submitted.

[15] M. Warmuth and D. Kuzmin, "Online variance minimization," in *COLT*, 2006, pp. 514–528.

[16] S. Kale, "Efficient algorithms using the multiplicative weights update method," Ph.D. dissertation, Princeton University, 2007.

[17] R. Jain and J. Watrous, "Parallel approximation of non-interactive zero-sum quantum games," in *CCC*, 2009, pp. 243–253, arXiv:0808.2775v1 [quant-ph].

[18] R. Jain, S. Upadhyay, and J. Watrous, "Two-message quantum interactive proofs are in PSPACE," in *FOCS*, 2009, pp. 534–543, arXiv:0905.1300v1 [quant-ph].

[19] A. Kitaev, "Quantum coin-flipping," Presentation at QIP Workshop, 2002.

[20] C. Fuchs and J. van de Graaf, "Cryptographic distinguishability measures for quantum mechanical states," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1216–1227, 1999, arXiv:quant-ph/9712042v2.

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

[22] J. Watrous, "Lecture notes: Theory of quantum information," 2011, available on the author's web page.

[23] J. von zur Gathen, "Parallel linear algebra," in *Synthesis of Parallel Algorithms*, J. H. Reif, Ed. Morgan Kaufmann Publishers, Inc., 1993, ch. 13.

[24] G. Gutoski and J. Watrous, "Toward a general theory of quantum games," in *STOC*, 2007, pp. 565–574, arXiv:quant-ph/0611234v2.

[25] J. von Neumann, "Zur theorie der gesellschaftspiele," *Mathematische Annalen*, vol. 100, no. 1, pp. 295–320, 1928.

[26] K. Fan, "Minimax theorems," *Proceedings of the National Academy of Sciences*, vol. 39, pp. 42–47, 1953.

[27] G. Gutoski, "Upper bounds for quantum interactive proofs with competing provers," in *CCC*, 2005, pp. 334–343.