Semidefinite programs for completely bounded norms

John Watrous

School of Computer Science
and
Institute for Quantum Computing

University of Waterloo

The trace norm and state distinguishability

Throughout this talk, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ and $\mathcal{W}$ are vector spaces of the form $\mathbb{C}^n$ (for possibly different $n \geqslant 1$).

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

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

It is commonly used in the theory of quantum information because it describes how well two given quantum states can be **distinguished** by means of a **measurement**.

Theorem (Holevo 1973, Helstrom 1976)

The minimum error probability to correctly distinguish two quantum states ρ_0 and ρ_1 by means of a measurement, assuming they are given with probabilities λ and $1 - \lambda$, respectively, is

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

The completely bounded trace norm

There is an analogous norm for 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 first consider the norm induced by the trace norm:

$$\| \Phi \|_1 \stackrel{\text{def}}{=} \max \{ \| \Phi(X) \|_1 \, : \, X \in \mathrm{L}(\mathcal{X}), \; \|X\|_1 \leqslant 1 \}.$$

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

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

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

Quantum channel distinguishability

In the problem of **quantum channel distinguishability**, two quantum channels (or completely positive and trace preserving maps)

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

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 correctly identify which channel was given, using an “interactive measurement”.

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 one of the states

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

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

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

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

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

The completely bounded (spectral) norm

The same process can be attempted for any other operator norm. In particular, we may consider the **spectral norm**

$$\|X\|_\infty \stackrel{\text{def}}{=} \lambda_1\left(\sqrt{X^* X}\right).$$

This induces a norm on maps of the form $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ as

$$\|\Phi\|_\infty \stackrel{\text{def}}{=} \max\left\{\|\Phi(X)\|_\infty \,:\, X \in \mathrm{L}(\mathcal{X}),\ \|X\|_\infty \leqslant 1\right\}.$$

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

$$|||\Phi|||_\infty \stackrel{\text{def}}{=} \sup_{k \geqslant 1} \left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_\infty = \left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right\|_\infty .$$

This norm is commonly called the **completely bounded norm** and is denoted $\|\Phi\|_{\mathrm{cb}}$.

Comparison: spectral and trace CB-norms

The trace and spectral norms are dual, meaning

$$\|A\|_1 = \max\{|\langle B, A\rangle| : \|B\|_\infty \leqslant 1\},$$
$$\|A\|_\infty = \max\{|\langle B, A\rangle| : \|B\|_1 \leqslant 1\},$$

where $\langle B, A\rangle \stackrel{\text{def}}{=} \operatorname{Tr}(B^* A)$.

Given a mapping $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$, we define $\Phi^* : \mathrm{L}(\mathcal{Y}) \to \mathrm{L}(\mathcal{X})$ to be the unique mapping that satisfies

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

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

By the duality of the trace and spectral norms, it follows that

$$|||\Phi|||_1 = |||\Phi^*|||_\infty$$

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

Properties and uses of these norms

There are many nice properties of the completely bounded trace and spectral norms. For example, they are multiplicative with respect to tensor products:

$$|||\Phi \otimes \Psi|||_1 = |||\Phi|||_1 \; |||\Psi|||_1 \, .$$

The existence of an "auxiliary space" in their definitions allow them to be used in various computational and cryptographic settings.

They have found several applications in the theory of quantum information. For instance:

1. Bounding errors in quantum computations
   [Kitaev 1997; Aharonov, Kitaev & Nisan 1998].
2. Error reduction in quantum interactive proof systems
   [Kitaev & W. 2000]
3. Study of Bell inequality violations
   [Pérez–García, Wolf, Palazuelos, Villanueva and Junge 2008]

Statement of results

Main result of this work

The completely bounded trace and spectral norms can be expressed in a simple and efficient way through the use **semidefinite programming**.

There are two consequences of this fact:

1. For a given mapping Φ, the values $|||\Phi|||_1$ and $|||\Phi|||_\infty$ can be efficiently computed through the use of semidefinite programming algorithms.
2. The duality theory for semidefinite programming yields simple proofs for some interesting (previously known) facts.

