Imperial College
London

Algorithms & Complexity Theory Seminar Series
Oxford, 16 May 2019

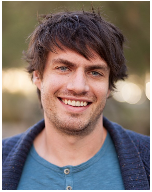
Semidefinite programming hierarchies for quantum information

Mario Berta

[SIAM Journal on Optimization '16] with Fawzi and Scholz
[arXiv:1810.12197] with Borderi, Fawzi, and Scholz

Imperial College
London

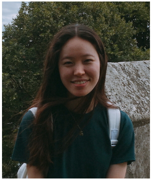
Theory of Quantum Information at Imperial



Mario Berta
Lecturer

Carlo Sparaciari
EPSRC Doctoral Prize

Francesco Borderi
PhD Department of Computing



Hyejung Jee
PhD Controlled Quantum Dynamics

Navneeth Ramakrishnan
Imperial President's PhD

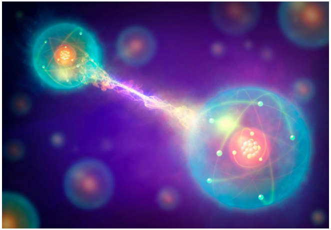
Samson Wang
PhD Quantum Systems Engineering

Quantum Information Processing

- The theory of information processing depends on underlying physical laws.
- Quantum information theory is based on (non-relativistic) quantum mechanics.

Research over the past two decades has shown:

Quantum information is in general fundamentally different from classical information.



Entangled quantum particles

- Quantum information is largely rooted in physics and thus often uses techniques motivated from physics.

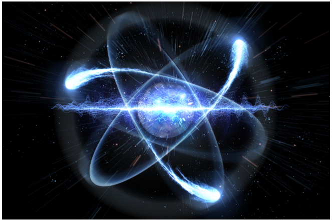
Imperial College London

Theoretical Computer Science for Quantum Information

Basic idea:

Many quantities of interest in information processing can be phrased as bilinear maximization programs $\Rightarrow$ make use of this structure.

- Winning probability of two-player games (CS) — violation of Bell inequalities (Physics)
- Information theory, e.g., success probability of error correcting codes
- Cryptography, e.g., cheating probability of adversaries
- Communication complexity
- Completely positive semidefinite cone, etc.



Quantum technologies @Imperial

Optimization methods for quantum information processing:

Approximate bilinear maximization programs via semidefinite programming hierarchies!

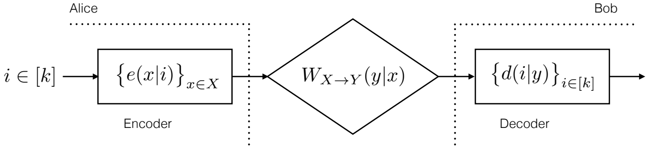
Imperial College London

This Talk

General approach:

Goal is to quantify the difference between classical and quantum information for operational problems — via semidefinite programming hierarchies.

- Develop by means of concrete example — noisy channel coding:

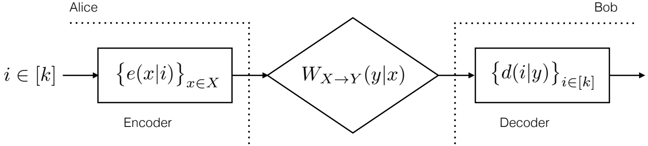


Noisy channel $W_{X \to Y}$ mapping X to Y with transition probability $W_{X \to Y}(y|x)$.

- Outline:
  0. Classical channel coding
  1. Quantum-assisted channel coding
  2. Quantum channel coding

Imperial College London

Part 0: Noisy Channel Coding



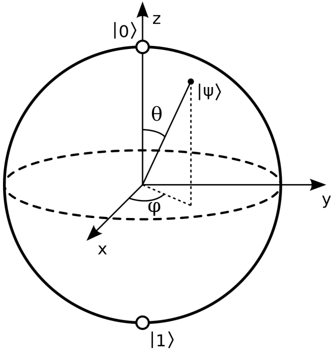
- The goal is to send k bits using $W_{X\to Y}$ while maximizing the success probability for decoding:

$$
\begin{aligned}
p(W,k) = \underset{e,d}{\text{maximize}} \quad & \frac{1}{2^k}\sum_{x,y,i} W_{X\to Y}(y|x)\,d(i|y)\,e(x|i) \\
\text{subject to} \quad & \sum_x e(x|i) = 1\ \forall i \in [2^k], \quad \sum_i d(i|y) = 1\ \forall y \in Y \\
& 0 \le e(x|i) \le 1, \quad 0 \le d(i|y) \le 1
\end{aligned}
$$

- Information-theoretic approach to error correction: Bilinear optimisation with linear constraints.

