On a measure of distance for quantum strategies

Gus Gutoski

Institute for Quantum Computing and School of Computer Science
University of Waterloo, Waterloo, Ontario, Canada

October 6, 2011
(Revised: February 17, 2012)

Abstract

The present paper studies an operator norm that captures the distinguishability of quantum strategies in the same sense that the trace norm captures the distinguishability of quantum states or the diamond norm captures the distinguishability of quantum channels. Characterizations of its unit ball and dual norm are established via strong duality of a semidefinite optimization problem. A full, formal proof of strong duality is presented for the semidefinite optimization problem in question. This norm and its properties are employed to generalize a state discrimination result of Ref. [GW05]. The generalized result states that for any two convex sets $\mathbf{S}_0, \mathbf{S}_1$ of strategies there exists a fixed interactive measurement scheme that successfully distinguishes any choice of $S_0 \in \mathbf{S}_0$ from any choice of $S_1 \in \mathbf{S}_1$ with bias proportional to the minimal distance between the sets $\mathbf{S}_0$ and $\mathbf{S}_1$ as measured by this norm. A similar discrimination result for channels then follows as a special case.

1 Introduction

1.1 Quantum strategies

A *quantum strategy* is a complete specification of the actions of one party in an interaction involving the exchange of multiple rounds of quantum messages with one or more other parties. Fundamental objects in the study of quantum information such as states, measurements, and channels may be viewed as special cases of strategies. A particularly useful representation for quantum strategies is presented in Ref. [GW07]. (See also Ref. [CDP09b].)

Briefly and informally, this representation associates with each strategy a single positive semidefinite operator S, the dimensions of which depend upon the size of the messages exchanged in the interaction. It is shown in Ref. [GW07] that the set of all positive semidefinite operators which are valid representations of strategies is characterized by a simple and efficiently-verifiable collection of linear equality conditions. (Essentially, these conditions reflect the intuitive causality constraint that outgoing messages in early rounds of the interaction cannot depend upon incoming messages from later rounds.) An explicit list of these conditions is given in Section 2.

In order to extract useful classical information from an interaction, a strategy might call for one or more quantum measurements throughout the interaction. In this case, the strategy is instead represented by a set $\{S_a\}$ of positive semidefinite operators indexed by all the possible combinations of outcomes of the measurements. These strategies are called *measuring strategies* and satisfy $\sum_a S_a = S$ for some ordinary

(non-measuring) strategy S. (By comparison, an ordinary POVM-type quantum measurement $\{P_a\}$ satisfies $\sum_a P_a = I$.)

Conveniently, the relationship between measuring and non-measuring strategies is analogous to that between ordinary measurements and states. In particular, the postulates of quantum mechanics dictate that for any ordinary measurement $\{P_a\}$ with outcomes indexed by a and any density operator ρ it holds that the probability with which $\{P_a\}$ yields outcome a when applied to a quantum system whose state is represented by ρ is given by the inner product

$$\Pr[\{P_a\} \text{ yields outcome } a \text{ on } \rho] = \langle P_a, \rho \rangle = \mathrm{Tr}(P_a \rho).$$

Similarly, it is shown in Ref. [GW07] that the probability with which a measuring strategy $\{S_a\}$ yields outcome a after an interaction with a compatible quantum strategy T is given by

$$\Pr[\{S_a\} \text{ yields outcome } a \text{ when interacting with } T] = \langle S_a, T \rangle = \mathrm{Tr}(S_a T).$$

A more formal review of quantum strategies is given in Section 2.

1.2 Distance measures

In the study of quantum information the need often arises for a distance measure that quantifies the observable difference between two states or channels. For states, such a distance measure is induced by the trace norm of the difference between two density operators. For channels, the measure of choice is induced by the diamond norm of the difference between two completely positive and trace-preserving linear maps.

The use of the trace norm to measure distance between states can be traced back to the 1960s (see Nielsen and Chuang [NC00, Chapter 9]). The diamond norm was defined by Kitaev for the explicit purpose of measuring distance between channels [Kit97, AKN98]. It was later noticed that the diamond norm is related via the notion of duality to the *norm of complete boundedness* for linear maps, an object of study in mathematics circles since the 1980s. (See Paulsen [Pau02].)

A suitable distance measure for quantum strategies was first considered by Chiribella, D'Ariano, and Perinotti [CDP08]. This distance measure captures the distinguishability of quantum strategies in the same sense that the trace norm captures the distinguishability of states or the diamond norm captures the distinguishability of channels. Given the strikingly similar relationships between states and measurements and between strategies and measuring strategies, the new distance measure suggests itself: whereas the trace norm $\|\rho - \sigma\|_{\mathrm{Tr}}$ for quantum states ρ, σ is easily seen to satisfy

$$\|\rho - \sigma\|_{\mathrm{Tr}} = \max\left\{\langle P_0 - P_1, \rho - \sigma \rangle : \{P_0, P_1\} \text{ is a quantum measurement}\right\},$$

the *strategy r-norm* $\|R - S\|_{\diamond r}$ for quantum strategies R, S can be informally defined by

$$\|R - S\|_{\diamond r} \stackrel{\text{def}}{=} \max\left\{\langle T_0 - T_1, R - S \rangle : \{T_0, T_1\} \text{ is a compatible measuring strategy}\right\}.$$

(A formal definition of this norm appears in Section 3 after due discussion of preliminary material.)

Here the subscript r denotes the number of rounds of messages in the protocol for which R, S are strategies. In particular, each positive integer r induces a different strategy norm. The choice of notation is inspired by the fact that this norm coincides with the diamond norm for the case $r = 1$ [CDP08]. Hence, the strategy r-norm can be viewed as a generalization of the diamond norm for Hermitian-preserving linear maps.

Little else is known of the strategy r-norm. It was noted in Ref. [CDP08] that this norm differs from the diamond norm for $r > 1$. The norm was also mentioned in Refs. [CDP09a, CDP09b] and Chiribella *et al.* proved a continuity bound for the strategy r-norm as part of their short impossibility proof for quantum bit commitment schemes [CDP$^+$09c].

Recent work in the mathematical physics literature has focussed on an extension of the diamond norm (or, equivalently, the norm of complete boundedness) to k-minimal and k-maximal operator spaces and operator systems and the relationships of these norms and spaces with entangled quantum states and with the k-positive and k-superpositive cones of operators—see Refs. [JKPP11, SSŻ09] and the references therein. We briefly elaborate upon this extension of the diamond norm at the end of Section 3.2. However, all appearances indicate that the these objects have little to do with the strategy r-norm or the cone generated by r-round strategies.

1.3 Results

Characterizations of the unit balls of the strategy r-norm and its dual norm are presented as Theorem 3 in Section 4. This theorem is proven via strong duality of a semidefinite optimization problem. A full, formal proof of strong duality for this problem is given in Appendix A.

These characterizations are then used to generalize a state discrimination result of Ref. [GW05], which asserts that for any two convex sets $\mathbf{A}_0, \mathbf{A}_1$ of states there exists a fixed measurement that successfully distinguishes any choice of $\rho_0 \in \mathbf{A}_0$ from any choice of $\rho_1 \in \mathbf{A}_1$ with bias proportional to the minimal trace norm distance between the sets $\mathbf{A}_0$ and $\mathbf{A}_1$.

By analogy, Theorem 5 in Section 5 asserts that for any two convex sets $\mathbf{S}_0, \mathbf{S}_1$ of r-round strategies there exists a fixed compatible r-round measuring strategy that successfully distinguishes any choice of $S_0 \in \mathbf{S}_0$ from any choice of $S_1 \in \mathbf{S}_1$ with bias proportional to the minimum distance between the sets $\mathbf{S}_0$ and $\mathbf{S}_1$ as measured by the strategy r-norm. Just as in Ref. [GW05], it therefore follows that

1. This compatible measuring strategy can be used to discriminate between *any* choices of strategies from $\mathbf{S}_0, \mathbf{S}_1$ at least as well as *any other* compatible measuring strategy could discriminate between the two *closest* strategies from those sets.

2. Even if two (or more) distinct pairs (S_0, S_1) and (S_0', S_1') both minimize the distance between $\mathbf{S}_0$ and $\mathbf{S}_1$ then *both* pairs may be *optimally* discriminated by the *same compatible measuring strategy*.

As a special case of Theorem 5, a similar discrimination result is obtained for convex sets of channels with the diamond norm in place of the strategy r-norm.

Strong duality of the aforementioned semidefinite optimization problem also yields an alternate and arguably simpler proof of a property of strategies established in Ref. [GW07]. This property, listed as Theorem 4 in Section 4.1, establishes a useful formula for the maximum probability with which a measuring strategy can be forced to produce a given measurement outcome by a compatible interacting strategy.

1.4 Notation

The following table summarizes the notation used in this paper.

$\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$	Calligraphic letters denote finite-dimensional complex Euclidean spaces of the form $\mathbb{C}^n$.
$\mathcal{X}_{1\ldots n}$	Shorthand notation for the tensor product $\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n$.
$\mathbf{S}, \mathbf{T}, \mathbf{A}, \mathbf{B}$	Bold letters denote sets of operators.
$\mathbf{L}(\mathcal{X})$	The (complex) space of all linear operators $A : \mathcal{X} \to \mathcal{X}$, implicitly identified with $\mathbb{C}^{n \times n}$.
$\mathbf{Her}(\mathcal{X})$	The (real) subspace of Hermitian operators within $\mathbf{L}(\mathcal{X})$.
$\mathbf{Pos}(\mathcal{X})$	The cone of positive semidefinite operators within $\mathbf{Her}(\mathcal{X})$.
$\succeq, \succ, \preceq, \prec$	The semidefinite partial ordering on $\mathbf{Her}(\mathcal{X})$.
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 two operators $A, B : \mathcal{X} \to \mathcal{Y}$. Defined by $\langle A, B \rangle \stackrel{\text{def}}{=} \mathrm{Tr}(A^* B)$.
$I_{\mathcal{X}}$	The identity operator acting on $\mathcal{X}$.
$\mathbb{1}_{\mathcal{X}}$	The identity linear map acting on $\mathbf{L}(\mathcal{X})$.
$\mathrm{Tr}_{\mathcal{X}}$	The partial trace over $\mathcal{X}$. For any space $\mathcal{Y}$ this linear map is defined by $$\mathrm{Tr}_{\mathcal{X}} : \mathbf{L}(\mathcal{X} \otimes \mathcal{Y}) \to \mathbf{L}(\mathcal{Y}) : X \otimes Y \mapsto \mathrm{Tr}(X) Y.$$ (This definition extends to all of $\mathbf{L}(\mathcal{X} \otimes \mathcal{Y})$ by linearity on operators of the form $X \otimes Y$.)
$J(\Phi)$	The Choi-Jamiołkowski operator representation of a linear map Φ. (See below.)

A *density operator* or *quantum state* is a positive semidefinite operator with unit trace. A *quantum measurement* with (finitely many) outcomes indexed by a is a finite set $\{P_a\} \subset \mathbf{Pos}(\mathcal{X})$ of positive semidefinite operators with $\sum_a P_a = I_{\mathcal{X}}$.

