Semidefinite Programming in the Theory of Quantum Information

John Watrous
Institute for Quantum Computing
University of Waterloo

Introduction

Semidefinite programming is a powerful tool in quantum information theory with many interesting applications.

Two aspects of semidefinite programming that contribute to its uses:

1. It may be used as an **analytical** tool to prove interesting theorems, often based on semidefinite programming duality.

   Examples: lower-bounds on strong quantum coin-flipping, perfect parallel repetition of entangled XOR games, optimality of the generalized adversary method, and a simple proof of the quantum substate theorem.

2. It is also a useful **algorithmic** paradigm: there are efficient algorithms for solving typical semidefinite programs.

   Implementations are available (CVX for MATLAB, for instance) and can be used for concrete problem instances.

Outline of the talk

1. Preliminary remarks (mostly about notation).
2. Fundamentals aspects of semidefinite programming.
   - Definitions.
   - Duality.
   - Basic examples.
   - Alternate forms of semidefinite programs.
3. Applications of semidefinite programming in quantum information theory.
   - Optimizing over measurements.
   - Optimizing over channels.
   - Parallel repetition of XOR games.
   - Completely bounded trace norm.
   - Quantum interactions (such as coin-flipping).

1. Preliminary remarks

Complex Euclidean spaces and operators

Complex Euclidean spaces (or finite-dimensional Hilbert spaces)

For any finite, non-empty set Σ, we write $\mathbb{C}^\Sigma$ to denote the space of all complex vectors indexed by Σ. (You may imagine $\Sigma = \{1, \ldots, n\}$ and write $\mathbb{C}^n$ rather than $\mathbb{C}^\Sigma$, if you prefer.)

In this talk, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ will always denote such spaces.

Standard inner product:

$$\langle u, v \rangle = \sum_{a \in \Sigma} \overline{u(a)}\, v(a).$$

Linear operators

Let $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ denote the set of all linear mappings (or *operators*) from a vector space $\mathcal{X}$ to a vector space $\mathcal{Y}$. Shorthand: $\mathrm{L}(\mathcal{X})$ means $\mathrm{L}(\mathcal{X}, \mathcal{X})$.

If $\mathcal{X} = \mathbb{C}^\Sigma$ and $\mathcal{Y} = \mathbb{C}^\Gamma$, one may identify $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ with the set of all matrices with rows indexed by Γ and columns indexed by Σ.

Adjoints and inner products of operators

The adjoint of an operator

For every operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, there is a uniquely defined operator $A^* \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ (called the *adjoint* of A) such that

$$\langle y, Ax \rangle = \langle A^* y, x \rangle$$

for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

As a matrix, A^* is the *conjugate transpose* (or *Hermitian transpose*) of A, and is often denoted $A^\dagger$ by physicists and quantum information theorists.

Inner products of operators

For operators $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, we define

$$\langle A, B \rangle = \mathrm{Tr}\big(A^* B\big).$$

(This is sometimes called the *Hilbert-Schmidt inner product*.)

Hermitian and positive semidefinite operators

Hermitian operators

An operator $\mathrm{H} \in \mathrm{L}(\mathcal{X})$ is *Hermitian* if $\mathrm{H} = \mathrm{H}^*$. We will write $\mathrm{Herm}(\mathcal{X})$ to denote the set of all such operators.

For any choice of a Hermitian operator $\mathrm{H} \in \mathrm{Herm}(\mathcal{X})$, it holds that (1) every eigenvalue of H is a real number, and (2) there must exist an orthonormal basis of $\mathcal{X}$ consisting of eigenvectors of H.

Positive semidefinite operators

An operator $\mathrm{P} \in \mathrm{L}(\mathcal{X})$ is *positive semidefinite* if $\mathrm{P} \in \mathrm{Herm}(\mathcal{X})$ and all eigenvalues of P are nonnegative. We will write $\mathrm{Pos}(\mathcal{X})$ to denote the set of all such operators.

Positive semidefinite operators having trace equal to one are called *density operators*: $\mathrm{D}(\mathcal{X}) = \big\{\rho \in \mathrm{Pos}(\mathcal{X}) : \mathrm{Tr}(\rho) = 1\big\}$.

If $\mathrm{A}, \mathrm{B} \in \mathrm{Herm}(\mathcal{X})$, then the notations $\mathrm{A} \leq \mathrm{B}$ and $\mathrm{B} \geq \mathrm{A}$ mean that $\mathrm{B} - \mathrm{A} \in \mathrm{Pos}(\mathcal{X})$.

Linear mappings on operators

Linear mappings of the form

$$\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$$

are important in quantum information theory. (E.g., *quantum channels* are important examples of mappings of this form.) The space of all such mappings will be denoted $\mathrm{T}(\mathcal{X}, \mathcal{Y})$.

For every $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, we define the *adjoint mapping* $\Phi^* \in \mathrm{T}(\mathcal{Y}, \mathcal{X})$ to be the unique mapping that satisfies

$$\langle \mathrm{Y}, \Phi(\mathrm{X}) \rangle = \langle \Phi^*(\mathrm{Y}), \mathrm{X} \rangle$$

for all $\mathrm{X} \in \mathrm{L}(\mathcal{X})$ and $\mathrm{Y} \in \mathrm{L}(\mathcal{Y})$.

A mapping $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is *Hermiticity-preserving* if $\Phi(\mathrm{X}) \in \mathrm{Herm}(\mathcal{Y})$ for all $\mathrm{X} \in \mathrm{Herm}(\mathcal{X})$.

2. Fundamental aspects of semidefinite programming

Semidefinite programs

A **semidefinite program** (or **SDP** for short) is a pair of optimization problems, specified by a triple $(\Phi, \mathrm{A}, \mathrm{B})$, where:

1. $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is a Hermiticity-preserving mapping,
2. $\mathrm{A} \in \mathrm{Herm}(\mathcal{X})$, and
3. $\mathrm{B} \in \mathrm{Herm}(\mathcal{Y})$.

The pair of optimization problems is as follows:

Primal problem		Dual problem	
maximize:	$\langle \mathrm{A}, \mathrm{X} \rangle$	minimize:	$\langle \mathrm{B}, \mathrm{Y} \rangle$
subject to:	$\Phi(\mathrm{X}) = \mathrm{B}$	subject to:	$\Phi^*(\mathrm{Y}) \geq \mathrm{A}$
	$\mathrm{X} \in \mathrm{Pos}(\mathcal{X})$		$\mathrm{Y} \in \mathrm{Herm}(\mathcal{Y})$

(There are other formulations of these problems, including the so-called *standard form*, that we will discuss shortly.)

Optimal values

The **optimal value** of the **primal problem**

$$\begin{aligned} \text{maximize:} \quad & \langle A, X \rangle \\ \text{subject to:} \quad & \Phi(X) = B \\ & X \in \mathrm{Pos}(\mathcal{X}) \end{aligned}$$

is defined as $\alpha = \sup\big\{\langle A, X\rangle \,:\, X \in \mathrm{Pos}(\mathcal{X}),\, \Phi(X) = B\big\}$.

Similarly, the **optimal value** of the **dual problem**

$$\begin{aligned} \text{minimize:} \quad & \langle B, Y \rangle \\ \text{subject to:} \quad & \Phi^*(Y) \geq A \\ & Y \in \mathrm{Herm}(\mathcal{Y}) \end{aligned}$$

is defined as $\beta = \inf\big\{\langle B, Y\rangle \,:\, Y \in \mathrm{Herm}(\mathcal{Y}),\, \Phi^*(Y) \geq A\big\}$.

(The supremum/infimum cannot be replaced by a maximum/minimum in general; in some cases the optimal values will not be achieved.)

Example

For an arbitrary choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, a Hermiticity-preserving map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, and Hermitian operators $A \in \mathrm{Herm}(\mathcal{X})$ and $B \in \mathrm{Herm}(\mathcal{Y})$, we will have a valid SDP.

As a very simple example, we may take:

$$\mathcal{X} = \mathbb{C}^n \text{ (arbitrary } n\text{)} \qquad \text{and} \qquad \mathcal{Y} = \mathbb{C},$$

let $A \in \mathrm{Herm}(\mathcal{X})$ be arbitrary, and let $\Phi = \mathrm{Tr}$ and $B = 1$.

The primal problem looks like this:

maximize:	$\langle A, X \rangle$
subject to:	$\Phi(X) = B$
	$X \in \mathrm{Pos}(\mathcal{X})$

$\xrightarrow{\text{simplify}}$

maximize:	$\langle A, \rho \rangle$
subject to:	$\rho \in \mathrm{D}(\mathcal{X})$

(Note: it is typical that one only sees the simplified form.)

Example (continued)

To determine the **optimal value** of the **primal problem**, it is helpful to consider a spectral decomposition

$$A = \sum_{k=1}^{n} \lambda_k x_k x_k^*.$$

Here, $\lambda_1 \geq \cdots \geq \lambda_n$ are eigenvalues of A and $\{x_1, \ldots, x_n\}$ is an orthonormal basis of $\mathbb{C}^n$ consisting of corresponding eigenvectors.

For any $\rho \in \mathrm{D}(\mathcal{X})$, we have

$$\langle A, \rho \rangle = \sum_{k=1}^{n} \lambda_k \left\langle x_k x_k^*, \rho \right\rangle,$$

which is a **convex combination** of the eigenvalues $\lambda_1, \ldots, \lambda_n$.

The **optimal value** is $\alpha = \lambda_1$ (the **largest eigenvalue** of A).