Semidefinite programming

A **semidefinite program** is a pair of optimization problems, determined by a mapping

$$\Psi : \mathrm{Herm}(\mathcal{X}) \to \mathrm{Herm}(\mathcal{Y})$$

and a pair of operators

$$C \in \mathrm{Herm}(\mathcal{X}) \qquad \text{and} \qquad D \in \mathrm{Herm}(\mathcal{Y}),$$

with the following form:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \langle C, X \rangle \\ \text{subject to:} \quad & \Psi(X) \leqslant D, \\ & X \in \mathrm{Pos}(\mathcal{X}). \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \langle D, Y \rangle \\ \text{subject to:} \quad & \Psi^*(Y) \geqslant C, \\ & Y \in \mathrm{Pos}(\mathcal{Y}). \end{aligned}$$

(Other equivalent formulations include the so-called *standard form*.)

An SDP for the completely bounded trace norm

Suppose $\Phi : \mathrm{L}(\mathcal{X}) \rightarrow \mathrm{L}(\mathcal{Y})$ is a mapping given by

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

for $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$. Consider this semidefinite program:

Primal problem

maximize $\langle BB^*, P \rangle$

subject to:

$$\mathrm{Tr}_{\mathcal{Y}}(P) = \mathrm{Tr}_{\mathcal{Y}}(A\rho A^*),$$
$$\rho \in \mathrm{D}(\mathcal{X}),$$
$$P \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{Z}).$$

Dual problem

minimize $\left\| A^*(\mathbb{1}_{\mathcal{Y}} \otimes Q)A \right\|_{\infty}$

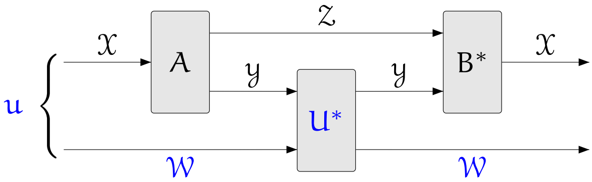
subject to:

$$\mathbb{1}_{\mathcal{Y}} \otimes Q \geqslant BB^*,$$
$$Q \in \mathrm{Pos}(\mathcal{Z}).$$

Claim: the optimal value (of both problems) is $|||\Phi|||_1^2$.

An abstract game

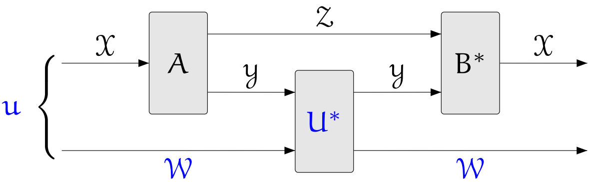
To see why the claim is true, consider an abstract game as follows.



Rules of play:

1. Choose $\mathcal{W}$ and a unit vector $u \in \mathcal{X} \otimes \mathcal{W}$.
2. The operator A is applied to $\mathcal{X}$.
3. Choose a unitary operator $U \in \mathrm{U}(\mathcal{Y} \otimes \mathcal{W})$ and apply U^* to $\mathcal{Y} \otimes \mathcal{W}$.
4. The operator B^* is applied to $\mathcal{Z} \otimes \mathcal{Y}$.

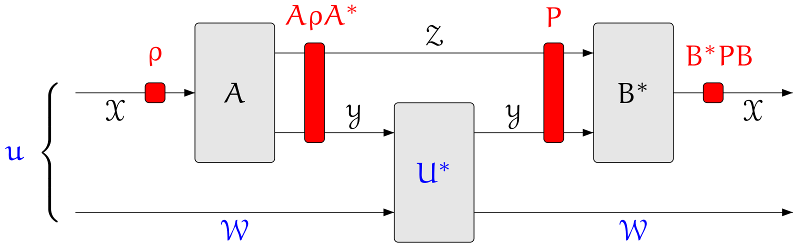
Goal: maximize $\|(B^* \otimes \mathbb{1}_{\mathcal{W}})(\mathbb{1}_{\mathcal{Z}} \otimes U^*)(A \otimes \mathbb{1}_{\mathcal{W}})\, u\|$, which is the length of the resulting vector.