A linear map Φ is *positive* if $\Phi(X) \succeq 0$ whenever $X \succeq 0$ and *completely positive* if $\Phi \otimes \mathbb{1}_{\mathcal{W}}$ is positive for all choices of the space $\mathcal{W}$. A linear map Φ is *trace-preserving* if $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$ for all X. As usual, the set of all possible physically realizable operations on quantum states is identified with the set of completely positive and trace-preserving linear maps. Such a map is often called a *channel*.

The *Choi-Jamiołkowski isomorphism* associates with each linear map $\Phi : \mathbf{L}(\mathcal{X}) \to \mathbf{L}(\mathcal{Y})$ a unique operator $J(\Phi) \in \mathbf{L}(\mathcal{Y} \otimes \mathcal{X})$ via the formula

$$J(\Phi) = \sum_{i,j=1}^{\dim(\mathcal{X})} \Phi(E_{i,j}) \otimes E_{i,j}$$

where $\{E_{i,j}\}$ is the standard orthonormal basis for $\mathbf{L}(\mathcal{X})$. It holds that Φ is completely positive if and only if $J(\Phi)$ is positive semidefinite and that that $\Phi : \mathbf{L}(\mathcal{X}) \to \mathbf{L}(\mathcal{Y})$ is trace-preserving if and only if $\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = I_{\mathcal{X}}$. A linear map Φ is *Hermitian-preserving* if $\Phi(X)$ is Hermitian whenever X is Hermitian. It holds that Φ is Hermitian-preserving if and only if $J(\Phi)$ is a Hermitian operator.

1.5 Table of contents

The rest of this paper is organized as follows.

2 Review of quantum strategies

This section reviews the formalism of quantum strategies as presented in Ref. [GW07]. The curious reader is referred to Refs. [GW07, CDP09b] for additional detail.

2.1 Operational formalism

At a high level, a *strategy* is a complete description of one party's actions in a multiple-round interaction involving the exchange of quantum information with one or more other parties. For convenience, let us call this party *Alice*. As we are only concerned for the moment with Alice's actions during the interaction, it is convenient to bundle the remaining parties into one party, whom we call *Bob*.

From Alice's point of view every finite interaction decomposes naturally into a finite number r of *rounds*. In a typical round a message comes in, the message is processed, and a reply is sent out. Naturally, this reply might depend upon messages exchanged during previous rounds of the interaction. To account for such a dependence, we allow for a memory workspace to be maintained between rounds.

The complex Euclidean spaces corresponding to the incoming and outgoing messages in an arbitrary round i shall be denoted $\mathcal{X}_i$ and $\mathcal{Y}_i$, respectively. The space corresponding to the memory workspace to be stored for the next round shall be denoted $\mathcal{Z}_i$. In a typical round i of the quantum interaction, Alice's actions are faithfully represented by a channel

$$\Phi_i : \mathbf{L}(\mathcal{X}_i \otimes \mathcal{Z}_{i-1}) \to \mathbf{L}(\mathcal{Y}_i \otimes \mathcal{Z}_i).$$

The first round of the interaction is a special case: there is no need for an incoming memory space for this round, so the channel Φ_1 has the form

$$\Phi_1 : \mathbf{L}(\mathcal{X}_1) \to \mathbf{L}(\mathcal{Y}_1 \otimes \mathcal{Z}_1).$$

The final round of the interaction is also a special case: there is no immediate need for an outgoing memory space for this round. However, the presence of this final memory space better facilitates the forthcoming discussion of strategies involving measurements. Thus, the channel Φ_r representing Alice's actions in the final round of the interaction has the same form as those from previous rounds:

$$\Phi_r : \mathbf{L}(\mathcal{X}_r \otimes \mathcal{Z}_{r-1}) \to \mathbf{L}(\mathcal{Y}_r \otimes \mathcal{Z}_r).$$

In order to extract classical information from the interaction it suffices to permit Alice to perform a single quantum measurement on her final memory workspace. (Sufficiency of a single measurement at the end of the interaction follows immediately from foundational results on mixed state quantum computations [AKN98], which tell us that any process calling for one or more intermediate measurements can be efficiently simulated by a channel with a single measurement at the end.)

Formally then, the *operational description* of an r*-round strategy* for an interaction with *input spaces* $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and *output spaces* $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ is specified by:

1. Complex Euclidean spaces $\mathcal{Z}_1, \ldots, \mathcal{Z}_r$, called *memory spaces*, and
2. An r-tuple of channels $(\Phi_1, \ldots, \Phi_r)$ of the form

$$\begin{aligned} \Phi_1 &: \mathbf{L}(\mathcal{X}_1) \to \mathbf{L}(\mathcal{Y}_1 \otimes \mathcal{Z}_1) \\ \Phi_i &: \mathbf{L}(\mathcal{X}_i \otimes \mathcal{Z}_{i-1}) \to \mathbf{L}(\mathcal{Y}_i \otimes \mathcal{Z}_i) \quad (2 \leq i \leq r). \end{aligned}$$

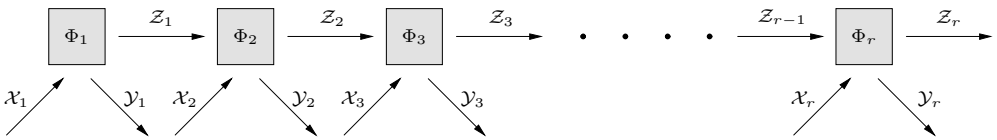

Figure 1: An r-round strategy.

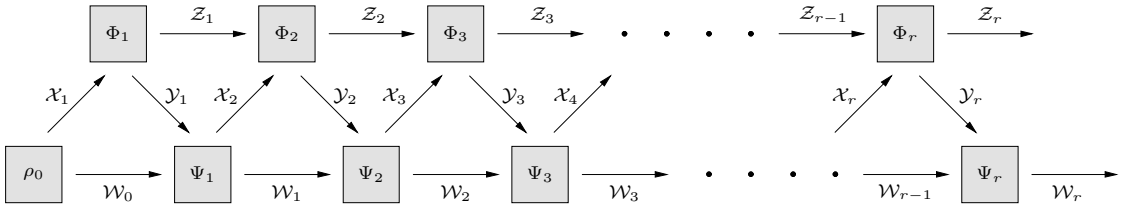

Figure 2: An interaction between an r-round strategy and co-strategy.

The operational description of an r-round *measuring* strategy with outcomes indexed by a is specified by items 1 and 2 above, as well as:

3. A measurement $\{P_a\} \subset \mathbf{Pos}(\mathcal{Z}_r)$ on the last memory space $\mathcal{Z}_r$.

We use the words "operational description" to distinguish this representation for strategies from the representation to be described in Section 2.2.

A strategy without a measurement is referred to a *non-measuring* strategy. A non-measuring strategy may be viewed as a measuring strategy in which the measurement has only one outcome, so that $\{P_a\} = \{I\}$ is the singleton set containing the identity. Figure 1 illustrates an r-round non-measuring strategy.

Note that input and output spaces may have dimension one, which corresponds to an empty message. One can therefore view simple actions such as the preparation of a quantum state or performing a measurement without producing a quantum output as special cases of strategies. (Special cases such as this are also discussed at the end of Section 2.3.)

In order for interaction to occur Bob must supply the incoming messages $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and process the outgoing messages $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ as suggested by Figure 2. Due to the inherently asymmetric nature of any interaction (only one of the parties can send the first message or receive the final message), the actions of Bob are described not by a *strategy*, but by a slightly different object called a *co-strategy*.

Formally, the operational description of an r-round *co-strategy* for an interaction with input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and output spaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ is specified by:

1. Complex Euclidean memory spaces $\mathcal{W}_0, \ldots, \mathcal{W}_r$,
2. A quantum state $\rho_0 \in \mathbf{Pos}(\mathcal{X}_1 \otimes \mathcal{W}_0)$, and
3. An r-tuple of channels $(\Psi_1, \ldots, \Psi_r)$ of the form

$$\Psi_i : \mathbf{L}(\mathcal{Y}_i \otimes \mathcal{W}_{i-1}) \to \mathbf{L}(\mathcal{X}_{i+1} \otimes \mathcal{W}_i) \quad (1 \leq i \leq r-1)$$
$$\Psi_r : \mathbf{L}(\mathcal{Y}_r \otimes \mathcal{W}_{r-1}) \to \mathbf{L}(\mathcal{W}_r).$$

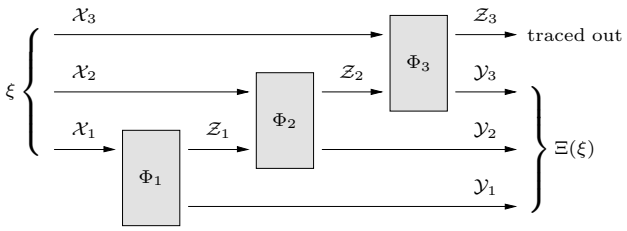

Figure 3: The linear map Ξ associated with a three-round strategy.

The operational description of an r-round *measuring* co-strategy with outcomes indexed by b is specified by items 1, 2 and 3 above, as well as:

4. A measurement $\{Q_b\} \subset \mathbf{Pos}(\mathcal{W}_r)$ on the last memory space $\mathcal{W}_r$.

The *output* of an interaction between a strategy and a co-strategy is the result of the measurements performed after the interaction. In particular, the postulates of quantum mechanics tell us that the probability with which Alice and Bob output the pair (a, b) is given by

$$\Pr[\text{output } (a,b)] = \mathrm{Tr}\left((P_a \otimes Q_b)\,\sigma_r\right)$$

where $\sigma_r \in \mathbf{Pos}(\mathcal{Z}_r \otimes \mathcal{W}_r)$ is the state of the system at the end of the interaction. (This state is most conveniently described by the recursive formula $\sigma_{i+1} = (\mathbb{1}_{\mathcal{Z}_i} \otimes \Psi_i) \circ (\Phi_i \otimes \mathbb{1}_{\mathcal{W}_{i-1}})(\sigma_i)$ with $\sigma_0 = \rho_0$.)

2.2 Choi-Jamiołkowski formalism

While intuitive from an operational perspective, the operational description of a strategy by an r-tuple of channels and a measurement is often inconvenient. In this subsection we describe the alternate formalism for strategies presented in Ref. [GW07] derived from the Choi-Jamiołkowski representation for linear maps.

Let us first restrict attention to r-round non-measuring strategies. To the r-round strategy specified by channels $(\Phi_1, \ldots, \Phi_r)$ we associate a single channel

$$\Xi : \mathbf{L}(\mathcal{X}_{1\ldots r}) \to \mathbf{L}(\mathcal{Y}_{1\ldots r}).$$

This channel takes a given r-partite input state $\xi \in \mathbf{Pos}(\mathcal{X}_{1\ldots r})$ and feeds the portions of this state corresponding to the input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$ into the network pictured in Figure 1, one piece at a time. The final memory space $\mathcal{Z}_r$ is then traced out, leaving some element $\Xi(\xi) \in \mathbf{Pos}(\mathcal{Y}_{1\ldots r})$. Such a map is depicted in Figure 3 for the case $r = 3$. An *r-round non-measuring strategy* for input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and output spaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ is defined to be Choi-Jamiołkowski representation

$$J(\Xi) \in \mathbf{Pos}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r})$$

of the channel Ξ we have just described. (This definition of a strategy is distinguished from the operational description of Section 2.1 by the absence of the words "operational description.")