Example (continued)

Now consider the **dual problem** for our example. Recall:

$$\mathcal{X} = \mathbb{C}^n,\ \mathcal{Y} = \mathbb{C},\ A \in \mathrm{Herm}(\mathcal{X}) \text{ (arbitrary)},\ \Phi = \mathrm{Tr},\ \text{and } B = 1.$$

The dual problem looks like this:

minimize:	$\langle B, Y \rangle$
subject to:	$\Phi^*(Y) \geq A$
	$Y \in \mathrm{Herm}(\mathcal{Y})$

simplify $\longrightarrow$

minimize:	λ
subject to:	$A \leq \lambda \mathbb{1}$
	$\lambda \in \mathbb{R}$

Here we used the fact that the adjoint of the mapping $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C})$ defined as $\Phi(X) = \mathrm{Tr}(X)$ is the mapping $\Phi^* \in \mathrm{T}(\mathbb{C}, \mathbb{C}^n)$ defined as

$$\Phi^*(\lambda) = \lambda \mathbb{1}.$$

This must be so, because $\langle \lambda, \mathrm{Tr}(X) \rangle = \langle \lambda \mathbb{1}, X \rangle$.

Example (continued)

To determine the **optimal value** of the **dual problem**, it is again helpful to consider a spectral decomposition

$$A = \sum_{k=1}^{n} \lambda_k x_k x_k^*.$$

The inequality $A \leq \lambda \mathbb{1}$ is equivalent to

$$\sum_{k=1}^{n} (\lambda - \lambda_k) x_k x_k^* \geq 0.$$

This holds if and only if $\lambda \geq \lambda_k$ for $k = 1, \ldots, n$.

The optimal value of the dual problem is therefore $\beta = \lambda_1$.

It is not an accident that the optimal primal value α and the optimal dual value β coincide—this usually happens.

Duality

For a semidefinite program specified by (Φ, A, B), we have defined the optimal primal value α and the optimal dual value β as

$$\alpha = \sup\bigl\{\langle A, X\rangle \,:\, X \in \mathrm{Pos}(\mathcal{X}),\ \Phi(X) = B\bigr\},$$
$$\beta = \inf\bigl\{\langle B, Y\rangle \,:\, Y \in \mathrm{Herm}(\mathcal{Y}),\ \Phi^*(Y) \geq A\bigr\}.$$

Weak duality: it always holds that $\alpha \leq \beta$.

To see this, suppose X is **primal feasible** and Y is **dual feasible**:

$$X \in \mathrm{Pos}(\mathcal{X}) \text{ and } \Phi(X) = B,$$
$$Y \in \mathrm{Herm}(\mathcal{Y}) \text{ and } \Phi^*(Y) \geq A.$$

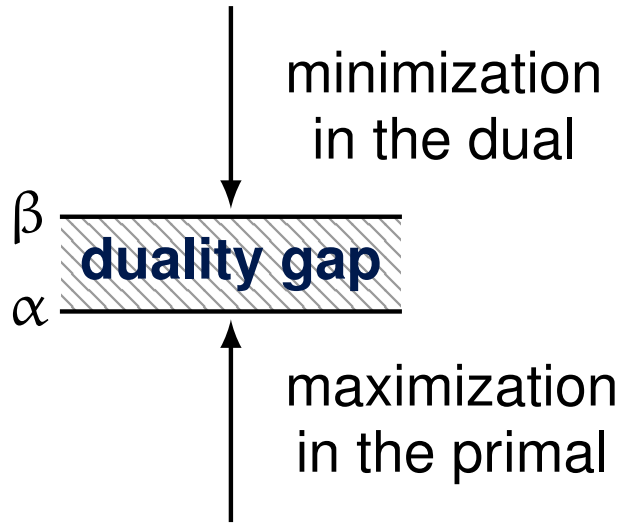
It follows that

$$\langle A, X\rangle \leq \langle \Phi^*(Y), X\rangle = \langle Y, \Phi(X)\rangle = \langle Y, B\rangle = \langle B, Y\rangle.$$

(The inequality follows from $\langle P, Q\rangle \geq 0$ for $P, Q \geq 0$.)

Duality (continued)

The relationship between the primal and dual optimal values is represented by this figure:



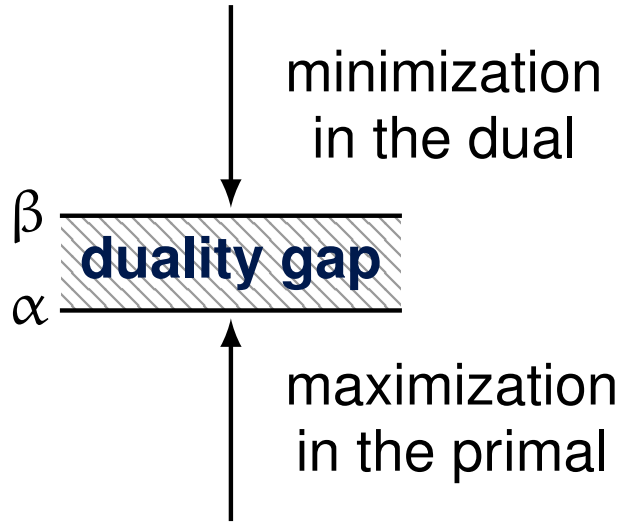
If we ever find a primal feasible X and a dual feasible Y such that

$$\langle A, X \rangle = \langle B, Y \rangle,$$

then we know we have found the optimal values; and moreover we have $\alpha = \beta$ (a condition known as **strong duality**).

Duality (continued)

The relationship between the primal and dual optimal values is represented by this figure:



Also note:

$$Y \text{ dual feasible} \Rightarrow \alpha \leq \langle B, Y \rangle$$
$$X \text{ primal feasible} \Rightarrow \beta \geq \langle A, X \rangle.$$

Every feasible point provides a bound on the complementary problem.

Strong duality

Strong duality refers to the situation in which the optimal primal and dual values are equal: $\alpha = \beta$. It is possible to construct SDPs for which **strong duality fails**. (We could have $\alpha = 0$ and $\beta = 1$, for instance.)

However, for most SDPs that arise naturally, strong duality will hold.

Slater conditions: strong duality holds under either of the following conditions:

1. The primal is **feasible** (there exists $X \in \mathrm{Pos}(\mathcal{X})$ with $\Phi(X) = B$) and the dual is **strictly feasible** (there exists $Y \in \mathrm{Herm}(\mathcal{Y})$ with $\Phi^*(Y) > A$).

   (Moreover, the optimal primal value is achieved in this case.)

2. The primal is **strictly feasible** (there exists $X > 0$ with $\Phi(X) = B$) and the dual is **feasible** (there exists $Y \in \mathrm{Herm}(\mathcal{Y})$ with $\Phi^*(Y) \geq A$).

   (Moreover, the optimal dual value is achieved in this case.)

Another example

Another example of an SDP is as follows. Assume $\rho, \sigma \in \mathrm{D}(\mathbb{C}^n)$ are density operators, for some positive integer n. Let

$$\mathcal{X} = \mathbb{C}^n \oplus \mathbb{C}^n = \mathcal{Y}$$

and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, $A \in \mathrm{Herm}(\mathcal{X})$, and $B \in \mathrm{Herm}(\mathcal{Y})$ be as follows:

$$\Phi\begin{pmatrix} Z & \cdot \\ \cdot & W \end{pmatrix} = \begin{pmatrix} Z & 0 \\ 0 & W \end{pmatrix}, \quad A = \frac{1}{2}\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \rho & 0 \\ 0 & \sigma \end{pmatrix}.$$

After simplification, the primal problem looks like this:

$$\begin{aligned} \text{maximize:} \quad & \frac{1}{2}\mathrm{Tr}(X) + \frac{1}{2}\mathrm{Tr}(X^*) \\ \text{subject to:} \quad & \begin{pmatrix} \rho & X \\ X^* & \sigma \end{pmatrix} \geq 0 \\ & X \in \mathrm{L}(\mathbb{C}^n) \end{aligned}$$

Another example (continued)

The optimal value of the primal problem

$$\begin{aligned} \text{maximize:} \quad & \frac{1}{2}\operatorname{Tr}(X) + \frac{1}{2}\operatorname{Tr}(X^*) \\ \text{subject to:} \quad & \begin{pmatrix} \rho & X \\ X^* & \sigma \end{pmatrix} \geq 0 \\ & X \in \mathrm{L}(\mathbb{C}^n) \end{aligned}$$

from the previous slide is a well-known quantity: it is the **fidelity**

$$\mathrm{F}(\rho, \sigma) = \operatorname{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} = \left\| \sqrt{\rho}\sqrt{\sigma} \right\|_1$$

between ρ and σ. This can be proved using the very useful relationship

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \geq 0 \quad \Longleftrightarrow \quad P, Q \geq 0 \text{ and } X = \sqrt{P}K\sqrt{Q} \text{ for } \|K\| \leq 1.$$

Another example (continued)

Now let us consider the dual problem. Recall that our SDP is given by

$$\Phi\begin{pmatrix} Z & \cdot \\ \cdot & W \end{pmatrix} = \begin{pmatrix} Z & 0 \\ 0 & W \end{pmatrix}, \quad A = \frac{1}{2}\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \rho & 0 \\ 0 & \sigma \end{pmatrix}.$$

The dual problem is given by:

$$\begin{aligned} \text{minimize:} \quad & \langle B, Y \rangle \\ \text{subject to:} \quad & \Phi^*(Y) \geq A \\ & Y \in \mathrm{Herm}(\mathbb{C}^n \oplus \mathbb{C}^n) \end{aligned}$$

After simplification it looks like this:

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\langle \rho, Z \rangle + \frac{1}{2}\langle \sigma, W \rangle \\ \text{subject to:} \quad & \begin{pmatrix} Z & \mathbb{1} \\ \mathbb{1} & W \end{pmatrix} \geq 0 \end{aligned}$$

Another example (continued)

Now let us consider the dual problem. Recall that our SDP is given by

$$\Phi\begin{pmatrix} Z & \cdot \\ \cdot & W \end{pmatrix} = \begin{pmatrix} Z & 0 \\ 0 & W \end{pmatrix}, \quad A = \frac{1}{2}\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \rho & 0 \\ 0 & \sigma \end{pmatrix}.$$

The dual problem is given by:

$$\begin{aligned} \text{minimize:} \quad & \langle B, Y \rangle \\ \text{subject to:} \quad & \Phi^*(Y) \geq A \\ & Y \in \mathrm{Herm}(\mathbb{C}^n \oplus \mathbb{C}^n) \end{aligned}$$

Further simplification:

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\langle \rho, P \rangle + \frac{1}{2}\langle \sigma, P^{-1} \rangle \\ \text{subject to:} \quad & P > 0 \end{aligned}$$

(Equivalent to a well-known expression of the fidelity due to Alberti.)