Quantum Assistance: Bits versus Qubits

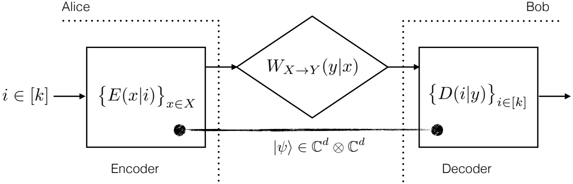
- Classical information unit:
  Bits take values 0 or 1 with certain probabilities
- Quantum information unit:
  Qubits take values $|\psi\rangle$ on the Bloch sphere $S^2 \subset \mathbb{R}^3$
- Quantum effects such as uncertainty principle for measurements or entanglement from quantum correlations



Quantum state:

For d degrees of freedom general quantum state given by unit vector $|\psi\rangle \in \mathbb{C}^d$.

Part 1: Quantum-Assisted Channel Coding



- Use $W_{X\to Y}$ and quantum assistance — that is to be optimized over:

$$p^*(W,k) = \underset{|\psi\rangle\in\mathbb{C}^d, E, D}{\text{maximize}} \quad \frac{1}{2^k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle$$

$$\text{subject to} \quad \sum_x E(x|i) = \mathrm{id}_{d\times d}\ \forall i\in[2^k], \quad \sum_i D(i|y) = \mathrm{id}_{d\times d}\ \forall y \in Y$$

$$0 \preceq E(x|i) \preceq \mathrm{id}_{d\times d}, \quad 0 \preceq D(i|y) \preceq \mathrm{id}_{d\times d}, \quad \||\psi\rangle\| = 1$$

- Versus classical value: $p(W,k) = \max_{e,d} \frac{1}{2^k}\sum_{x,y,i} W_{X\to Y}(y|x)d(i|y)e(x|i)$.

Imperial College London

Quantum-Assisted Channel Coding (continued)

Goal is to understand the possible separation:

$p(W,k) \leq p^*(W,k) \leq ?$ Is this even computable?

- Example separation for $d = 2$ quantum assistance [Prevedel *et al.* '11]:

$$W = \begin{pmatrix} 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 1/3 \\ 1/3 & 0 & 0 & 1/3 \\ 0 & 1/3 & 1/3 & 0 \end{pmatrix} \text{ has } p(W,2) = \frac{5}{6} \approx 0.833 < 0.902 \approx \frac{2+2^{-1/2}}{3} \leq p^*(W,2).$$

- Also optimal for $d = 2$ quantum assistance [Hemenway *et al.* '13] but remains unclear what $p^*(W,2) = ?$
- Lower bounds on $p^*(W,k)$ via feasible points and, e.g., see-saw optimisation.

Quantum-Assisted Channel Coding (continued)

$$
\begin{aligned}
p^*(W,k) = \underset{|\psi\rangle\in\mathbb{C}^d, E, D}{\text{maximize}} \quad & \frac{1}{2^k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle \\
\text{subject to} \quad & \sum_x E(x|i) = \mathrm{id}_{d\times d}\ \forall i\in[2^k], \quad \sum_i D(i|y) = \mathrm{id}_{d\times d}\ \forall y\in Y \\
& 0 \preceq E(x|i) \preceq \mathrm{id}_{d\times d}, \quad 0 \preceq D(i|y) \preceq \mathrm{id}_{d\times d}, \quad \||\psi\rangle\| = 1
\end{aligned}
$$

- See-saw based lower bounds on $p^*(W,k)$ lead to semidefinite programs.

Semidefinite programs (SDP):

Optimization of a linear objective function over the intersection of the cone of positive semidefinite matrices with an affine space.
$\Rightarrow$ solved efficiently in terms of the size of the matrices and the approximation error

- Question: How to generate upper bounds on $p^*(W,k)$?
- Answer: Non-commutative version of sum-of-squares semidefinite programming hierarchies of [Lasserre '01] and [Parrilo '03]!

Imperial College London

SDPs for Quantum-Assistance

Asymptotically converging hierarchy of SDP relaxations:

Motivated by [Navascues *et al.* '07 '08] [Doherty *et al.* '08] [Pironio *et al.* '10] we give

$$p^*(W,k) \le \mathrm{sdp}_n(W,k) \le \cdots \le \mathrm{sdp}_1(W,k) \text{ with } p^*(W,k) = \lim_{n\to\infty} \mathrm{sdp}_n(W,k).$$

- For our example channel W the first level gives

$$p^*(W,2) \le \mathrm{sdp}_1(W,2) \approx 0.908 = \frac{1}{2} + \frac{1}{\sqrt{6}}$$

  versus $p(W,2) = \frac{5}{6} \approx 0.833$ and $d = 2$ lower bound $p^*(W,2) \ge 0.902$.
  Then found $d = 4$ lower bound $p^*(W,2) \ge \frac{1}{2} + \frac{1}{\sqrt{6}}$!

- Follow-up [Barnam and Fawzi '18] give rounding for linear program $\mathrm{lp}(W,k)$ as

$$p(W,k) \le p^*(W,k) = \mathrm{sdp}_\infty(W,k) \le \mathrm{sdp}_1(W,k) \le \mathrm{lp}(W,k) \le \frac{1}{1-e^{-1}} \cdot p(W,k)$$

- Better than $(1-e^{-1})$-approximation NP-hard and simple polynomial-time algorithms for classical error correction codes that achieve this approximation.