To a measuring strategy with measurement $\{P_a\} \subset \mathbf{Pos}(\mathcal{Z}_r)$ we associate not a single channel, but instead a set $\{\Xi_a\}$ of linear maps, one for each measurement outcome a, each of the same form

$$\Xi_a : \mathbf{L}(\mathcal{X}_{1\ldots r}) \to \mathbf{L}(\mathcal{Y}_{1\ldots r}).$$

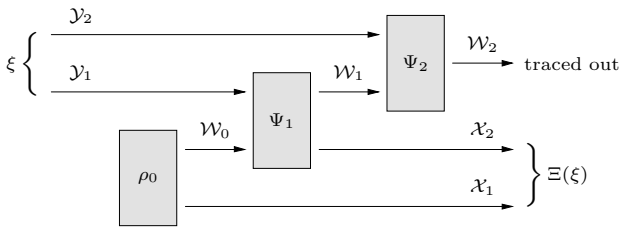

Figure 4: The linear map Ξ associated with a two-round co-strategy.

Each Ξ_a is defined precisely as in the non-measuring case except that the partial trace over $\mathcal{Z}_r$ is replaced by the mapping

$$X \mapsto \mathrm{Tr}_{\mathcal{Z}_r}((P_a \otimes I_{\mathcal{Y}_{1\ldots r}})X).$$

Each of the linear maps Ξ_a is completely positive and trace non-increasing, but not necessarily trace-preserving. Notice that

$$\sum_a \Xi_a = \Xi$$

where Ξ is the channel defined as in the non-measuring case. This observation is consistent with the view that Ξ represents a measuring strategy with only one outcome.

Non-measuring co-strategies are defined similarly to non-measuring strategies except that we take the Choi-Jamiołkowski representation $J(\Xi^*)$ of the adjoint linear mapping Ξ^* of the channel Ξ described above. So, for example, an r-round non-measuring co-strategy specified by $(\rho_0, \Psi_1, \ldots, \Psi_r)$ induces a channel

$$\Xi : \mathbf{L}(\mathcal{Y}_{1\ldots r}) \to \mathbf{L}(\mathcal{X}_{1\ldots r})$$

as suggested by Figure 4. Notice that the domain $\mathbf{L}(\mathcal{Y}_{1\ldots r})$ and range $\mathbf{L}(\mathcal{X}_{1\ldots r})$ are switched when the mapping Ξ is derived from a co-strategy instead of a strategy. The domain and range are switched back again by working with the adjoint mapping Ξ^*. One implication of this choice to work with the adjoint mapping for co-strategies is that the Choi-Jamiołkowski representations for both strategies and co-strategies are always elements of $\mathbf{Pos}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r})$. (Otherwise, co-strategies would lie in $\mathbf{Pos}(\mathcal{X}_{1\ldots r} \otimes \mathcal{Y}_{1\ldots r})$.)

The extension from non-measuring co-strategies to measuring co-strategies is completely analogous to that for strategies.

2.3 Properties of strategies

This subsection lists several useful properties of strategies, each of which was first established in Ref. [GW07].

The first such property is that the set of all linear maps that represent legal non-measuring strategies is conveniently characterized by a collection of linear constraints on the Choi-Jamiołkowski matrix. Specifically, an arbitrary operator $S \in \mathbf{L}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r})$ is the representation of some r-round non-measuring strategy for input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and output spaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ if and only if S is positive semidefinite and there exist positive semidefinite operators $S_{[1]}, \ldots, S_{[r]}$ of the form

$$S_{[i]} \in \mathbf{Pos}(\mathcal{Y}_{1\ldots i} \otimes \mathcal{X}_{1\ldots i}) \qquad (1 \leq i \leq r)$$

such that $S = S_{[r]}$ and

$$\begin{aligned}
\mathrm{Tr}_{\mathcal{Y}_r}\left(S_{[r]}\right) &= S_{[r-1]} \otimes I_{\mathcal{X}_r} \\
&\vdots \\
\mathrm{Tr}_{\mathcal{Y}_2}\left(S_{[2]}\right) &= S_{[1]} \otimes I_{\mathcal{X}_2} \\
\mathrm{Tr}_{\mathcal{Y}_1}\left(S_{[1]}\right) &= I_{\mathcal{X}_1}.
\end{aligned}$$

In other words, there exist memory spaces $\mathcal{Z}_1, \ldots, \mathcal{Z}_r$ and channels $(\Phi_1, \ldots, \Phi_r)$ such that the channel Ξ induced by these objects as described in Section 2.2 satisfies $J(\Xi) = S$ if and only if S meets the above criteria.

Similarly, an operator $T \in \mathbf{L}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r})$ is the representation of some r-round non-measuring co-strategy for input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and output spaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ if and only if T is positive semidefinite and there exist positive semidefinite operators $T_{[1]}, \ldots, T_{[r]}$ of the form

$$T_{[i]} \in \mathbf{Pos}(\mathcal{Y}_{1\ldots i-1} \otimes \mathcal{X}_{1\ldots i}) \qquad\qquad (1 \leq i \leq r)$$

such that

$$\begin{aligned}
T &= T_{[r]} \otimes I_{\mathcal{Y}_r} \\
\mathrm{Tr}_{\mathcal{X}_r}\left(T_{[r]}\right) &= T_{[r-1]} \otimes I_{\mathcal{Y}_{r-1}} \\
&\vdots \\
\mathrm{Tr}_{\mathcal{X}_2}\left(T_{[2]}\right) &= T_{[1]} \otimes I_{\mathcal{Y}_1} \\
\mathrm{Tr}\left(T_{[1]}\right) &= 1.
\end{aligned}$$

Measuring strategies also admit a simple characterization: a set $\{S_a\} \subset \mathbf{Pos}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r})$ is the representation of some r-round measuring strategy for input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and output spaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ if and only if $\sum_a S_a$ is the representation of some r-round non-measuring strategy for the same input and output spaces. A similar characterization holds for measuring co-strategies.

When an r-round measuring strategy $\{S_a\}$ for input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and output spaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ interacts with an r-round measuring co-strategy $\{T_b\}$ for the same input and output spaces the probability with which the output pair (a, b) occurs is given by the inner product

$$\Pr[\text{interaction between } \{S_a\} \text{ and } \{T_b\} \text{ yields output } (a,b)] = \langle S_a, T_b \rangle = \mathrm{Tr}(S_a T_b).$$

The standard inner product relationship between ordinary states and measurements is recovered in the special case $r = 1$ and $\dim(\mathcal{Y}_1) = 1$. To see this, notice that the set of all non-measuring co-strategies coincides in this case with the set of all density operators on $\mathcal{X}_1$. Any measuring strategy $\{S_a\} \subset \mathbf{Pos}(\mathcal{X}_1)$ satisfies $\sum_a S_a = I_{\mathcal{X}_1}$ and hence acts as an ordinary measurement on $\mathcal{X}_1$. The previous inner product formula therefore tells us

$$\Pr[\{S_a\} \text{ yields output } a \text{ when applied to } \rho] = \langle S_a, \rho \rangle = \mathrm{Tr}(S_a \rho),$$

which is the familiar postulate of quantum mechanics.

3 Discrimination problems and norms

A formal definition of the strategy r-norm is given in Definition 1 in Section 3.3 after due discussions of the trace and diamond norms in Sections 3.1 and 3.2, respectively. A discrimination problem for convex sets of states, channels, and strategies is discussed in Section 3.4.

3.1 The trace norm as a distance measure for states

The *trace norm* $\|X\|_{\mathrm{Tr}}$ of an arbitrary operator X is defined as the sum of the singular values of X. If X is Hermitian then it is a simple exercise to verify that its trace norm is given by

$$\begin{aligned}\|X\|_{\mathrm{Tr}} &= \max\left\{\langle P_0 - P_1, X\rangle \;:\; P_0, P_1 \succeq 0,\; P_0 + P_1 = I\right\} \\ &= \max\left\{\langle P_0 - P_1, X\rangle \;:\; \{P_0, P_1\} \text{ is a two-outcome measurement}\right\}.\end{aligned}$$

The trace norm provides a physically meaningful distance measure for quantum states in the sense that it captures the maximum likelihood with which two states can be correctly discriminated. This fact is illustrated by a simple example involving two parties called *Alice* and *Bob* and a fixed pair of quantum states ρ_0, ρ_1. Suppose Bob selects a state $\rho \in \{\rho_0, \rho_1\}$ uniformly at random and gives Alice a quantum system prepared in state ρ. Alice has a complete description of both ρ_0 and ρ_1, but she does not know which of the two was selected by Bob. Her goal is to correctly guess which of $\{\rho_0, \rho_1\}$ was selected based upon the outcome of a measurement she conducts on ρ.

Since Alice's guess is binary-valued and completely determined by her measurement, that measurement can be assumed to be a two-outcome measurement $\{P_0, P_1\}$ wherein outcome $a \in \{0, 1\}$ indicates a guess that Bob prepared $\rho = \rho_a$. The probability with which Alice successfully discriminates ρ_0 from ρ_1 is easily shown to be

$$\Pr[\text{Alice guesses correctly}] = \frac{1}{2} + \frac{1}{4}\langle P_0 - P_1, \rho_0 - \rho_1\rangle \leq \frac{1}{2} + \frac{1}{4}\|\rho_0 - \rho_1\|_{\mathrm{Tr}}$$

with equality achieved at the optimal measurement $\{P_0, P_1\}$ for Alice. This fundamental observation was originally made by Helstrom [Hel69].

3.2 The diamond norm as a distance measure for channels

The *linear map trace norm* is induced by the operator trace norm via the formula

$$\|\Phi\|_{\mathrm{Tr}} \stackrel{\text{def}}{=} \max_{\|X\|_{\mathrm{Tr}}=1} \|\Phi(X)\|_{\mathrm{Tr}}.$$

Unfortunately, this norm does not lead to an overly useful distance measure for quantum channels. To achieve such a measure, the trace norm must be "stabilized" to form the *diamond norm* via the formula

$$\|\Phi\|_{\diamond} \stackrel{\text{def}}{=} \sup_{\mathcal{W}} \|\Phi \otimes \mathbb{1}_{\mathcal{W}}\|_{\mathrm{Tr}}$$

where the supremum is taken over all finite-dimensional complex Euclidean spaces $\mathcal{W}$.