Alternate forms of SDPs

Semidefinite programs in papers and books do not always take the same form as our general form:

Primal problem		Dual problem	
maximize:	$\langle A, X \rangle$	minimize:	$\langle B, Y \rangle$
subject to:	$\Phi(X) = B$	subject to:	$\Phi^*(Y) \geq A$
	$X \in \mathrm{Pos}(\mathcal{X})$		$Y \in \mathrm{Herm}(\mathcal{Y})$

For example, the primal problem could have an **inequality constraint**:

Primal problem		Dual problem
maximize:	$\langle A, X \rangle$	?
subject to:	$\boxed{\Phi(X) \leq B}$	
	$X \in \mathrm{Pos}(\mathcal{X})$	

Alternate forms of SDPs (continued)

It is not difficult to express an inequality constraint using an equality constraint plus a **slack variable**.

These two conditions are equivalent for any choice of $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ Hermiticity-preserving, $X \in \mathrm{Pos}(\mathcal{X})$, and $B \in \mathrm{Herm}(\mathcal{Y})$:

1. $\Phi(X) \leq B$
2. $\Phi(X) + Z = B$ for some $Z \in \mathrm{Pos}(\mathcal{Y})$.

The equality in the second condition is equivalent to

$$\Psi \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} = B$$

for $\Psi \in \mathrm{T}(\mathcal{X} \oplus \mathcal{Y}, \mathcal{Y})$ defined as

$$\Psi \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} = \Phi(X) + Z.$$

Alternate forms of SDPs (continued)

So, the primal problem with an inequality constraint

$$\begin{aligned} \text{maximize:} \quad & \langle \mathrm{A}, \mathrm{X} \rangle \\ \text{subject to:} \quad & \Phi(\mathrm{X}) \leq \mathrm{B} \\ & \mathrm{X} \in \mathrm{Pos}(\mathcal{X}) \end{aligned}$$

may be expressed using an equality constraint as follows:

$$\begin{aligned} \text{maximize:} \quad & \left\langle \begin{pmatrix} \mathrm{A} & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathrm{X} & \cdot \\ \cdot & \mathrm{Z} \end{pmatrix} \right\rangle \\ \text{subject to:} \quad & \Psi \begin{pmatrix} \mathrm{X} & \cdot \\ \cdot & \mathrm{Z} \end{pmatrix} \triangleq \Phi(\mathrm{X}) + \mathrm{Z} = \mathrm{B} \\ & \begin{pmatrix} \mathrm{X} & \cdot \\ \cdot & \mathrm{Z} \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{Y}) \end{aligned}$$

Alternate forms of SDPs (continued)

For the mapping $\Psi \in \mathrm{T}(\mathcal{X} \oplus \mathcal{Y}, \mathcal{Y})$ defined as

$$\Psi \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} = \Phi(X) + Z,$$

we have that $\Psi^* \in \mathrm{T}(\mathcal{Y}, \mathcal{X} \oplus \mathcal{Y})$ is given by

$$\Psi^*(Y) = \begin{pmatrix} \Phi^*(Y) & 0 \\ 0 & Y \end{pmatrix}.$$

This may be verified as follows:

$$\left\langle Y, \Psi \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} \right\rangle = \langle Y, \Phi(X) \rangle + \langle Y, Z \rangle = \langle \Phi^*(Y), X \rangle + \langle Y, Z \rangle$$

$$= \left\langle \begin{pmatrix} \Phi^*(Y) & 0 \\ 0 & Y \end{pmatrix}, \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} \right\rangle.$$

Alternate forms of SDPs (continued)

The dual problem for our SDP (expressed by an equality constraint plus a slack variable) is as follows:

$$\begin{aligned} \text{minimize:} \quad & \langle \mathrm{B}, \mathrm{Y} \rangle \\ \text{subject to:} \quad & \Psi^*(\mathrm{Y}) \triangleq \begin{pmatrix} \Phi^*(\mathrm{Y}) & 0 \\ 0 & \mathrm{Y} \end{pmatrix} \geq \begin{pmatrix} \mathrm{A} & 0 \\ 0 & 0 \end{pmatrix} \\ & \mathrm{Y} \in \mathrm{Herm}(\mathcal{Y}) \end{aligned}$$

At the end of the day, we obtain this pair of optimization problems:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \langle \mathrm{A}, \mathrm{X} \rangle \\ \text{subject to:} \quad & \Phi(\mathrm{X}) \leq \mathrm{B} \\ & \mathrm{X} \in \mathrm{Pos}(\mathcal{X}) \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \langle \mathrm{B}, \mathrm{Y} \rangle \\ \text{subject to:} \quad & \Phi^*(\mathrm{Y}) \geq \mathrm{A} \\ & \mathrm{Y} \in \mathrm{Pos}(\mathcal{Y}) \end{aligned}$$

Exercise 1: equality to inequality constraints

Show that a semidefinite program in our original form

Primal problem		Dual problem	
maximize:	$\langle \mathrm{A}, \mathrm{X} \rangle$	minimize:	$\langle \mathrm{B}, \mathrm{Y} \rangle$
subject to:	$\Phi(\mathrm{X}) = \mathrm{B}$	subject to:	$\Phi^*(\mathrm{Y}) \geq \mathrm{A}$
	$\mathrm{X} \in \mathrm{Pos}(\mathcal{X})$		$\mathrm{Y} \in \mathrm{Herm}(\mathcal{Y})$

can be expressed in the inequality form

Primal problem		Dual problem	
maximize:	$\langle \mathrm{C}, \mathrm{Z} \rangle$	minimize:	$\langle \mathrm{D}, \mathrm{W} \rangle$
subject to:	$\Psi(\mathrm{Z}) \leq \mathrm{D}$	subject to:	$\Psi^*(\mathrm{W}) \geq \mathrm{C}$
	$\mathrm{Z} \in \mathrm{Pos}(\mathcal{Z})$		$\mathrm{W} \in \mathrm{Pos}(\mathcal{W})$

Hint: $\Phi(\mathrm{X}) = \mathrm{B}$ if and only if $\Phi(\mathrm{X}) \geq \mathrm{B}$ and $-\Phi(\mathrm{X}) \geq -\mathrm{B}$.

Exercise 2: the "standard form" for SDPs

The "standard form" for an SDP is as follows:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \langle A, X \rangle \\ \text{subject to:} \quad & \langle B_1, X \rangle = b_1 \\ & \vdots \\ & \langle B_m, X \rangle = b_m \\ & X \in \mathrm{Pos}(\mathcal{X}) \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \sum_{k=1}^{m} b_k y_k \\ \text{subject to:} \quad & \sum_{k=1}^{m} y_k B_k \geq A \\ & y_1, \ldots, y_m \in \mathbb{R} \end{aligned}$$

Prove that this form is equivalent to the form we have taken as our definition.

Hints:

- One direction of the equivalence is easy if we take $\mathcal{Y} = \mathbb{C}^m$.
- For the other direction, let $m = \dim(\mathcal{Y})^2$ and start by choosing an orthogonal basis $\{H_1, \ldots, H_m\}$ of $\mathrm{Herm}(\mathcal{Y})$.

Algorithms for approximating semidefinite programs

There exist efficient algorithms for approximating optimal solutions of semidefinite programs, provided that they meet certain conditions.

1. **Ellipsoid method:** Not useful in practice, but provides a provably polynomial-time algorithm for a very general class of SDPs (having "well-bounded" feasible sets).
2. **Interior point algorithms:** These algorithms are useful in practice. (The CVX system for MATLAB provides a very nice, easy-to-use implementation.)
3. **Matrix multiplicative weights update method:** this is a "meta-algorithm" that describes certain highly-efficient algorithms for special classes of semidefinite programs.

Not all semidefinite programs can be solved efficiently. Some have exponential-size solutions, and others are related to the notorious sum of square-roots problem.

3. Applications of semidefinite programming in quantum information theory

Optimizing over measurements

Suppose that we have an **ensemble** of states

$$\mathcal{E} = \{(p_1, \sigma_1), \ldots, (p_m, \sigma_m)\},$$

where $(p_1, \ldots, p_m)$ is a probability vector and $\sigma_1, \ldots, \sigma_m \in \mathrm{D}(\mathcal{Y})$ are density operators. We will consider the following task:

1. An index $k \in \{1, \ldots, m\}$ is selected randomly according to the distribution $(p_1, \ldots, p_m)$.
2. We are presented with a quantum system in the state σ_k, and our goal is to identify the correct value of k by means of a measurement on this system.

In other words, we wish to choose a measurement $\mathcal{M} = \{P_1, \ldots, P_m\}$ so as to maximize the quantity

$$\sum_{k=1}^{m} p_k \langle P_k, \sigma_k \rangle.$$

An SDP for optimal measurements

In general (assuming $m \geq 3$), there is no closed-form expression for an optimal measurement.