Using an optimal strategy, the length of the resulting vector is:

$$
\begin{aligned}
\max_{U,\,u} \left\| (B^* \otimes \mathbb{1}_{\mathcal{W}})(\mathbb{1}_{\mathcal{Z}} \otimes U^*)(A \otimes \mathbb{1}_{\mathcal{W}})\, u \right\| \\
&= \max_{U,\,u,\,v} \left| v^* (B^* \otimes \mathbb{1}_{\mathcal{W}})(\mathbb{1}_{\mathcal{Z}} \otimes U^*)(A \otimes \mathbb{1}_{\mathcal{W}})\, u \right| \\
&= \max_{U,\,u,\,v} \left| \operatorname{Tr}\left[(\mathbb{1}_{\mathcal{Z}} \otimes U^*)(A \otimes \mathbb{1}_{\mathcal{W}})\, uv^* (B^* \otimes \mathbb{1}_{\mathcal{W}}) \right] \right| \\
&= \max_{U,\,u,\,v} \left| \left\langle U, \operatorname{Tr}_{\mathcal{Z}}\left[(A \otimes \mathbb{1}_{\mathcal{W}})\, uv^* (B^* \otimes \mathbb{1}_{\mathcal{W}}) \right] \right\rangle \right| \\
&= \max_{u,v} \left\| \operatorname{Tr}_{\mathcal{Z}}\left[(A \otimes \mathbb{1}_{\mathcal{W}})\, uv^* (B^* \otimes \mathbb{1}_{\mathcal{W}}) \right] \right\|_1 \\
&= |||\Phi|||_1 \qquad (\text{ for } \Phi(X) = \operatorname{Tr}_{\mathcal{Z}}(AXB^*)).
\end{aligned}
$$

Now consider the possible “states” of various subsystems as the game is played:



The possible choices of $P \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{Z})$ are **precisely** those that satisfy $\mathrm{Tr}_{\mathcal{Y}}(P) = \mathrm{Tr}_{\mathcal{Y}}(A\rho A^*)$ for some choice of $\rho \in \mathrm{D}(\mathcal{X})$.

Maximizing over all such choices of P gives:

$$|||\Phi|||_1^2 = \max_P \mathrm{Tr}(B^*PB) = \max_P \langle BB^*, P \rangle .$$

This maximization corresponds to the primal problem in our SDP.

CB trace and spectral norm computation

- Algorithms for computing the completely bounded trace and spectral norms have been known prior to this work:

  [V. ZARIKIAN, 2006]

  [N. JOHNSTON, D. KRIBS & V. PAULSEN, 2009]

  These are iterative methods, and bounds on their rates of convergence have not been established.

- Computation of $|||\Phi_0 - \Phi_1|||_1$ for *channels* Φ_0 and Φ_1 was previously argued to reduce to a convex optimization problem.

  [A. GILCHRIST, N. LANGFORD & M. NIELSEN, 2005]

- This work shows that $|||\Phi|||_1$ and $|||\Phi|||_\infty$ can be efficiently computed for general maps Φ using algorithms for semidefinite programming.

- Independently, [A. BEN-AROYA AND A. TA-SHMA, 2009] gave a different way to efficiently compute $|||\Phi|||_1$ and $|||\Phi|||_\infty$ for general maps (using convex optimization).

An analytic application

Let $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$, and consider the set of all pairs (A, B) for which

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(AXB^*) \qquad \text{(for all } X \in \mathrm{L}(\mathcal{X})\text{)}.$$

It is known [PAULSEN 2002, KITAEV, SHEN & VYALYI 2002] that

$$|||\Phi|||_1 = \inf_{(A,B)} \|A\|_\infty \|B\|_\infty . \tag{1}$$

An alternate proof of this fact follows from the dual problem for our SDP. For any fixed choice of A_0 and B_0, we have that $|||\Phi|||_1^2$ is given by

$$\begin{aligned} \text{minimize :} \quad & \|A_0^*(\mathbb{1}_{\mathcal{Y}} \otimes Q)A_0\|_\infty \\ \text{subject to:} \quad & \mathbb{1}_{\mathcal{Y}} \otimes Q \geqslant B_0 B_0^*, \quad Q \in \mathrm{Pos}(\mathcal{Z}) . \end{aligned}$$

Letting Q range over positive definite operators and taking

$$A = \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{1/2}\right) A_0 \qquad \text{and} \qquad B = \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right) B_0$$

establishes the non-trivial inequality in (1).

Conclusion

In this work it has been shown that the completely bounded trace and spectral norms can be expressed in a simple and efficient way through the use of semidefinite programming.

This provides a provably efficient and practical way to compute these norms, and gives simple proofs of a couple of known facts.

Thanks to Patrick, Marius, David, Debbie, Mary Beth, and Andreas for the invitation to speak, and thank you for your attention!