Much is known of the diamond norm. For example, if Φ has the form $\Phi : \mathbf{L}(\mathcal{X}) \to \mathbf{L}(\mathcal{Y})$ then the supremum in the definition of $\|\Phi\|_{\diamond}$ is always achieved by some space $\mathcal{W}$ whose dimension does not exceed that of the input space $\mathcal{X}$. (This fact was originally established for the completetly bounded norm by Smith

[Smi83] and independently rediscovered for the diamond norm by Kitaev [Kit97, AKN98].) As a consequence, the supremum in the definition of the diamond norm can be replaced by a maximum. Moreover, if Φ is Hermitian-preserving and $\dim(\mathcal{W}) \geq \dim(\mathcal{X})$ then the maximum in the definition of $\|\Phi \otimes \mathbb{1}_{\mathcal{W}}\|_{\mathrm{Tr}}$ is always achieved by some positive semidefinite operator X [RW05, Wat05, GLN05]. Thus, if Φ is Hermitian-preserving then its diamond norm is given by

$$\begin{aligned} \|\Phi\|_{\diamond} &= \max \|(\Phi \otimes \mathbb{1}_{\mathcal{W}})(\rho)\|_{\mathrm{Tr}} \\ &= \max \langle P_0 - P_1, (\Phi \otimes \mathbb{1}_{\mathcal{W}})(\rho) \rangle \end{aligned}$$

where the maxima in these two expressions are taken over all spaces $\mathcal{W}$ with dimension at most $\mathcal{X}$, all states $\rho \in \mathbf{Pos}(\mathcal{X} \otimes \mathcal{W})$, and all two-outcome measurements $\{P_0, P_1\} \subset \mathbf{Pos}(\mathcal{Y} \otimes \mathcal{W})$.

The diamond norm is to channels as the trace norm is to states: it provides a physically meaningful distance measure for channels in the sense that the value $\|\Phi_0 - \Phi_1\|_{\diamond}$ quantifies the observable difference between two channels Φ_0, Φ_1. As before, this fact may be illustrated with a simple example. Suppose Bob selects a channel from $\Phi \in \{\Phi_0, \Phi_1\}$ uniformly at random. Alice is granted "one-shot, black-box" access to Φ and her goal is to correctly guess which of Φ_0, Φ_1 was applied. Specifically, Alice may prepare a quantum system in state ρ and send a portion of that system to Bob, who applies Φ to that portion and then returns it to Alice. Finally, Alice performs a two-outcome measurement $\{P_0, P_1\}$ on the resulting state $(\Phi \otimes \mathbb{1})(\rho)$ where outcome $a \in \{0, 1\}$ indicates a guess that $\Phi = \Phi_a$.

Repeating the derivation from Section 3.1, the probability with which Alice successfully discriminates Φ_0 from Φ_1 is seen to be

$$\Pr[\text{Alice guesses correctly}] = \frac{1}{2} + \frac{1}{4}\langle P_0 - P_1, (\Phi_0 \otimes \mathbb{1})(\rho) - (\Phi_1 \otimes \mathbb{1})(\rho) \rangle \leq \frac{1}{2} + \frac{1}{4}\|\Phi_0 - \Phi_1\|_{\diamond}$$

with equality achieved at the optimal input state ρ and measurement $\{P_0, P_1\}$ for Alice.

It is interesting to note that the ability to send only *part* of the input state ρ to Bob and keep the rest for herself can enhance Alice's ability to distinguish some pairs of channels, as compared to a simpler test that involves sending the *entire* input state to Bob. Indeed, there exist pairs Φ_0, Φ_1 of channels that are perfectly distinguishable when applied to half of a maximally entangled input state—that is, $\|\Phi_0 - \Phi_1\|_{\diamond} = 2$—yet they appear nearly identical when an auxiliary system is not used—that is, $\|\Phi_0 - \Phi_1\|_{\mathrm{Tr}} \approx 0$. An example of such a pair of linear maps can be found in Watrous [Wat08], along with much of the discussion that has occurred thus far in this section. It is this phenomenon that renders the linear map trace norm less useful than the diamond norm in the study of quantum information.

One might also consider an interpolation between $\|\Phi\|_{\mathrm{Tr}}$ and $\|\Phi\|_{\diamond}$ in which the dimension of the auxiliary space $\mathcal{W}$ is restricted to be at most k for some $1 \leq k \leq \dim(\mathcal{X})$. Johnston *et al.* studied the relationship between these norms and k-minimal operator spaces [JKPP11]. They also showed that for each k the same norm is achieved by replacing the restriction $\dim(\mathcal{W}) \leq k$ with the restriction that the Schmidt rank of the input state ρ be no larger than k. Timoney [Tim03] and Watrous [Wat08] studied conditions on Φ and k under which this norm is equal to the diamond norm. (In quantum information theoretic terms, their results esablish conditions under which an auxiliary space $\mathcal{W}$ of dimension k is sufficient for optimal distinguishability of two channels.)

3.3 The strategy r-norm as a distance measure for strategies

The simple guessing game played by Alice and Bob extends naturally from channels to strategies. Let S_0, S_1 be arbitrary r-round strategies and suppose Bob selects $S \in \{S_0, S_1\}$ uniformly at random. Alice's task is to interact with Bob and then decide after the interaction whether Bob selected $S = S_0$ or $S = S_1$.

Thanks to the inner product relationship between measuring strategies and co-strategies, much of discussion from Section 3.1 concerning the task of discriminating *states* can be re-applied to the task of discriminating *strategies*. In particular, Alice can be assumed to act according to some two-outcome r-round measuring co-strategy $\{T_0, T_1\}$ for Bob's input and output spaces, with outcome $a \in \{0, 1\}$ indicating a guess that Bob acted according to strategy S_a. As before, the probability with which Alice guesses correctly is given by

$$\Pr[\text{Alice guesses correctly}] = \frac{1}{2} + \frac{1}{4}\langle T_0 - T_1, S_0 - S_1 \rangle.$$

Naturally, Alice maximizes her chance of success by maximizing this expression over all r-round measuring co-strategies $\{T_0, T_1\}$.

Of course, this guessing game is symmetric with respect to strategies and co-strategies. In particular, if Bob's actions S_0, S_1 are *co-strategies* instead of strategies then Alice's actions $\{T_0, T_1\}$ must be a measuring *strategy* instead of a measuring co-strategy. Alice's maximum success probability is given by the same formula, except that Alice now maximizes this probability over all r-round measuring *strategies* $\{T_0, T_1\}$.

With this discrimination problem in mind, the distance measure of Ref. [CDP08] is recast in the present paper in the form of two norms—one that captures the distinguishability of strategies and one that captures the distinguishability of co-strategies.

Definition 1 (Strategy r-norm—see Ref. [CDP08])**.** For any Hermitian operator $X \in \mathbf{Her}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r})$ let

$$\begin{aligned}
\|X\|_{\diamond r} &\stackrel{\text{def}}{=} \max\left\{\langle T_0 - T_1, X\rangle : \{T_0, T_1\} \text{ is an } r\text{-round measuring co-strategy}\right\}, \\
\|X\|_{\diamond r}^* &\stackrel{\text{def}}{=} \max\left\{\langle S_0 - S_1, X\rangle : \{S_0, S_1\} \text{ is an } r\text{-round measuring strategy}\right\}.
\end{aligned}$$

These norms could also be viewed as linear map norms rather than operator norms. In this case, for any Hermitian-preserving linear map $\Phi : \mathbf{L}(\mathcal{X}_{1\ldots r}) \to \mathbf{L}(\mathcal{Y}_{1\ldots r})$ one may write

$$\begin{aligned}
\|\Phi\|_{\diamond r} &\stackrel{\text{def}}{=} \|J(\Phi)\|_{\diamond r}, \\
\|\Phi\|_{\diamond r}^* &\stackrel{\text{def}}{=} \|J(\Phi)\|_{\diamond r}^*.
\end{aligned}$$

The present paper leaves these norms undefined when X is not Hermitian, or, equivalently, when Φ is not Hermitian-preserving. □

It is not difficult to see that the functions $\|\cdot\|_{\diamond r}$ and $\|\cdot\|_{\diamond r}^*$ are norms.

Proposition 2. *The functions $\|\cdot\|_{\diamond r}$ and $\|\cdot\|_{\diamond r}^*$ from Definition 1 are norms.*

Proof. The defining properties of a norm can be verified directly. It follows immediately from Definition 1 that these functions obey the triangle inequality and that they are homogenous (meaning that $\|aX\|_{\diamond r} = |a|\|X\|_{\diamond r}$ for all $a \in \mathbb{R}$). To see that these functions are positive (meaning that $\|X\|_{\diamond r} \geq 0$ with equality only when $X = 0$), it suffices to establish the lower bounds

$$\begin{aligned}
\|X\|_{\diamond r} &\geq \frac{1}{\dim(\mathcal{X}_{1\ldots r})}\|X\|_{\mathrm{Tr}}, \\
\|X\|_{\diamond r}^* &\geq \frac{1}{\dim(\mathcal{Y}_{1\ldots r})}\|X\|_{\mathrm{Tr}}.
\end{aligned}$$

To this end, let Π_+, Π_- denote the projections (divided by $\dim(\mathcal{X}_{1\ldots r})$) onto the positive and nonpositive eigenspaces of X, respectively. Note that $\Pi_+ + \Pi_- = \frac{1}{\dim(\mathcal{X}_{1\ldots r})} I_{\mathcal{Y}_{1\ldots r}\otimes\mathcal{X}_{1\ldots r}}$, which is an r-round non-measuring co-strategy. Hence, $\{\Pi_+, \Pi_-\}$ is an r-round measuring co-strategy. We have

$$\|X\|_{\diamond r} \geq \langle \Pi_+ - \Pi_-, X\rangle = \frac{1}{\dim(\mathcal{X}_{1\ldots r})}\|X\|_{\mathrm{Tr}}$$

as desired. A similar argument for $\|X\|_{\diamond r}^*$ follows from the observation that $\frac{1}{\dim(\mathcal{Y}_{1\ldots r})} I_{\mathcal{Y}_{1\ldots r}\otimes\mathcal{X}_{1\ldots r}}$ is an r-round non-measuring strategy. □

If S_0, S_1 are strategies for input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$ and output spaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$ then it follows immediately from the discussion in Sections 3.1 and 3.2 that the maximum probability with which Alice can correctly distinguish S_0 from S_1 is

$$\frac{1}{2} + \frac{1}{4}\|S_0 - S_1\|_{\diamond r}.$$

Likewise, if S_0, S_1 are co-strategies rather than strategies then the maximum probability with which Alice can correctly distinguish S_0 from S_1 is

$$\frac{1}{2} + \frac{1}{4}\|S_0 - S_1\|_{\diamond r}^*.$$

It may seem superfluous to allow both strategies and co-strategies as descriptions for Bob's actions in this simple example, as every co-strategy may be written as a strategy via suitable relabelling of input and output spaces. But there is something to be gained by considering both the norms $\|\cdot\|_{\diamond r}$ and $\|\cdot\|_{\diamond r}^*$. Indeed, it is established by Theorem 3 that these norms are dual to each other.

3.4 Discrimination problems for convex sets of states, channels, and strategies

The guessing game played by Alice and Bob as discussed thus far in this section can be further generalized from a problem of discriminating *individual* states, channels, or strategies to discriminating *convex sets* of states, channels, or strategies.