An optimal measurement does, however, arise as a solution to a semidefinite programming problem:

$$\begin{aligned} \text{maximize:} \quad & p_1 \langle \sigma_1, P_1 \rangle + \cdots + p_m \langle \sigma_m, P_m \rangle \\ \text{subject to:} \quad & P_1 + \cdots + P_m = \mathbb{1} \\ & P_1, \ldots, P_m \in \mathrm{Pos}(\mathcal{Y}) \end{aligned}$$

This is the primal problem corresponding to (Φ, A, B) where

$$\Phi \begin{pmatrix} P_1 & \cdots & \cdot \\ \vdots & \ddots & \vdots \\ \cdot & \cdots & P_m \end{pmatrix} = \sum_{k=1}^{m} P_k, \quad A = \begin{pmatrix} p_1 \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_m \sigma_m \end{pmatrix},$$

and $B = \mathbb{1}$, where we have taken $\mathcal{X} = \mathcal{Y} \oplus \cdots \oplus \mathcal{Y}$ (m times).

An SDP for optimal measurements (continued)

To determine the dual problem, we first observe:

$$\Phi\begin{pmatrix} P_1 & \cdots & \cdot \\ \vdots & \ddots & \vdots \\ \cdot & \cdots & P_m \end{pmatrix} = \sum_{k=1}^{m} P_m \quad \Longrightarrow \quad \Phi^*(Y) = \begin{pmatrix} Y & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Y \end{pmatrix}.$$

The dual problem is as follows:

$$\begin{aligned} \text{minimize:} \quad & \operatorname{Tr}(Y) \\ \text{subject to:} \quad & \begin{pmatrix} Y & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Y \end{pmatrix} \geq \begin{pmatrix} p_1\sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_m\sigma_m \end{pmatrix} \\ & Y \in \operatorname{Herm}(\mathcal{Y}) \end{aligned}$$

An SDP for optimal measurements (continued)

The final (simplified) SDP for optimizing over measurements is as follows:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \sum_{k=1}^{m} p_k \langle \sigma_k, P_k \rangle \\ \text{subject to:} \quad & P_1 + \cdots + P_m = \mathbb{1} \\ & P_1, \ldots, P_m \in \mathrm{Pos}(\mathcal{Y}) \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \mathrm{Tr}(Y) \\ \text{subject to:} \quad & Y \geq p_k \sigma_k \quad (k = 1, \ldots, m) \\ & Y \in \mathrm{Herm}(\mathcal{Y}) \end{aligned}$$

Complementary slackness

We may say something more about optimal measurements using a property known as complementary slackness.

Complementary slackness: Suppose (Φ, A, B) is a semidefinite program, X is a primal feasible point, Y is a dual feasible point, and $\langle A, X\rangle = \langle B, Y\rangle$. It must hold that

$$\Phi^*(Y)X = AX. \tag{1}$$

Proof. We have

$$\langle \Phi^*(Y), X\rangle = \langle Y, \Phi(X)\rangle = \langle Y, B\rangle = \langle A, X\rangle,$$

so

$$\langle \Phi^*(Y) - A, X\rangle = 0.$$

If $\langle P, Q\rangle = 0$ for $P, Q \geq 0$, then $PQ = 0$, and therefore (1) holds. □

Optimal measurement complementary slackness

Suppose we have an **optimal** measurement $\{P_1, \ldots, P_m\}$ for our state distinguishability problem for ensemble $\mathcal{E} = \{(p_1, \sigma_1), \ldots, (p_m, \sigma_m)\}$.

Let Y be any optimal solution to the dual problem. By the Slater conditions, Y must exist, and $\mathrm{Tr}(Y)$ must be equal to the optimal value of the primal problem.

By complementary slackness:

$$\begin{pmatrix} YP_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & YP_m \end{pmatrix} = \begin{pmatrix} p_1\sigma_1P_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_m\sigma_mP_m \end{pmatrix}$$

Summing the diagonal entries yields:

$$Y = \sum_{k=1}^{m} p_k\sigma_kP_k.$$

Holevo's criterion for measurement optimality

It turns out that Y is uniquely determined by $\{P_1, \ldots, P_m\}$:

$$Y = \sum_{k=1}^{m} p_k \sigma_k P_k.$$

Also note that

$$\mathrm{Tr}(Y) = \mathrm{Tr}\left(\sum_{k=1}^{m} p_k \sigma_k P_k\right) = \sum_{k=1}^{m} p_k \langle \sigma_k, P_k \rangle,$$

which is the primal objective value obtained by $\{P_1, \ldots, P_m\}$.

We obtain the **Holevo criterion for measurement optimality**: $\{P_1, \ldots, P_m\}$ is optimal for $\mathcal{E} = \{(p_1, \sigma_1), \ldots, (p_m, \sigma_m)\}$ if and only if

$$\sum_{k=1}^{m} p_k \sigma_k P_k \geq p_j \sigma_j \quad \text{(for } j = 1, \ldots, m\text{)}.$$

Optimizing over channels

A linear mapping of the form

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

for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, is a **quantum channel** if it is both **completely positive** and **trace preserving**. The transformation $\rho \mapsto \Phi(\rho)$ may then be considered physically realizable (in principle).

There are several common ways that one may represent such a mapping. Examples include Kraus representations and Stinespring representations.

The **Choi operator** of Φ is another such representation:

$$\mathrm{J}(\Phi) = \sum_{a,b \in \Sigma} \Phi\big(|a\rangle\langle b|\big) \otimes |a\rangle\langle b| \in \mathrm{L}(\mathcal{Y} \otimes \mathcal{X}).$$

(Here it is assumed that $\{|a\rangle \,:\, a \in \Sigma\}$ is the standard basis of $\mathcal{X}$.)

Properties of Choi operators

There is a one-to-one and onto correspondence between linear mappings of the form $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ and their corresponding Choi operators $\mathrm{J}(\Phi) \in \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$.

An operator $\mathrm{J}(\Phi) \in \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$ is the Choi operator of a quantum channel if and only if

1. $\mathrm{J}(\Phi)$ is positive semidefinite, and
2. $\mathrm{Tr}_{\mathcal{Y}}(\mathrm{J}(\Phi)) = \mathbb{1}_{\mathcal{X}}$.

For any linear function φ mapping quantum channels to real numbers, there must exist a Hermitian operator A such that

$$\varphi(\Phi) = \langle A, \mathrm{J}(\Phi) \rangle$$

for every channel $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$.

These observations make it possible to represent any optimization of a real-valued linear function over all quantum channels as an SDP.

Wiesner's quantum money scheme

Wiesner's quantum money protocol was published in 1983 (in SIGACT News), but was over 10 years old at that time...

Preparation phase: a mint chooses n single-qubit states, each one selected independently and uniformly at random from the set

$$|0\rangle, \quad |1\rangle, \quad |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

Qubits $X_1, \ldots, X_n$ are initialized to these states and placed on a bank note, together with a unique serial number. A classical record of the states is privately distributed to banks.

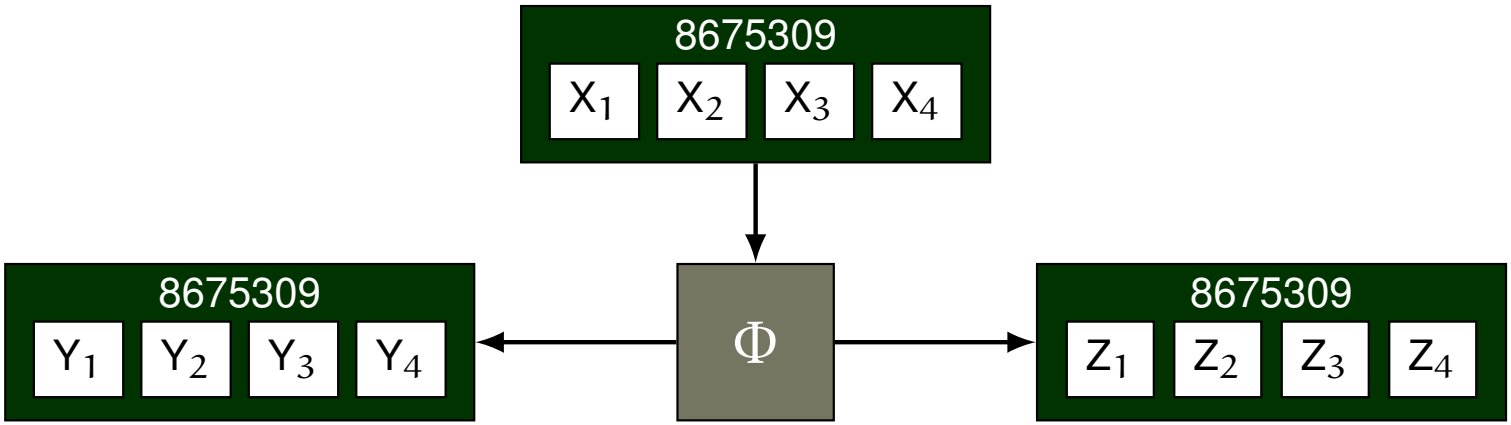
Verification phase: a bank verifies a bank note by measuring the note's qubits $X_1, \ldots, X_n$ against the known pure states selected in the preparation phase. If all (or most) of the measurements agree with the recorded states, the note is declared **valid**, otherwise it is **invalid**.

Attacks against the scheme

One can consider a variety of attacks against this scheme. . .

For instance, an attacker could try to learn the states corresponding to a given serial number by repeated verification attempts with a bank. (A brilliant attack of this sort was recently found by Nagaj and Sattath.)

Suppose that we consider only **simple counterfeiting schemes**: a counterfeiter with one copy of a given bank note attempts to produce two copies, aiming to maximize the probability that both are successfully verified.



Single-qubit counterfeiting

We may begin by considering the $n = 1$ (or single-qubit) version of the scheme. Assume $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ represent the input and two output spaces, respectively (each representing one qubit).

The probability that a given quantum channel $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y} \otimes \mathcal{Z})$ succeeds in simple counterfeiting is

$$\frac{1}{4} \sum_{k=1}^{4} \langle \psi_k \otimes \psi_k | \, \Phi\big(|\psi_k\rangle\langle\psi_k|\big) \, |\psi_k \otimes \psi_k\rangle$$

for $|\psi_1\rangle = |0\rangle$, $|\psi_2\rangle = |1\rangle$, $|\psi_3\rangle = |+\rangle$, and $|\psi_4\rangle = |-\rangle$.

An equivalent expression is $\langle Q, \mathrm{J}(\Phi) \rangle$ for

$$Q = \frac{1}{4} \sum_{k=1}^{4} \big|\psi_k \otimes \psi_k \otimes \overline{\psi_k}\big\rangle \big\langle \psi_k \otimes \psi_k \otimes \overline{\psi_k}\big|$$

and $\mathrm{J}(\Phi)$ being the Choi operator associated with Φ.

An SDP for single-qubit counterfeiting

We obtain the following semidefinite program representing the optimal simple counterfeiting probability:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \langle Q, X \rangle \\ \text{subject to:} \quad & \mathrm{Tr}_{\mathcal{Y}\otimes\mathcal{Z}}(X) = \mathbb{1}_{\mathcal{X}} \\ & X \in \mathrm{Pos}(\mathcal{Y}\otimes\mathcal{Z}\otimes\mathcal{X}) \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \mathrm{Tr}(Y) \\ \text{subject to:} \quad & \mathbb{1}_{\mathcal{Y}\otimes\mathcal{Z}} \otimes Y \geq Q \\ & Y \in \mathrm{Herm}(\mathcal{X}) \end{aligned}$$

where (as before)

$$Q = \frac{1}{4} \sum_{k=1}^{4} \left|\psi_k \otimes \psi_k \otimes \overline{\psi_k}\right\rangle \left\langle \psi_k \otimes \psi_k \otimes \overline{\psi_k}\right| .$$

(Exercise: verify that the dual problem above is indeed the correct dual problem corresponding to the primal problem.)

Optimal value of the SDP

The optimal value of this SDP (for both the primal and dual) is 3/4, as will now be argued.

A **primal solution** achieving the value $3/4$ is given by $X = J(\Phi)$ for

$$\Phi(X) = A_0 X A_0^* + A_1 X A_1^*$$

for

$$A_0 = \frac{1}{\sqrt{12}} \begin{pmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad A_1 = \frac{1}{\sqrt{12}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 3 \end{pmatrix}$$

A **dual solution** achieving the value $3/4$ is given by $Y = \frac{3}{8}\mathbb{1}_{\mathcal{X}}$. The feasibility of this solution follows from a computation $\lambda_1(Q) = 3/8$.

By weak duality, the value $3/4$ is therefore **optimal**.

The general (multiple qubit) case

Next we may ask: what happens in the general (n qubit) case? It is reasonable to guess that $(3/4)^n$ is optimal, but this requires a proof: **a counterfeiter may not treat the individual qubits independently**.

To verify that the bound $(3/4)^n$ is correct, consider the dual problem for an SDP for the n-qubit scheme:

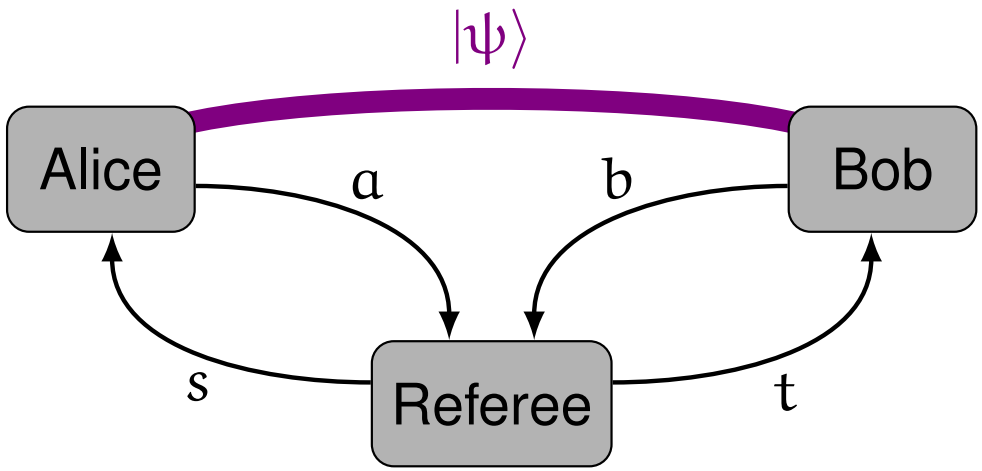
$$\begin{aligned} \text{minimize:} \quad & \mathrm{Tr}(Y) \\ \text{subject to:} \quad & \mathbb{1} \otimes Y \geq W Q^{\otimes n} W^* \\ & Y \in \mathrm{Herm}\big(\mathcal{X}^{\otimes n}\big) \end{aligned}$$

(The identity is on $\mathcal{Y}^{\otimes n} \otimes \mathcal{Z}^{\otimes n}$ and W represents a permutation of tensor factors.)

If Y_0 was dual-feasible for the single-qubit primal problem, then $Y = Y_0^{\otimes n}$ will be feasible for the n-qubit dual problem. It follows that $(3/4)^n$ is an upper bound on the counterfeiting probability.

Nonlocal games

A **nonlocal game** is a cooperative game played by two players (**Alice** and **Bob**) and run by a **Referee**.



1. Referee randomly selects questions: $s \in S$ for Alice, $t \in T$ for Bob.
2. Alice responds with $a \in A$, Bob responds with $b \in B$.
3. Referee evaluates some fixed predicate on (s, t, a, b) to determine one of two outcomes: **Alice and Bob win** or **Alice and Bob lose**.

Alice and Bob may not communicate once the game begins (but may share entanglement in the quantum setting).

XOR games

An **XOR game** is a nonlocal game in which $A = B = \{0, 1\}$, and where Alice and Bob win if and only if

$$a \oplus b = f(s, t)$$

for some function $f : S \times T \to \{0, 1\}$.

One of the most famous nonlocal games, the **CHSH game**, is an example of an XOR game:

1. The referee chooses $s, t \in \{0, 1\}$ uniformly at random.
2. Alice and Bob respond with $a, b \in \{0, 1\}$.
3. They win if and only if $a \oplus b = s \wedge t$.

Classically (i.e., without entanglement) Alice and Bob can win with probability at most 3/4. Using entanglement they can win with probability $\cos^2(\pi/8) \approx 0.85$.

XOR game strategies and values

Consider any strategy for Alice and Bob in an XOR game:

1. A shared entangled state $|\psi\rangle \in \mathcal{A} \otimes \mathcal{B}$.
2. A choice of binary-valued (and w.l.o.g. projective) measurements:

$$\left\{\Pi_0^s, \Pi_1^s\right\}_{s\in S} \text{ for Alice,} \qquad \left\{\Pi_0^t, \Pi_1^t\right\}_{t\in T} \text{ for Bob.}$$

The probability they output (a, b) on questions (s, t) is therefore

$$\left\langle \psi \,\middle|\, \Pi_a^s \otimes \Pi_b^t \,\middle|\, \psi \right\rangle .$$

The probability that such a strategy wins an XOR game defined by the function $f : S \times T \to \{0, 1\}$ is

$$\frac{1}{2} + \frac{1}{2} \sum_{s,t} \pi(s, t)(-1)^{f(s,t)} \left\langle \psi \,\middle|\, (\Pi_0^s - \Pi_1^s) \otimes (\Pi_0^t - \Pi_1^t) \,\middle|\, \psi \right\rangle .$$

(Follows from a calculation of the probability of winning minus losing.)

Tsirelson's correspondence

By the previous expression for the winning probability,

$$\frac{1}{2} + \frac{1}{2} \sum_{s,t} \pi(s,t)(-1)^{f(s,t)} \left\langle \psi \right| (\Pi_0^s - \Pi_1^s) \otimes (\Pi_0^t - \Pi_1^t) \left| \psi \right\rangle ,$$

one has that the winning probability for any strategy is determined by the values

$$\left\langle \psi \right| (\Pi_0^s - \Pi_1^s) \otimes (\Pi_0^t - \Pi_1^t) \left| \psi \right\rangle$$

ranging over all $s \in S$ and $t \in T$.

It is not difficult to prove that there must necessarily exist collections of real unit vectors $\{u_s : s \in S\}$ and $\{v_t : t \in T\}$ such that

$$\langle u_s, v_t \rangle = \left\langle \psi \right| (\Pi_0^s - \Pi_1^s) \otimes (\Pi_0^t - \Pi_1^t) \left| \psi \right\rangle .$$

Tsirelson's correspondence establishes that this is also a sufficient condition.