Specifically, suppose two convex sets $\mathbf{A}_0, \mathbf{A}_1$ of states are fixed. Suppose that Bob arbitrarily selects $\rho_0 \in \mathbf{A}_0$ and $\rho_1 \in \mathbf{A}_1$ and then selects $\rho \in \{\rho_0, \rho_1\}$ uniformly at random and gives Alice a quantum system prepared in state ρ. Alice's goal is to correctly guess whether $\rho \in \mathbf{A}_0$ or $\rho \in \mathbf{A}_1$ based upon the outcome of a measurement she conducts on ρ. It is clear that this problem is a generalization of that from Section 3.1, as the original problem is recovered by considering singleton sets $\mathbf{A}_0 = \{\rho_0\}$ and $\mathbf{A}_1 = \{\rho_1\}$.

As mentioned in the introduction, this problem of discriminating convex sets of states was solved in Ref. [GW05] wherein it was shown that there exists a *single* measurement $\{P_0, P_1\}$ that depends only upon the sets $\mathbf{A}_0, \mathbf{A}_1$ with the property that *any* pair $\rho_0 \in \mathbf{A}_0$, $\rho_1 \in \mathbf{A}_1$ may be correctly discriminated with probability at least

$$\frac{1}{2} + \frac{1}{4}\min_{\sigma_a \in \mathbf{A}_a} \|\sigma_0 - \sigma_1\|_{\mathrm{Tr}}.$$

What can be said about this discrimination problem for convex sets of channels or strategies? Nothing was known of either problem prior to the work of the present paper. It is established by Theorem 5 that the discrimination result for convex sets of states extends unhindered to both channels and strategies. In particular, it is proven that two convex sets $\mathbf{S}_0, \mathbf{S}_1$ of r-round strategies can be correctly discriminated with probability at least

$$\frac{1}{2} + \frac{1}{4}\min_{S_a \in \mathbf{S}_a} \|S_0 - S_1\|_{\diamond r}.$$

It then follows trivially that two convex sets $\mathbf{T}_0, \mathbf{T}_1$ of r-round co-strategies can be correctly discriminated with probability at least

$$\frac{1}{2} + \frac{1}{4} \min_{T_a \in \mathbf{T}_a} \|T_0 - T_1\|_{\diamond r}^* .$$

As a special case, it holds that two convex sets $\mathbf{\Phi}_0, \mathbf{\Phi}_1$ of channels can be discriminated with probability at least

$$\frac{1}{2} + \frac{1}{4} \min_{\Phi_a \in \mathbf{\Phi}_a} \|\Phi_0 - \Phi_1\|_{\diamond} .$$

4 Unit ball of the strategy r-norm and its dual

By employing the characterization of r-round strategies mentioned in Section 2, the quantity $\|X\|_{\diamond r}$ can easily be written as a semidefinite optimization problem:

$$\begin{array}{ll} \text{maximize} & \langle X, T_0 - T_1 \rangle \\ \text{subject to} & T_0 + T_1 \text{ is an } r\text{-round non-measuring co-strategy} \\ & T_0, T_1 \succeq 0 \end{array} \tag{1}$$

In Appendix A it is shown that the dual optimization problem is given by

$$\begin{array}{ll} \text{minimize} & p \\ \text{subject to} & -pS \preceq X \preceq pS \\ & S \succeq 0 \text{ is an } r\text{-round non-measuring strategy} \end{array} \tag{2}$$

Moreover, it is also shown in Appendix A that *strong duality* holds for the optimization problems (1), (2), meaning that these problems have the same optimal value. Given that, we prove the following theorem.

Theorem 3 (Unit ball of the strategy r-norm and its dual)**.** *For every Hermitian operator $X \in \mathbf{Her}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r})$ it holds that*

1. $\|X\|_{\diamond r} \leq 1$ *if and only if* $X = S_0 - S_1$ *for some r-round measuring strategy* $\{S_0, S_1\}$.
2. $\|X\|_{\diamond r}^* \leq 1$ *if and only if* $X = T_0 - T_1$ *for some r-round measuring co-strategy* $\{T_0, T_1\}$.

Moreover, the norms $\|\cdot\|_{\diamond r}$ and $\|\cdot\|_{\diamond r}^*$ are dual to each other, meaning that

$$\|X\|_{\diamond r} = \max_{\|Y\|_{\diamond r}^* \leq 1} \langle Y, X \rangle,$$
$$\|X\|_{\diamond r}^* = \max_{\|Y\|_{\diamond r} \leq 1} \langle Y, X \rangle.$$

Proof. We begin with a proof of item 1. One direction is easy: if $X = S_0 - S_1$ for some r-round measuring strategy $\{S_0, S_1\}$ then for every r-round measuring co-strategy $\{T_0, T_1\}$ it holds that

$$\langle X, T_0 - T_1 \rangle = \langle S_0 - S_1, T_0 - T_1 \rangle \leq \langle S_0 + S_1, T_0 + T_1 \rangle = 1$$

and so $\|X\|_{\diamond r} \leq 1$.

For the other direction, suppose $\|X\|_{\diamond r} \leq 1$. By the strong duality of the optimization problems (1), (2) (see Appendix A) there exists an r-round non-measuring strategy S with $-S \preceq X \preceq S$. Let

$$S_0 = \frac{1}{2}(S + X), \qquad S_1 = \frac{1}{2}(S - X).$$

By construction it holds that $S_0 - S_1 = X$, that $S_0 + S_1 = S$, and that $S_0, S_1 \succeq 0$. The proof of item 1 is now complete.

That the norm $\|\cdot\|_{\diamond r}^*$ is dual to $\|\cdot\|_{\diamond r}$ now follows immediately:

$$\|X\|_{\diamond r}^* = \max\{\langle S_0 - S_1, X\rangle : \{S_0, S_1\} \text{ is an } r\text{-round measuring strategy}\} = \max_{\|Y\|_{\diamond r} \leq 1} \langle Y, X\rangle.$$

(The first equality is by definition and the second is item 1.)

The remaining claims of the theorem are symmetric to those already proved. One way to finish the proof would be to formulate a semidefinite optimization problem similar to (1) for $\|X\|_{\diamond r}^*$ and then derive its dual as in Appendix A. Alternately, the Duality Theorem (see Horn and Johnson [HJ85]) can be used to achieve a more direct proof.

To that end, note first that the Duality Theorem immediately implies that $\|\cdot\|_{\diamond r}$ is also dual to $\|\cdot\|_{\diamond r}^*$:

$$\|X\|_{\diamond r} = \max_{\|Y\|_{\diamond r}^* \leq 1} \langle Y, X\rangle.$$

To prove item 2, let $\mathbf{B}$ denote the set of all operators of the form $T_0 - T_1$ for some r-round measuring co-strategy $\{T_0, T_1\}$. We claim that $\mathbf{B}$ is the unit ball for some norm. This claim can be established by verifying that the set $\mathbf{B}$ is compact, convex, symmetric (meaning that $-B \in \mathbf{B}$ whenever $B \in \mathbf{B}$), and contains the origin in its interior [HJ85]. All but the last of these properties are immediate. To see that $\mathbf{B}$ contains the origin in its interior select any Hermitian operator X with $\|X\| \leq \frac{1}{\dim(\mathcal{X}_{1\ldots r})}$ and let $X = X_+ - X_-$ be an orthogonal decomposition of X. Write

$$D = \frac{1}{\dim(\mathcal{X}_{1\ldots r})} I - X_+ - X_-, \qquad T_0 = X_+ + \frac{1}{2}D, \qquad T_1 = X_- + \frac{1}{2}D.$$

Then $\{T_0, T_1\}$ is an r-round measuring co-strategy and $X = T_0 - T_1$ so $X \in \mathbf{B}$ and thus $\mathbf{B}$ contains the origin in its interior.

Let $\|\cdot\|_{\mathbf{B}}$ denote the unique norm whose unit ball is $\mathbf{B}$. We already know that

$$\|X\|_{\diamond r} = \max_{\|Y\|_{\mathbf{B}} \leq 1} \langle Y, X\rangle = \max_{\|Y\|_{\diamond r}^* \leq 1} \langle Y, X\rangle.$$

(The first equality is by definition and the second by duality of $\|\cdot\|_{\diamond r}$ and $\|\cdot\|_{\diamond r}^*$.) In particular, each of the norms $\|\cdot\|_{\diamond r}^*$ and $\|\cdot\|_{\mathbf{B}}$ has $\|\cdot\|_{\diamond r}$ as its dual norm. By the Duality Theorem, these norms must be equal. □

4.1 Alternate proof of maximum output probabilities

Incidentally, strong duality of the problems (1), (2) also yields an alternate proof of a result from Ref. [GW07] about maximum output probabilities. That result is stated as follows.

Theorem 4 (Maximum output probabilities [GW07])**.** *Let $\{S_a\}$ be an r-round measuring strategy. The maximum probability with which $\{S_a\}$ can be forced to produce a given outcome a by any r-round co-strategy is given by $\|S_a\|_{\diamond r}$. Furthermore, this quantity equals the minimum value p for which there exists an r-round non-measuring strategy S with $S_a \preceq pS$. An analogous result holds when $\{S_a\}$ is a co-strategy.*

Theorem 4 was originally proven via convex polarity. While semidefinite optimization duality and convex polarity are nominally different manifestations of the same underlying idea, some readers might be more familiar with semidefinite optimization duality than with convex polarity; the proof presented in the present paper should be more digestible to those readers.

New proof of Theorem 4. It is easy to see that the maximum probability with which $\{S_a\}$ can be forced to produce outcome a is expressed by the semidefinite optimization problem (1) with S_a in place of X. (As $S_a \succeq 0$, it is clear that the maximum is attained for operators T_0, T_1 with $T_1 = 0$, implying that T_0 is a non-measuring co-strategy.) By definition, this quantity is $\| S_a \|_{\diamond r}$.

By the strong duality of (1), (2), this quantity equals the minimum over all p such that there exists an r-round non-measuring strategy S with $-pS \preceq S_a \preceq pS$. As $S_a \succeq 0$, the first inequality is trivially satisfied by any $S \succeq 0$ and nonnegative p , and so the theorem follows. □

5 Distinguishability of convex sets of strategies

Our proof of the distinguishability of convex sets of strategies is an adaptation of the proof appearing in Ref. [GW05] with states and measurements replaced by strategies and co-strategies and the trace and operator norms replaced with the strategy r-norm and its dual. The requisite properties of these new norms were established by Theorem 3.

Theorem 5 (Distinguishability of convex sets of strategies)**.** *Let $\mathbf{S}_0, \mathbf{S}_1 \subset \mathbf{Pos}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r})$ be nonempty convex sets of r-round strategies. There exists an r-round measuring co-strategy $\{T_0, T_1\}$ with the property that*

$$\langle T_0 - T_1, S_0 - S_1 \rangle \geq \min_{R_a \in \mathbf{S}_a} \| R_0 - R_1 \|_{\diamond r}$$

for all choices of $S_0 \in \mathbf{S}_0$ and $S_1 \in \mathbf{S}_1$. A similar statement holds in terms of the dual norm $\| \cdot \|_{\diamond r}^*$ for convex sets of co-strategies.

Proof. The proof for co-strategies is completely symmetric to the proof for strategies, so we address only strategies here. Let d denote the minimum distance between $\mathbf{S}_0$ and $\mathbf{S}_1$ as stated in the theorem. If $d = 0$ then the theorem is satisfied by the trivial r-round measuring co-strategy corresponding to a random coin flip. (For this trivial co-strategy, both T_0 and T_1 are equal to the identity divided by $2 \dim(\mathcal{X}_{1\ldots r})$.) For the remainder of this proof, we shall restrict our attention to the case $d > 0$.

Define

$$\mathbf{S} \stackrel{\text{def}}{=} \mathbf{S}_0 - \mathbf{S}_1 = \{ S_0 - S_1 : S_0 \in \mathbf{S}_0, S_1 \in \mathbf{S}_1 \}$$

and let

$$\mathbf{B} \stackrel{\text{def}}{=} \{ B \in \mathbf{Her}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r}) : \| B \|_{\diamond r} < d \}$$

denote the open ball of radius d with respect to the $\| \cdot \|_{\diamond r}$ norm. The sets $\mathbf{S}$ and $\mathbf{B}$ are nonempty disjoint sets of Hermitian operators, both are convex, and $\mathbf{B}$ is open. By the Separation Theorem from convex analysis, there exists a Hermitian operator H and a scalar α such that

$$\langle H, S \rangle \geq \alpha > \langle H, B \rangle$$

for all $S \in \mathbf{S}$ and $B \in \mathbf{B}$.

For every choice of $B \in \mathbf{B}$ it holds that $-B \in \mathbf{B}$, from which it follows that $|\langle H, B\rangle| < \alpha$ for all $B \in \mathbf{B}$ and hence $\alpha > 0$. Moreover, as $\mathbf{B}$ is the open ball of radius d in the norm $\|\cdot\|_{\diamond r}$, it follows from the duality of the strategy r-norms (Theorem 3) that

$$\|H\|_{\diamond r}^{*} \leq \alpha/d.$$

Now let $\hat{H} = \frac{d}{\alpha}H$ be the normalization of H so that $\|\hat{H}\|_{\diamond r}^{*} \leq 1$. It follows from Theorem 3 that

$$\hat{H} = T_0 - T_1$$

for some r-round measuring co-strategy $\{T_0, T_1\}$. It remains only to verify that $\{T_0, T_1\}$ has the desired property: for every choice of $S_0 \in \mathbf{S}_0$ and $S_1 \in \mathbf{S}_1$ we have

$$\langle T_0 - T_1, S_0 - S_1\rangle = \langle \hat{H}, S_0 - S_1\rangle = \frac{d}{\alpha}\langle H, S_0 - S_1\rangle \geq d$$

as desired. □

The claimed result regarding the distinguishability of convex sets of strategies now follows immediately. To recap, let $\mathbf{S}_0, \mathbf{S}_1$ be convex sets of strategies and let $\{T_0, T_1\}$ denote the measuring co-strategy from Theorem 5 that distinguishes elements in $\mathbf{S}_0$ from elements in $\mathbf{S}_1$. Suppose Bob selects $S_0 \in \mathbf{S}_0$ and $S_1 \in \mathbf{S}_1$ arbitrarily and then selects $S \in \{S_0, S_1\}$ uniformly at random. As derived in Section 3, if Alice acts according to $\{T_0, T_1\}$ then the probability with which she correctly guesses whether $S \in \mathbf{S}_0$ or $S \in \mathbf{S}_1$ is given by

$$\frac{1}{2} + \frac{1}{4}\langle T_0 - T_1, S_0 - S_1\rangle \geq \frac{1}{2} + \frac{1}{4}\min_{R_a \in \mathbf{S}_a} \|R_0 - R_1\|_{\diamond r}$$

as desired.

A Appendix to Section 4: formal proof of semidefinite optimization duality

This appendix contains a formal proof that the semidefinite optimization problems (1), (2) from Section 4 satisfy strong duality. In other words, their optimal values are equal and are both achieved by feasible solutions.

A.1 Review of the linear map form for semidefinite optimization

The semidefinite optimization problem discussed in this appendix is expressed in *linear map form*. While the linear map form differs superficially from the more conventional *standard form* for these problems, the two forms can be shown to be equivalent and the linear map form is more convenient for our purpose. Watrous provides a helpful overview of this form of semidefinite optimization [Wat09]. For completeness, that overview is reproduced here.

A *semidefinite optimization problem* for spaces $\mathcal{P}, \mathcal{Q}$ is specified by a triple (Ψ, A, B) where $\Psi : \mathbf{L}(\mathcal{P}) \to \mathbf{L}(\mathcal{Q})$ is a Hermitian-preserving linear map and $A \in \mathbf{Her}(\mathcal{P})$ and $B \in \mathbf{Her}(\mathcal{Q})$. This triple specifies two optimization problems:

Primal problem		Dual problem	
maximize	$\langle A, P\rangle$	minimize	$\langle B, Q\rangle$
subject to	$\Psi(P) \preceq B$	subject to	$\Psi^{*}(Q) \succeq A$
	$P \in \mathbf{Pos}(\mathcal{P})$		$Q \in \mathbf{Pos}(\mathcal{Q})$

(Here $\Psi^* : \mathbf{L}(\mathcal{Q}) \to \mathbf{L}(\mathcal{D})$ denotes the adjoint of Ψ.) An operator P obeying the constraints of the primal problem is said to be *primal feasible*, while an operator Q obeying the constraints of the dual problem is called *dual feasible*. The functions $P \mapsto \langle A, P \rangle$ and $Q \mapsto \langle B, Q \rangle$ are called the primal and dual *objective functions*, respectively. Let

$$\alpha \stackrel{\text{def}}{=} \sup \{ \langle A, P \rangle : \ P \text{ is primal feasible} \}$$
$$\beta \stackrel{\text{def}}{=} \inf \{ \langle B, Q \rangle : \ Q \text{ is dual feasible} \}$$

denote the *optimal values* of the primal and dual problems. (If there are no primal or dual feasible operators then we adopt the convention $\alpha = -\infty$ and $\beta = \infty$, respectively.)

Semidefinite optimization problems derive great utility from the notions of *weak* and *strong duality*. Weak duality asserts that $\alpha \leq \beta$ for all triples (Ψ, A, B), whereas strong duality provides conditions on (Ψ, A, B) under which $\alpha = \beta$. Two such conditions are stated explicitly as follows.

Fact 6 (Strong duality conditions—see Ref. [BV04])**.** *Let (Ψ, A, B) be a semidefinite optimization problem. The following hold:*

1. *(Strict primal feasibility.) Suppose β is finite and there exists $P \succ 0$ with $\Psi(P) \prec B$. Then $\alpha = \beta$ and β is achieved by some dual feasible operator.*
2. *(Strict dual feasibility.) Suppose α is finite and there exists $Q \succ 0$ with $\Psi^*(Q) \succ A$. Then $\alpha = \beta$ and α is achieved by some primal feasible operator.*

A.2 A semidefinite optimization problem for the strategy r-norm

Let us construct a triple (Ψ, A, B) whose primal problem is equivalent to the problem (1) from Section 4. To this end, it is helpful to observe that (1) can be written more explicitly via the linear characterization of co-strategies mentioned in Section 2:

$$\begin{aligned}
\text{maximize} \quad & \langle X, T_0 - T_1 \rangle \\
\text{subject to} \quad & T_0 + T_1 = T_{[r]} \otimes I_{\mathcal{Y}_r} \\
& \operatorname{Tr}_{\mathcal{X}_r} \left(T_{[r]} \right) = T_{[r-1]} \otimes I_{\mathcal{Y}_{r-1}} \\
& \qquad \vdots \\
& \operatorname{Tr}_{\mathcal{X}_2} \left(T_{[2]} \right) = T_{[1]} \otimes I_{\mathcal{Y}_1} \\
& \operatorname{Tr} \left(T_{[1]} \right) = 1 \\
& T_0, T_1 \in \mathbf{Pos}(\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r}) \\
& T_{[i]} \in \mathbf{Pos}(\mathcal{Y}_{1\ldots i-1} \otimes \mathcal{X}_{1\ldots i}) \qquad (1 \leq i \leq r)
\end{aligned}$$

The triple (Ψ, A, B) is chosen so that its primal problem captures an inequality relaxation of the above problem. The components of (Ψ, A, B) are most conveniently expressed in block diagonal form via the intuitive shorthand notation $\operatorname{diag}(\cdot)$ defined so that, for example,

$$\operatorname{diag}\left(P, P'\right) \stackrel{\text{def}}{=} \operatorname{diag}\begin{pmatrix} P \\ P' \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} P & 0 \\ 0 & P' \end{pmatrix}.$$

The operators A, B are given by

$$A = \operatorname{diag}(X, -X, 0, \ldots, 0) \qquad B = \operatorname{diag}(0, \ldots, 0, 1)$$

and the linear map Ψ is given by

$$\Psi : \operatorname{diag}\begin{pmatrix} T_0 \\ T_1 \\ T_{[r]} \\ \vdots \\ T_{[1]} \end{pmatrix} \mapsto \operatorname{diag}\begin{pmatrix} T_0 + T_1 - T_{[r]} \otimes I_{\mathcal{Y}_r} \\ \operatorname{Tr}_{\mathcal{X}_r}\left(T_{[r]}\right) - T_{[r-1]} \otimes I_{\mathcal{Y}_{r-1}} \\ \vdots \\ \operatorname{Tr}_{\mathcal{X}_2}\left(T_{[2]}\right) - T_{[1]} \otimes I_{\mathcal{Y}_1} \\ \operatorname{Tr}\left(T_{[1]}\right) \end{pmatrix}$$

It is straightforward but tedious to verify that the optimal value of the primal problem described by (Ψ, A, B) is equal to $\|X\|_{\diamond r}$. To this end, let T be any positive semidefinite operator with diagonal blocks $T_0, T_1, T_{[r]}, \ldots, T_{[1]}$. The primal objective value at T is given by

$$\langle A, T\rangle = \langle X, T_0\rangle + \langle -X, T_1\rangle = \langle X, T_0 - T_1\rangle$$

as desired, so it remains only to verify that the constraint $\Psi(T) \preceq B$ enforces the property that $T_0 + T_1$ is an r-round non-measuring co-strategy. The following lemma serves that purpose.

Lemma 7 (Correctness of the primal problem)**.** *The optimal value of the primal problem (Ψ, A, B) is achieved by a primal feasible solution $T^\star$ whose diagonal blocks $T_0^\star, T_1^\star, T_{[r]}^\star, \ldots, T_{[1]}^\star$ have the property that $T_0^\star + T_1^\star$ is an r-round non-measuring co-strategy.*

Proof. The proof is a standard "slackness" argument: any feasible solution with unsaturated inequality constraints can be "inflated" so as to saturate all the constraints without decreasing the objective value of that solution.