Tsirelson's correspondence

That is, given **any two collections** of real unit vectors

$$\{u_s : s \in S\} \quad \text{and} \quad \{v_t : t \in T\},$$

there must exist

1. a shared entangled state $|\psi\rangle \in \mathcal{A} \otimes \mathcal{B}$, and
2. a choice of binary-valued measurements: $\left\{\Pi_0^s, \Pi_1^s\right\}_{s\in S}$ for Alice, $\left\{\Pi_0^t, \Pi_1^t\right\}_{t\in T}$ for Bob,

such that the same relation as the previous slide is satisfied:

$$\langle u_s, v_t\rangle = \left\langle \psi \,\middle|\, (\Pi_0^s - \Pi_1^s) \otimes (\Pi_0^t - \Pi_1^t) \,\middle|\, \psi \right\rangle .$$

The proof is based on the Weyl-Brauer matrices:

$$\sigma_z \otimes \cdots \otimes \sigma_z \otimes \sigma_x \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}$$
$$\sigma_z \otimes \cdots \otimes \sigma_z \otimes \sigma_y \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}$$

From Tsirelson's correspondence to SDPs

Next, we will use Tsirelson's correspondence to represent the optimal winning probability of an XOR game as an SDP. More precisely, we will describe an SDP for the maximum value of the expression

$$\sum_{s,t} \pi(s,t)(-1)^{f(s,t)} \langle u_s, v_t \rangle,$$

over all collections of real unit vectors $\{u_s : s \in S\}$ and $\{v_t : t \in T\}$.

The first step is to define a matrix C (indexed by $S \times T$) as

$$C(s,t) = \pi(s,t)(-1)^{f(s,t)},$$

and an operator $A \in \mathrm{Herm}(\mathbb{C}^S \oplus \mathbb{C}^T)$ as

$$A = \frac{1}{2} \begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix}.$$

From Tsirelson's correspondence to SDPs

Next, for a given collection of real unit vectors, consider the matrix X, indexed by the disjoint union $S \cup T$, as follows:

$$X = \begin{array}{c} \\ s_0 \in S \\ \\ t_0 \in T \end{array} \overbrace{\begin{pmatrix} \langle u_{s_0}, u_{s_1} \rangle & \langle u_{s_0}, v_{t_1} \rangle \\ \langle v_{t_0}, u_{s_1} \rangle & \langle v_{t_0}, v_{t_1} \rangle \end{pmatrix}}^{s_1 \in S \quad t_1 \in T}$$

In words, this is the **Gram matrix** of the collection of vectors $\{u_s : s \in S\} \cup \{v_t : t \in T\}$. It is necessarily positive semidefinite, has real entries, and diagonal entries equal to 1.

Conversely, any matrix with these properties must be obtained from a collection of real unit vectors $\{u_s : s \in S\} \cup \{v_t : t \in T\}$ in this way.

Semidefinite programs for XOR games

We may conclude that the optimal value of the following SDP determines the optimal winning probability of a given XOR game:

Primal problem		Dual problem	
maximize:	$\langle A, X \rangle$	minimize:	$\mathrm{Tr}(Y)$
subject to:	$\Delta(X) = \mathbb{1}$	subject to:	$\Delta(Y) \geq A$
	$X \in \mathrm{Pos}\big(\mathbb{C}^S \oplus \mathbb{C}^T\big)$		$Y \in \mathrm{Herm}\big(\mathbb{C}^S \oplus \mathbb{C}^T\big)$

The operator

$$A = \frac{1}{2}\begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix} \in \mathrm{Herm}\big(\mathbb{C}^S \oplus \mathbb{C}^T\big)$$

is determined by the game (as described before), and the mapping Δ is the **completely dephasing channel**, which zeroes out all off-diagonal entries and leaves diagonal entries unchanged.

Parallel repetition of XOR games

The problem of **parallel repetition** concerns the relationship between the optimal winning probability of a given nonlocal game G as compared to G^n, defined as follows:

- The referee behaves as if he is playing n **independent copies** of G with Alice and Bob simultaneously: he sends Alice questions $(s_1, \ldots, s_n)$ and Bob $(t_1, \ldots, t_n)$.
- Alice and Bob send responses $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$. They may choose to **correlate** the different copies of the games.
- Alice and Bob win G^n if and only if they win every one of the iterations of G.

The optimal probability $\omega(G^n)$ to win G^n is larger than $\omega(G)^n$ for some nonlocal games (such as the **Fortnow-Feige-Lovász game**).

Cleve, Slofstra, Unger, and Upadhyay proved that $\omega(G^n) = \omega(G)^n$ (for quantum strategies) for every XOR game G using the SDP formulation just described.

The trace norm and state distinguishability

The **trace norm** of an operator $X \in \mathrm{L}(\mathcal{X})$ is defined as

$$\|X\|_1 = \mathrm{Tr}\,\sqrt{X^*X}.$$

It is commonly used in the theory of quantum information as a measure of **distinguishability** between states.

Theorem (Holevo 1973, Helstrom 1976).
The minimum error probability to correctly distinguish quantum states ρ_0 and ρ_1 by means of a measurement, assuming they are given with probabilities λ and $1-\lambda$, respectively, is

$$\frac{1}{2} - \frac{1}{2}\|\lambda\rho_0 - (1-\lambda)\rho_1\|_1.$$

(This theorem may be proved using the SDP formulation of state distinguishability we discussed previously, although there are more direct proofs.)

An SDP for the trace norm

Incidentally, there is a simple SDP for the trace norm of an arbitrary operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \frac{1}{2}\langle A, X\rangle + \frac{1}{2}\langle A^*, X^*\rangle \\ \text{subject to:} \quad & \begin{pmatrix} \mathbb{1}_{\mathcal{X}} & X^* \\ X & \mathbb{1}_{\mathcal{Y}} \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{Y}) \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\,\mathrm{Tr}(Y) + \frac{1}{2}\,\mathrm{Tr}(Z) \\ \text{subject to:} \quad & \begin{pmatrix} Y & A^* \\ A & Z \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{Y}) \end{aligned}$$

The completely bounded trace norm

There is an analogous norm for linear mappings of the form

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

called the **completely bounded trace norm** (and also commonly called the **diamond norm**).

To define this norm, we consider the norm induced by the trace norm:

$$\|\Phi\|_1 = \max\left\{\|\Phi(X)\|_1 \,:\, X \in \mathrm{L}(\mathcal{X}),\ \|X\|_1 \leq 1\right\}.$$

The **completely bounded trace norm** is now defined as

$$|||\Phi|||_1 = \sup_{k \geq 1} \left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_1 = \left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\right\|_1.$$

(Other notations include $\|\Phi\|_{\mathrm{cb},1}$ and $\|\Phi\|_\diamond$.)

Quantum channel distinguishability

In the problem of **quantum channel distinguishability**, two channels

$$\Phi_0, \Phi_1 : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$$

are fixed.

A single evaluation of one of the two channels is made available. With probability λ the given channel is Φ_0 and with probability $1 - \lambda$ it is Φ_1.

The goal is to identify which channel was given by a procedure as follows:

1. A quantum state of the form $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{W})$ is prepared.
2. The given channel is applied to $\mathcal{X}$, resulting in the state

$$\sigma_0 = \big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\rho) \quad \text{or} \quad \sigma_1 = \big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\rho).$$

3. The states σ_0 and σ_1 are distinguished by a measurement.

Optimal quantum channel distinguishability

The minimum error probability to distinguish the outcomes is

$$\frac{1}{2} - \frac{1}{2}\left\| \lambda\big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\rho) - (1-\lambda)\big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\rho) \right\|_1$$

Optimizing over all choices of $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{W})$ gives a quantum channel analogue to the Holevo-Helstrom theorem.

Theorem. The minimum error probability to correctly distinguish channels Φ_0 and Φ_1 given with probabilities λ and $1-\lambda$, respectively, is

$$\frac{1}{2} - \frac{1}{2}\left\| \lambda\Phi_0 - (1-\lambda)\Phi_1 \right\|_1.$$

(There are simple examples, such as the so-called Werner-Holevo channels, that demonstrate that the auxiliary space $\mathcal{W}$ is necessary for an optimal discrimination of channels Φ_0 and Φ_1.)

Computing the completely-bounded trace norm

There is no closed-form expression known for the completely bounded trace norm of a given map. It may, however, be represented as the optimal value of a semidefinite program. (I am aware of three different ways to do this.)

Suppose hereafter that

$$\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$$

is any linear mapping, and that we are interested in computing $|||\Phi|||_1$.

One may write Φ in its (generalized) Stinespring form as

$$\Phi(\mathrm{X}) = \mathrm{Tr}_{\mathcal{Z}}\left(\mathrm{A}\mathrm{X}\mathrm{B}^*\right)$$

for all $\mathrm{X} \in \mathrm{L}(\mathcal{X})$, where $\mathrm{A}, \mathrm{B} \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ are operators and $\mathcal{Z}$ is a sufficiently large complex Euclidean space.

(It is always sufficient to take $\dim(\mathcal{Z}) = \dim(\mathcal{X} \otimes \mathcal{Y})$.)

Computing the completely-bounded trace norm

From the expression

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(AXB^*\big)$$

we define two new mappings $\Psi_A, \Psi_B : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Z})$ as

$$\Psi_A(X) = \mathrm{Tr}_{\mathcal{Y}}\big(AXA^*\big) \quad \text{and} \quad \Psi_B(X) = \mathrm{Tr}_{\mathcal{Y}}\big(BXB^*\big)$$

for all $X \in \mathrm{L}(\mathcal{X})$. (Note that $\mathcal{Y}$ is traced out rather than $\mathcal{Z}$.)

The mappings Ψ_A and Ψ_B are completely positive (but not necessarily trace-preserving) linear maps.

The following relation is not difficult to prove:

$$\|\!|\Phi|\!\|_1 = \max\big\{F(\Psi_A(\rho_A), \Psi_B(\rho_B)) \,:\, \rho_A, \rho_B \in \mathrm{D}(\mathcal{X})\big\}. \qquad (2)$$

Here it is assumed that the fidelity function is extended to arbitrary positive semidefinite operators.

(The formula (2) appears as an exercise in Kitaev, Shen, and Vyalyi.)