Formally, we begin by observing that the optimal value must be achieved by some primal feasible T, as the set of feasible solutions is easily seen to be compact. (In particular, each block of T has trace not exceeding $\dim(\mathcal{Y}_{1\ldots r})$.) Let $T_0, T_1, T_{[r]}, \ldots, T_{[1]}$ denote the diagonal blocks of T. As T is primal feasible it holds that $\Psi(T) \preceq B$ and hence

$$\begin{aligned} T_0 + T_1 &\preceq T_{[r]} \otimes I_{\mathcal{Y}_r} \\ \operatorname{Tr}_{\mathcal{X}_r}\left(T_{[r]}\right) &\preceq T_{[r-1]} \otimes I_{\mathcal{Y}_{r-1}} \\ &\vdots \\ \operatorname{Tr}_{\mathcal{X}_2}\left(T_{[2]}\right) &\preceq T_{[1]} \otimes I_{\mathcal{Y}_1} \\ \operatorname{Tr}\left(T_{[1]}\right) &\leq 1. \end{aligned}$$

To prove the lemma it suffices to construct a feasible solution $T^\star$ whose objective value equals that of T and whose diagonal blocks $T_0^\star, T_1^\star, T_{[r]}^\star, \ldots, T_{[1]}^\star$ meet the above constraints with equality.

To this end, the desired blocks $T_{[1]}^\star, \ldots, T_{[r]}^\star$ are constructed inductively from $T_{[1]}, \ldots, T_{[r]}$ so as to satisfy $T_{[i]}^\star \succeq T_{[i]}$ for each $i = 1, \ldots, r$. For the base case, it is clear that there is a $T_{[1]}^\star \succeq T_{[1]}$ with $\operatorname{Tr}(T_{[1]}^\star) = 1$. For the inductive step, it holds that

$$\operatorname{Tr}_{\mathcal{X}_i}\left(T_{[i]}\right) \preceq T_{[i-1]} \otimes I_{\mathcal{Y}_{i-1}} \preceq T_{[i-1]}^\star \otimes I_{\mathcal{Y}_{i-1}}.$$

(Here we have used the operator inequality $P^\star \succeq P \implies P^\star \otimes I \succeq P \otimes I$, an observation that follows from the fact that $A \succeq 0 \implies A \otimes I \succeq 0$ by substituting $A = P^\star - P$.) Thus, there must exist $Q \succeq 0$ with

$$\mathrm{Tr}_{\mathcal{X}_i}(T_{[i]}) + Q = T^\star_{[i-1]} \otimes I_{\mathcal{Y}_{i-1}}.$$

Choose any $R \succeq 0$ with $\mathrm{Tr}_{\mathcal{X}_i}(R) = Q$. Selecting $T^\star_{[i]} = T_{[i]} + R$, it holds that $T^\star_{[i]} \succeq T_{[i]}$ and

$$\mathrm{Tr}_{\mathcal{X}_i}\left(T^\star_{[i]}\right) = T^\star_{[i-1]} \otimes I_{\mathcal{Y}_{i-1}}$$

as claimed.

The final blocks $T^\star_0, T^\star_1$ are constructed similarly. As $T^\star_{[r]} \succeq T_{[r]}$, it holds that

$$T_0 + T_1 \preceq T_{[r]} \otimes I_{\mathcal{Y}_r} \preceq T^\star_{[r]} \otimes I_{\mathcal{Y}_r}$$

and hence there exists $D \succeq 0$ with

$$T_0 + T_1 + D = T^\star_{[r]} \otimes I_{\mathcal{Y}_r}.$$

Selecting

$$T^\star_0 = T_0 + \frac{1}{2}D, \qquad T^\star_1 = T_1 + \frac{1}{2}D,$$

it holds that $T^\star_0, T^\star_1 \succeq 0$, that $T^\star_0 + T^\star_1$ is an r-round co-strategy, and that $T^\star_0 - T^\star_1 = T_0 - T_1$, from which it follows that $T^\star$ and T have the same objective value. □

A.3 The dual problem

In this section it is shown that the dual problem for (Ψ, A, B) is equivalent to the optimization problem (2) from Section 4. To this end, it is helpful to observe that (2) can be written more explicitly via the linear characterization of strategies mentioned in Section 2:

$$\begin{array}{ll}
\text{minimize} & p \\
\text{subject to} & S_{[r]} \succeq \pm X \\
& \mathrm{Tr}_{\mathcal{Y}_r}\left(S_{[r]}\right) = S_{[r-1]} \otimes I_{\mathcal{X}_r} \\
& \qquad \vdots \\
& \mathrm{Tr}_{\mathcal{Y}_2}\left(S_{[2]}\right) = S_{[1]} \otimes I_{\mathcal{X}_2} \\
& \mathrm{Tr}_{\mathcal{Y}_1}\left(S_{[1]}\right) = pI_{\mathcal{X}_1} \\
& S_{[i]} \in \mathbf{Pos}(\mathcal{Y}_{1\ldots i} \otimes \mathcal{X}_{1\ldots i}) \qquad (1 \leq i \leq r) \\
& p \geq 0
\end{array}$$

In order to demonstrate the desired equivalence between (2) and the dual problem for (Ψ, A, B) we require an explicit formula for the adjoint linear map Ψ^*. It is straightforward but tedious to derive such a formula. To this end, let S, T be operators with diagonal blocks $S_{[r]}, \ldots, S_{[1]}, p$ and $T_0, T_1, T_{[r]}, \ldots, T_{[1]}$, respectively. As $\langle \Psi(T), S\rangle = \langle T, \Psi^*(S)\rangle$, a formula for Ψ^* may be derived by writing $\langle \Psi(T), S\rangle$ in terms

of the blocks of T:

$$
\begin{aligned}
&\langle \Psi(T), S \rangle \\
&= \underbrace{\left\langle T_0 + T_1 - T_{[r]} \otimes I_{\mathcal{Y}_r}, S_{[r]} \right\rangle}_{\text{expand}} + \left(\sum_{i=1}^{r-1} \underbrace{\left\langle \mathrm{Tr}_{\mathcal{X}_{i+1}} \left(T_{[i+1]} \right) - T_{[i]} \otimes I_{\mathcal{Y}_i}, S_{[i]} \right\rangle}_{\text{expand}} \right) + p \, \mathrm{Tr} \left(T_{[1]} \right) \\
&= \left\langle T_0, S_{[r]} \right\rangle + \left\langle T_1, S_{[r]} \right\rangle - \underbrace{\left\langle T_{[r]} \otimes I_{\mathcal{Y}_r}, S_{[r]} \right\rangle}_{\text{isolate } T_{[r]}} + \left(\sum_{i=1}^{r-1} \underbrace{\left\langle \mathrm{Tr}_{\mathcal{X}_{i+1}} \left(T_{[i+1]} \right), S_{[i]} \right\rangle}_{\text{isolate } T_{[i+1]}} - \underbrace{\left\langle T_{[i]} \otimes I_{\mathcal{Y}_i}, S_{[i]} \right\rangle}_{\text{isolate } T_{[i]}} \right) + \underbrace{p \, \mathrm{Tr} \left(T_{[1]} \right)}_{\text{isolate } T_{[1]}} \\
&= \left\langle T_0, S_{[r]} \right\rangle + \left\langle T_1, S_{[r]} \right\rangle - \underbrace{\left\langle T_{[r]}, \mathrm{Tr}_{\mathcal{Y}_r} \left(S_{[r]} \right) \right\rangle}_{\substack{\text{absorb into summation,} \\ \text{remove } T_{[1]} \text{ term}}} + \left(\sum_{i=1}^{r-1} \left\langle T_{[i+1]}, S_{[i]} \otimes I_{\mathcal{X}_{i+1}} \right\rangle - \left\langle T_{[i]}, \mathrm{Tr}_{\mathcal{Y}_i} \left(S_{[i]} \right) \right\rangle \right) + \left\langle T_{[1]}, p I_{\mathcal{X}_1} \right\rangle \\
&= \left\langle T_0, S_{[r]} \right\rangle + \left\langle T_1, S_{[r]} \right\rangle + \left(\sum_{i=2}^{r} \underbrace{\left\langle T_{[i]}, S_{[i-1]} \otimes I_{\mathcal{X}_i} \right\rangle - \left\langle T_{[i]}, \mathrm{Tr}_{\mathcal{Y}_i} \left(S_{[i]} \right) \right\rangle}_{\text{collect } T_{[i]} \text{ terms}} \right) + \underbrace{\left\langle T_{[1]}, p I_{\mathcal{X}_1} \right\rangle - \left\langle T_{[1]}, \mathrm{Tr}_{\mathcal{Y}_1} \left(S_{[1]} \right) \right\rangle}_{\text{collect } T_{[1]} \text{ terms}} \\
&= \left\langle T_0, S_{[r]} \right\rangle + \left\langle T_1, S_{[r]} \right\rangle + \left(\sum_{i=2}^{r} \left\langle T_{[i]}, S_{[i-1]} \otimes I_{\mathcal{X}_i} - \mathrm{Tr}_{\mathcal{Y}_i} \left(S_{[i]} \right) \right\rangle \right) + \left\langle T_{[1]}, p I_{\mathcal{X}_1} - \mathrm{Tr}_{\mathcal{Y}_1} \left(S_{[1]} \right) \right\rangle .
\end{aligned}
$$

It is now clear that Ψ^* is given by

$$
\Psi^* : \mathrm{diag} \begin{pmatrix} S_{[r]} \\ \vdots \\ S_{[1]} \\ p \end{pmatrix} \mapsto \mathrm{diag} \begin{pmatrix} S_{[r]} \\ S_{[r]} \\ S_{[r-1]} \otimes I_{\mathcal{X}_r} - \mathrm{Tr}_{\mathcal{Y}_r} \left(S_{[r]} \right) \\ \vdots \\ S_{[1]} \otimes I_{\mathcal{X}_2} - \mathrm{Tr}_{\mathcal{Y}_2} \left(S_{[2]} \right) \\ p I_{\mathcal{X}_1} - \mathrm{Tr}_{\mathcal{Y}_1} \left(S_{[1]} \right) \end{pmatrix} .
$$

As was done in section A.2 for the primal problem, it is now argued that the dual problem for (Ψ, A, B) is an inequality relaxation of (2). To this end, Let S be any positive semidefinite operator with diagonal blocks $S_{[r]}, \ldots, S_{[1]}, p$. The dual objective value at S is given by

$$
\langle B, S \rangle = \langle 1, p \rangle = p
$$

as desired, so it remains only to verify that the constraint $\Psi^*(S) \succeq A$ enforces the property that $S_{[r]}$ is an r-round non-measuring strategy multiplied by p. The following lemma serves that purpose.

Lemma 8 (Correctness of the dual problem)**.** *For each dual feasible solution S to (Ψ, A, B) (including optimal or near-optimal solutions) there exists another dual feasible solution $S^\star$ whose objective value $p^\star$ equals that of S and whose diagonal blocks $S^\star_{[r]}, \ldots, S^\star_{[1]}, p^\star$ have the property that $S^\star_{[r]}$ is an r-round non-measuring strategy multiplied by $p^\star$.*

Proof. The proof closely follows the slackness argument used in the proof of Lemma 7. Let $S_{[r]}, \ldots, S_{[1]}, p$ denote the diagonal blocks of S. As S is dual feasible it holds that $\Psi^*(S) \succeq A$ and hence

$$\begin{aligned} S_{[r]} &\succeq \pm X \\ S_{[r-1]} \otimes I_{\mathcal{X}_r} &\succeq \mathrm{Tr}_{\mathcal{Y}_r}(S_{[r]}) \\ &\vdots \\ S_{[1]} \otimes I_{\mathcal{X}_2} &\succeq \mathrm{Tr}_{\mathcal{Y}_2}(S_{[2]}) \\ p I_{\mathcal{X}_1} &\succeq \mathrm{Tr}_{\mathcal{Y}_1}(S_{[1]}). \end{aligned}$$

To prove the lemma it suffices to construct a dual feasible solution $S^\star$ whose objective value equals that of S and whose diagonal blocks $S^\star_{[r]}, \ldots, S^\star_{[1]}, p^\star$ meet the above constraints with equality (except the constraint $S_{[r]} \succeq \pm X$).

To this end, the desired blocks $S^\star_{[1]}, \ldots, S^\star_{[r]}$ are constructed inductively from $S_{[1]}, \ldots, S_{[r]}$ so as to satisfy $S^\star_{[i]} \succeq S_{[i]}$ for each $i = 1, \ldots, r$. For the base case, it is clear that there is an $S^\star_{[1]} \succeq S_{[1]}$ with $p I_{\mathcal{X}_1} = \mathrm{Tr}_{\mathcal{Y}_1}(S^\star_{[1]})$. For the inductive step, it holds that

$$S^\star_{[i-1]} \otimes I_{\mathcal{X}_i} \succeq S_{[i-1]} \otimes I_{\mathcal{X}_i} \succeq \mathrm{Tr}_{\mathcal{Y}_i}(S_{[i]}).$$

(Again, we have used the operator inequality $P^\star \succeq P \implies P^\star \otimes I \succeq P \otimes I$ for any $P \succeq 0$.) Thus, there must exist $Q \succeq 0$ with

$$S^\star_{[i-1]} \otimes I_{\mathcal{X}_i} = \mathrm{Tr}_{\mathcal{Y}_i}(S_{[i]}) + Q.$$

Choose any $R \succeq 0$ with $\mathrm{Tr}_{\mathcal{Y}_i}(Q) = R$. Selecting $S^\star_{[i]} = S_{[i]} + R$, it holds that $S^\star_{[i]} \succeq S_{[i]}$ and

$$S^\star_{[i-1]} \otimes I_{\mathcal{X}_i} = \mathrm{Tr}_{\mathcal{Y}_i}(S^\star_{[i]})$$

as claimed.

Selecting $p^\star = p$, it holds that $S^\star_{[r]}$ is an r-round non-measuring strategy multiplied by $p^\star$ as desired. As $S^\star_{[r]} \succeq S_{[r]} \succeq \pm X$ and $p^\star = p$, it follows that $S^\star$ is a dual feasible solution that achieves the same objective value as S. □

A.4 Strong duality

Thus far, it has been argued that the optimal values of the problems (1), (2) from Section 4 are captured by the primal and dual semidefinite optimization problems associated with the triple (Ψ, A, B). It remains only to show that these two quantities are equal. Equality is established by showing that (Ψ, A, B) satisfies the conditions for strong duality from Fact 6.

Theorem 9 (Strong duality of (Ψ, A, B))**.** *There exists a primal feasible operator T and a dual feasible operator S such that $\langle A, T \rangle = \langle B, S \rangle$.*

Proof. The proof is via item 1 of Fact 6 (Strong duality conditions). Specifically, it is shown that β is finite and the primal problem is strictly feasible. It then follows from Fact 6 that $\alpha = \beta$ and that β is achieved for some dual feasible operator. To complete the proof, it suffices to note that the optimal value α is also achieved by a primal feasible operator, as established in Lemma 7 (Correctness of the primal problem).

First, it is argued that β is finite. As $B \succeq 0$, any dual feasible solution has nonnegative objective value. Thus, to show that β is finite it suffices to exhibit a single dual feasible solution. That solution S is a block-diagonal matrix with blocks $S_{[r]}, \ldots, S_{[1]}, p$ given by

$$\begin{aligned} S_{[r]} &= \|X\| I_{\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r}} \\ S_{[r-1]} &= \|X\| \dim(\mathcal{Y}_r) I_{\mathcal{Y}_{1\ldots r-1} \otimes \mathcal{X}_{1\ldots r-1}} \\ &\vdots \\ S_{[1]} &= \|X\| \dim(\mathcal{Y}_{2\ldots r}) I_{\mathcal{Y}_1 \otimes \mathcal{X}_1} \\ p &= \|X\| \dim(\mathcal{Y}_{1\ldots r}). \end{aligned}$$

As X is Hermitian it holds that

$$-S_{[r]} = -\|X\| I_{\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r}} \preceq X \preceq \|X\| I_{\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r}} = S_{[r]}$$

and hence S is dual feasible as desired.

Finally, it is shown that the primal is strictly feasible. Choose $\delta \in (0, \frac{1}{r+1})$ and let T be the block-diagonal operator with diagonal blocks $T_0, T_1, T_{[r]}, \ldots, T_{[1]}$ given by

$$\begin{aligned} T_{[i]} &= \frac{1 - i\delta}{\dim(\mathcal{X}_{1\ldots i})} I_{\mathcal{Y}_{1\ldots i-1} \otimes \mathcal{X}_{1\ldots i}} \qquad (1 \leq i \leq r) \\ T_0 = T_1 &= \frac{1 - (r+1)\delta}{2 \dim(\mathcal{X}_{1\ldots r})} I_{\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r}}. \end{aligned}$$

It is clear that $T \succ 0$ and it is tedious but straightforward to verify that $\Psi(T) \prec B$. Specifically, we have

$$\begin{aligned} \operatorname{Tr}\left(T_{[1]}\right) = 1 - \delta &< 1 \\ \operatorname{Tr}_{\mathcal{X}_2}\left(T_{[2]}\right) = \frac{1 - 2\delta}{\dim(\mathcal{X}_1)} I_{\mathcal{Y}_1 \otimes \mathcal{X}_1} &\prec \frac{1 - \delta}{\dim(\mathcal{X}_1)} I_{\mathcal{Y}_1 \otimes \mathcal{X}_1} = T_{[1]} \otimes I_{\mathcal{Y}_1} \\ &\vdots \\ \operatorname{Tr}_{\mathcal{X}_r}\left(T_{[r]}\right) = \frac{1 - r\delta}{\dim(\mathcal{X}_{1\ldots r-1})} I_{\mathcal{Y}_{1\ldots r-1} \otimes \mathcal{X}_{1\ldots r-1}} &\prec \frac{1 - (r-1)\delta}{\dim(\mathcal{X}_{1\ldots r-1})} I_{\mathcal{Y}_{1\ldots r-1} \otimes \mathcal{X}_{1\ldots r-1}} = T_{[r-1]} \otimes I_{\mathcal{Y}_{r-1}} \\ T_0 + T_1 = \frac{1 - (r+1)\delta}{\dim(\mathcal{X}_{1\ldots r})} I_{\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r}} &\prec \frac{1 - r\delta}{\dim(\mathcal{X}_{1\ldots r})} I_{\mathcal{Y}_{1\ldots r} \otimes \mathcal{X}_{1\ldots r}} = T_{[r]} \otimes I_{\mathcal{Y}_r}. \end{aligned}$$

It now follows from item 1 of Fact 6 that $\alpha = \beta$ and that β is achieved by some dual feasible operator. □

Acknowledgements

The author is grateful to Giulio Chiribella and John Watrous for informative discussions. This research was 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. Part of this research was conducted while the author was a graduate student at the University of Waterloo, at which time this research was supported by Canada's NSERC and the David R. Cheriton School of Computer Science at the University of Waterloo.

References

[AKN98] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the Thirtieth ACM Symposium on Theory of Computing*, pages 20–30, 1998. arXiv:quant-ph/9806029v1.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[CDP08] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Memory effects in quantum channel discrimination. *Physical Review Letters*, 101:180501, 2008. arXiv:0803.3237v3 [quant-ph].

[CDP09a] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Optimal covariant quantum networks. In Alexander Lvovsky, editor, *Proceedings of the 9th International Conference on Quantum Communication, Measurement and Computing (QCMC '08)*, volume 1110, pages 47–56. American Institute of Physics, 2009. arXiv:0812.3922v1 [quant-ph].

[CDP09b] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Theoretical framework for quantum networks. *Physical Review A*, 80(2):022339, 2009. arXiv:0904.4483v2 [quant-ph].

[CDP+09c] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, Dirk Schlingemann, and Reinhard F. Werner. A short impossibility proof of quantum bit commitment. arXiv:0905.3801v1 [quant-ph], 2009.

[GLN05] Alexei Gilchrist, Nathan Langford, and Michael Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71:062310, 2005. arXiv:quant-ph/0408063v2.

[GW05] Gus Gutoski and John Watrous. Quantum interactive proofs with competing provers. In *Proceedings of the 22nd Symposium on Theoretical Aspects of Computer Science (STACS'05)*, volume 3404 of *Lecture Notes in Computer Science*, pages 605–616. Springer, 2005. arXiv:cs/0412102v1 [cs.CC].

[GW07] Gus Gutoski and John Watrous. Toward a general theory of quantum games. In *Proceedings of the 39th ACM Symposium on Theory of Computing (STOC 2007)*, pages 565–574, 2007. arXiv:quant-ph/0611234v2.

[Hel69] Carl Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, 1969.

[HJ85] Roger Horn and Charles Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[JKPP11] Nathaniel Johnston, David W. Kribs, Vern Paulsen, and Rajesh Pereira. Minimal and maximal operator spaces and operator systems in entanglement theory. *Journal of Functional Analysis*, 260:2407–2423, 2011. arXiv:1010.1432v1 [math.OA].

[Kit97] Alexei Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997.

[NC00] Michael Nielsen and Issac Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Pau02] Vern Paulsen. *Completely Bounded Maps and Operator Algebras*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2002.

[RW05] Bill Rosgen and John Watrous. On the hardness of distinguishing mixed-state quantum computations. In *Proceedings of the 20th Conference on Computational Complexity*, pages 344–354, 2005. arXiv:cs/0407056v1 [cs.CC].

[Smi83] Roger Smith. Completely bounded maps between C^*-algebras. *Journal of the London Mathematical Society*, 27:157–166, 1983.

[SSŻ09] Łukasz Skowronek, Erling Størmer, and Karol Życzkowski. Cones of positive maps and their duality relations. *Journal of Mathematical Physics*, 50:062106, 2009. arXiv:0902.4877v1 [quant-ph].

[Tim03] Richard Timoney. Computing the norms of elementary operators. *Illinois Journal of Mathematics*, 47(4):1207–1226, 2003.

[Wat05] John Watrous. Notes on super-operator norms induced by Schatten norms. *Quantum Information and Computation*, 5(1):58–68, 2005. arXiv:quant-ph/0411077v1.

[Wat08] John Watrous. Distinguishing quantum operations having few Kraus operators. *Quantum Information and Computation*, 8(9):819–833, 2008. arXiv:0710.0902v3 [quant-ph].

[Wat09] John Watrous. Semidefinite programs for completely bounded norms. *Theory of Computing*, 5:217–238, 2009. arXiv:0901.4709v2 [quant-ph].