Computing the completely-bounded trace norm

Using the SDP for the fidelity (discussed earlier in the talk), one obtains an SDP for the completely bounded trace norm:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \frac{1}{2}\operatorname{Tr}(X) + \frac{1}{2}\operatorname{Tr}(X^*) \\ \text{subject to:} \quad & \begin{pmatrix} \Psi_A(\rho_A) & X \\ X^* & \Psi_B(\rho_B) \end{pmatrix} \geq 0 \\ & \rho_A, \rho_B \in \mathrm{D}(\mathcal{X}),\ X \in \mathrm{L}(\mathcal{Z}). \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\left\|\Psi_A^*(Z)\right\| + \frac{1}{2}\left\|\Psi_B^*\left(Z^{-1}\right)\right\| \\ \text{subject to:} \quad & Z \in \mathrm{Pos}(\mathcal{Z}),\ Z > 0. \end{aligned}$$

A theorem due to R. Smith

Along similar lines to the dual SDP for the fidelity function being equivalent to Alberti's theorem, the dual problem

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\left\|\Psi_A^*(Z)\right\| + \frac{1}{2}\left\|\Psi_B^*\left(Z^{-1}\right)\right\| \\ \text{subject to:} \quad & Z \in \mathrm{Pos}(\mathcal{Z}),\ Z > 0 \end{aligned}$$

is equivalent to a theorem of R. Smith about the completely bounded trace norm (proved in 1983).

Theorem (Smith). For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ it holds that

$$\left\|\left\|\Phi\right\|\right\|_1^2 = \inf_{(A,B)\in S_\Phi} \|A\|\|B\|$$

where the infimum is over all $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ satisfying

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\left(AXB^*\right).$$

Alternate SDP for the CB trace norm

Here is a different SDP for $\|\!|\!|\Phi|\!|\!|_1$ based on the Choi operator of Φ:

Primal problem

$$\text{maximize:} \quad \frac{1}{2}\langle J(\Phi), X\rangle + \frac{1}{2}\langle J(\Phi)^*, X^*\rangle$$

$$\text{subject to:} \quad \begin{pmatrix} \mathbb{1}_{\mathcal{Y}} \otimes \rho_0 & X \\ X^* & \mathbb{1}_{\mathcal{Y}} \otimes \rho_1 \end{pmatrix} \geq 0$$

$$\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X}), \; X \in \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$$

Dual problem

$$\text{minimize:} \quad \frac{1}{2}\left\| \mathrm{Tr}_{\mathcal{Y}}(Y_0) \right\| + \frac{1}{2}\left\| \mathrm{Tr}_{\mathcal{Y}}(Y_1) \right\|$$

$$\text{subject to:} \quad \begin{pmatrix} Y_0 & J(\Phi) \\ J(\Phi)^* & Y_1 \end{pmatrix} \geq 0$$

$$Y_0, Y_1 \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$$

Basic objects of quantum information

It is reasonable to say that the most basic objects of quantum information theory are:

1. **States:** represented by density operators.
2. **Measurements:** represented by sets of measurement operators (or POVM elements).
3. **Channels:** represented by completely positive, trace-preserving maps.

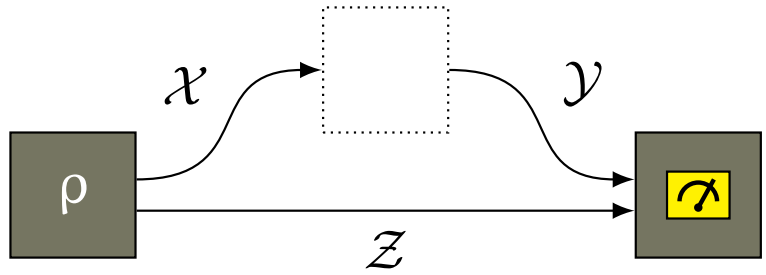
As we have observed, these objects can be represented by **positive semidefinite operators** subject to **linear constraints**.

Based on this observation, it should not be a surprise that semidefinite programming should be applicable to quantum information theory...

In fact, there is an interesting general family of objects, formed by composing these basic objects, that may be represented by semidefinite operators subject to linear constraints.

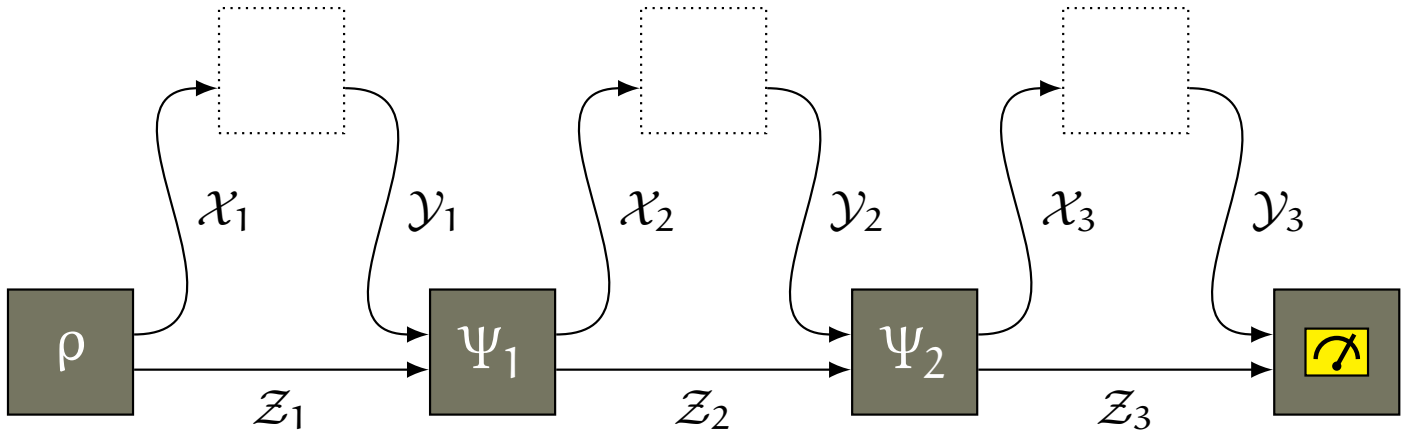
Interactive processes

Here is an example of such an object:



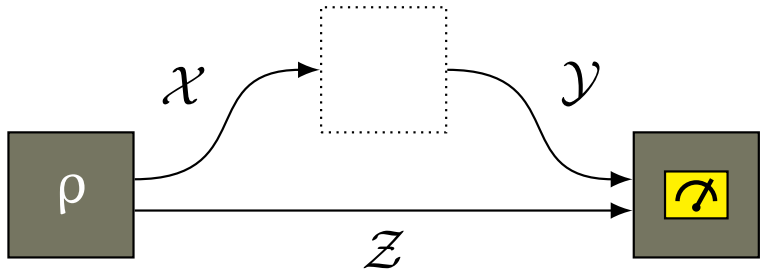
We may view this as an **interactive measurement**. It is essentially a measurement <u>of</u> a given channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$.

We may generalize such objects to **multiple-round interactions**:



Interactive measurements

Consider first the example of an interactive measurement:



Claim: We can represent this interactive measurement by a collection of positive semidefinite operators $\{Q_a\} \subset \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$.

- If we "interface" the interactive measurement with a channel Φ, each outcome $a \in \Sigma$ appears with probability $p(a) = \langle Q_a, J(\Phi) \rangle$.
- The collections $\{Q_a\}$ that arise in this way are characterized by a set of linear constraints (plus the positive semidefinite constraint). In particular, there must exist $\sigma \in \mathrm{D}(\mathcal{X})$ such that

$$\sum_a Q_a = \mathbb{1}_{\mathcal{Y}} \otimes \sigma.$$

An SDP for distinguishing mappings

For a given **ensemble of channels**, or more generally a **quantum instrument** $\{\Phi_1 \ \ldots, \Phi_m\}$, we obtain a semidefinite program for optimally distinguishing these mappings:

Primal problem

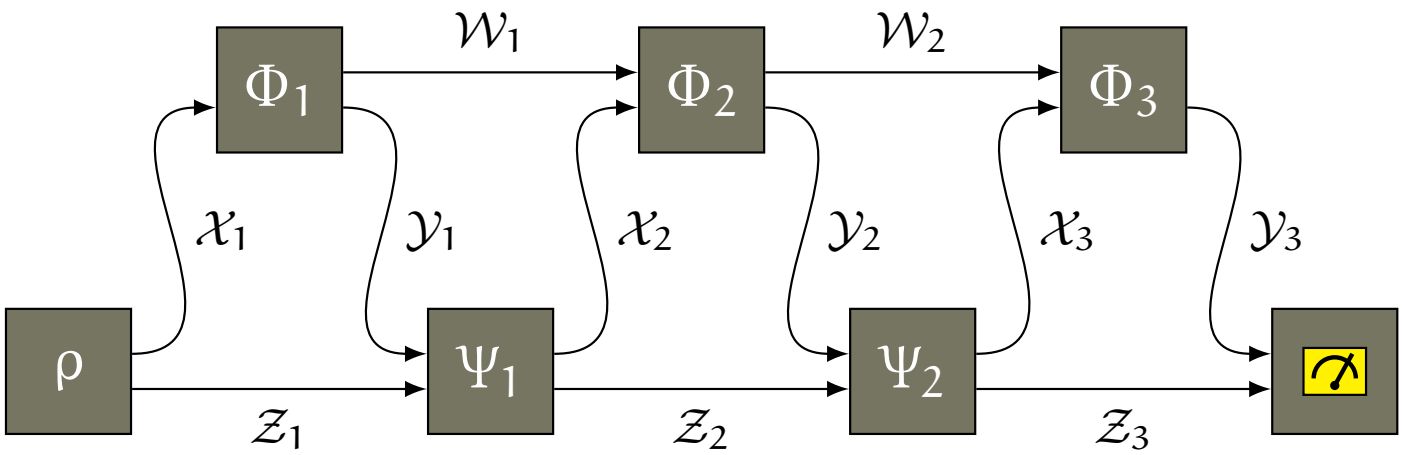
$$\begin{aligned} \text{maximize:} \quad & \langle J(\Phi_1), X_1 \rangle + \cdots + \langle J(\Phi_m), X_m \rangle \\ \text{subject to:} \quad & X_1 + \cdots + X_m = \mathbb{1}_{\mathcal{Y}} \otimes \sigma \\ & X_1, \ldots, X_m \subset \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X}), \ \sigma \in \mathrm{D}(\mathcal{X}) \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \| \mathrm{Tr}_{\mathcal{Y}}(Y) \| \\ \text{subject to:} \quad & Y \geq J(\Phi_k) \ \ (k = 1, \ldots, m) \\ & Y \in \mathrm{Herm}(\mathcal{Y} \otimes \mathcal{X}) \end{aligned}$$

Multiple-round interactions

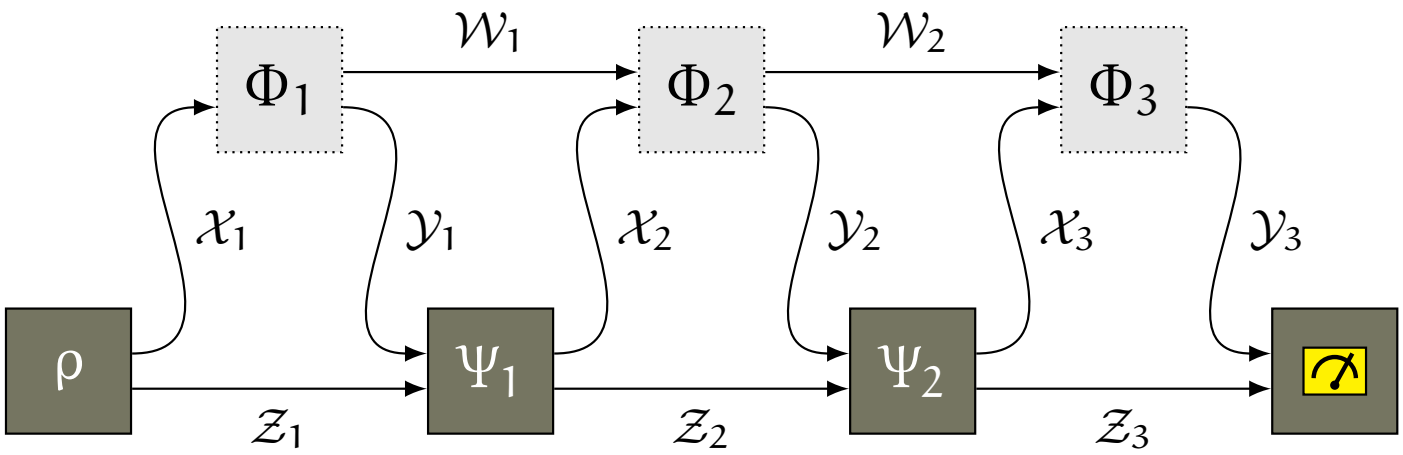
More generally, one may consider multiple-round processes, such as this one:



Such a process is expected to be interfaced with a **compatible** process.

An example of a question that one may consider: for a particular measurement outcome in the picture above, what is the highest probability with which channels Φ_1, Φ_2, and Φ_3 can cause this measurement outcome to appear?

An SDP for maximum output probability



For Q determined from ρ, Ψ_1, Ψ_2, and a chosen measurement operator, the maximum output probability is represented by this problem:

maximize: $\langle Q, X_3 \rangle$

subject to:

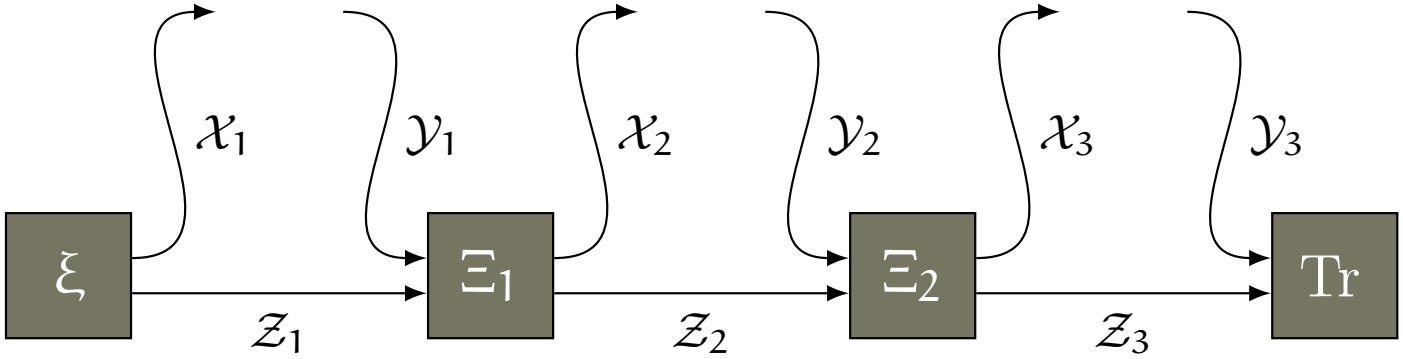
$$\begin{aligned}
\mathrm{Tr}_{\mathcal{Y}_3}(X_3) &= X_2 \otimes \mathbb{1}_{\mathcal{X}_3}, & X_3 &\in \mathrm{Pos}(\mathcal{Y}_1 \otimes \mathcal{Y}_2 \otimes \mathcal{Y}_3 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_3), \\
\mathrm{Tr}_{\mathcal{Y}_2}(X_2) &= X_1 \otimes \mathbb{1}_{\mathcal{X}_2}, & X_2 &\in \mathrm{Pos}(\mathcal{Y}_1 \otimes \mathcal{Y}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2), \\
\mathrm{Tr}_{\mathcal{Y}_1}(X_1) &= \mathbb{1}_{\mathcal{X}_1}, & X_1 &\in \mathrm{Pos}(\mathcal{Y}_1 \otimes \mathcal{X}_1).
\end{aligned}$$

Dual form of maximum output probability SDP

The dual form of the SDP described on the previous slide has is remarkably simple:

$$\begin{aligned} \text{minimize:} \quad & \lambda \\ \text{subject to:} \quad & \mathrm{Q} \leq \lambda \mathrm{R} \end{aligned}$$

for R ranging over all operators $\mathrm{R} \in \mathrm{Pos}(\mathcal{Y}_1 \otimes \mathcal{Y}_2 \otimes \mathcal{Y}_3 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_3)$ that arise from a similar representation to Q, for processes of this form:



From this fact, we may obtain a very simple proof of Kitaev's bound on strong quantum coin-flipping...

Strong quantum coin-flipping

A **strong quantum coin-flipping protocol** with bias ε is an interaction between two (honest) players Alice and Bob, both having output sets $\{0, 1, \text{abort}\}$.

Required properties:

1. The interaction between the honest players Alice and Bob produces the same outcome $b \in \{0, 1\}$ for both players, with probability $1/2$ for each outcome.
2. If one of the players does not follow the protocol but the other does, neither of the outcomes $b \in \{0, 1\}$ is output by the honest player with probability greater than $1/2 + \varepsilon$.

Theorem (Kitaev): All strong quantum coin-flipping protocols have bias at least

$$\frac{1}{\sqrt{2}} - \frac{1}{2} \approx 0.207.$$

Simple proof of Kitaev's coin-flipping bound

Consider any quantum coin-flipping protocol (any number of rounds). Let the honest strategies for Alice and Bob be represented by operators

$$\{A_0, A_1, A_{\text{abort}}\} \quad \text{and} \quad \{B_0, B_1, B_{\text{abort}}\}.$$

Let p be the maximum probability a cheating Bob can force outcome 0 for Alice. This implies that there must be a (cheating) strategy R for Alice such that $A_0 \leq pR$.

The strategy R would cause Bob to output 0 with probability

$$\langle R, B_0 \rangle \geq \frac{1}{p} \langle A_0, B_0 \rangle = \frac{1}{2p}.$$

For $p > 0$ we have

$$\max\left\{p, \frac{1}{2p}\right\} \geq \frac{1}{\sqrt{2}};$$

one of the players can force outcome 0 with probability at least $\frac{1}{\sqrt{2}}$. □

Other applications of SDPs in quantum information

There are many other topics in quantum information and computation where SDPs find applications. A few among many examples:

1. Quantum query complexity and the general adversary bound.
2. Quantum interactive proof systems (QIP = PSPACE, QRG = EXP, and many other known facts are based on SDPs).
3. Min- and max-entropy for quantum states, smoothed versions, and other variants are closely connected with SDPs.
4. A simple proof of the quantum substate theorem.
5. SDP hierarchies for separability and nonlocal correlations.
6. Non-commutative graphs and a quantum Lovász theta function.

Various objects of interest, such as symmetric extensions, PPT measurements and channels, steering witnesses, cloning channels, and a variety of norms, are naturally represented by SDPs.

The End

Thank you for your attention.

Primal problem		Dual problem	
maximize:	$\langle A, X \rangle$	minimize:	$\langle B, Y \rangle$
subject to:	$\Phi(X) = B$	subject to:	$\Phi^*(Y) \geq A$
	$X \in \mathrm{Pos}(\mathcal{X})$		$Y \in \mathrm{Herm}(\mathcal{Y})$