SDPs for Quantum-Assistance (continued)

- General form quantum information [B. *et al.*, SIAM Journal on Optimization '16]:

$$
\begin{aligned}
p^*[A,\mathcal{G},\mathcal{K}] = \underset{|\psi\rangle\in\mathcal{H},E_\alpha,D_\beta}{\text{maximize}} \quad & \sum_{\alpha,\beta} A_{\alpha,\beta}\langle\psi|E_\alpha D_\beta|\psi\rangle \\
\text{subject to} \quad & E_\alpha D_\beta = D_\beta E_\alpha \quad \forall(\alpha,\beta)\in[N]\times[M] \\
& g(E_1,\dots,E_N)\succeq 0\ \forall g\in\mathcal{G} \\
& k(D_1,\dots,D_M)\succeq 0\ \forall k\in\mathcal{K}
\end{aligned}
$$

  with normalized $|\psi\rangle\in\mathcal{H}$ (Hilbert space), Hermitian bounded operators E_α, D_β on $\mathcal{H}$, and sets of affine constraints $\mathcal{G},\mathcal{K}$.

- Proof of asymptotic convergence

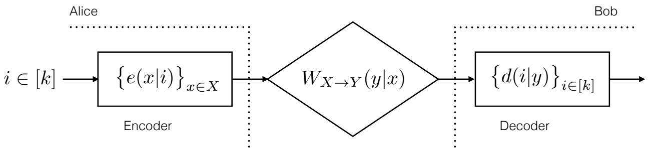
$$
p^*(A,\mathcal{G},\mathcal{K}) = \lim_{n\to\infty} \mathrm{sdp}_n(A,\mathcal{G},\mathcal{K})
$$

  via non-commutative Positivstellensatz [Helton and McCullough '04] and purifications in C^*-algebras [Woronowicz '73].

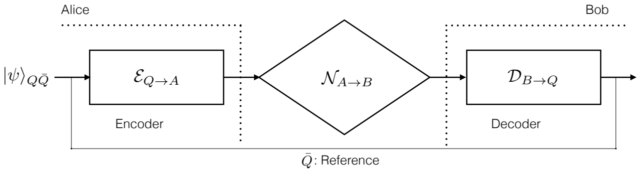
- Finite convergence unclear, related to deep problems in operator algebra theory: Connes' embedding conjecture.

Part 2: Quantum Channel Coding

- Classical channel coding (with or without quantum assistance):



- Make message and channel itself quantum — quantum channel coding:



Quantum Channel Coding (continued)

- The goal is to send k qubits using quantum channel $\mathcal{N}_{A\to B}$ while maximizing the quantum channel fidelity (uniform message quantum success probability)

$$F(\mathcal{N},k) = d_A d_B \cdot \underset{E,D}{\text{maximize}} \quad \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{AB} \otimes \Phi_{Q\bar{Q}}\right)\left(E_{AQ} \otimes D_{B\bar{Q}}\right)\right]$$
$$\text{subject to} \quad E_{AQ} \succeq 0, \quad E_A = \frac{1_{d_A\times d_A}}{d_A}$$
$$D_{B\bar{Q}} \succeq 0, \quad D_B = \frac{1_{d_B\times d_B}}{d_B}$$

  with $\Phi_{Q\bar{Q}}$ the 2^k dimensional maximally entangled state and $J^{\mathcal{N}}_{\bar{A}B}$ the Choi state of $\mathcal{N}_{A\to B}$.

- Information-theoretic approach to quantum error correction:
  Bilinear optimization with matrix valued variables and linear constraints.

Noisy Intermediate-Scale Quantum (NISQ) technology:

Near-term quantum hardware is only a few qubits in size and inherently noisy, in particular quantum memory (in stark contrast to classical memory!).

SDPs for Quantum Optimization

$$F(\mathcal{N},k) = d_A d_B \cdot \underset{E,D}{\text{maximize}} \quad \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{AB} \otimes \Phi_{Q\bar{Q}}\right)\left(E_{AQ} \otimes D_{B\bar{Q}}\right)\right]$$

$$\text{subject to} \quad E_{AQ} \succeq 0, \quad E_A = \frac{1_{d_A \times d_A}}{d_A}$$

$$D_{B\bar{Q}} \succeq 0, \quad D_B = \frac{1_{d_B \times d_B}}{d_B}$$

- See-saw based lower bounds on $F(\mathcal{N},k)$ lead to semidefinite programs, e.g., explored in [Reimpell & Werner '05]. Upper bounds on $F(\mathcal{N},k)$?

Non-commutative sum-of-squares hierarchy of SDP relaxations:

Motivated by [Doherty *et al.* '02] we give

$$F(\mathcal{N},k) \leq \mathrm{sdp}_n(\mathcal{N},k) \leq \cdots \leq \mathrm{sdp}_1(\mathcal{N},k) \text{ with (slow) finite convergence}$$

$$\mathrm{sdp}_n(\mathcal{N},k) - F(\mathcal{N},k) \leq \frac{\mathrm{poly}\left(d_A, d_B, 2^k\right)}{\sqrt{n}}$$

- Convergence proof information-theoretic based on quantum de Finetti theorems with linear constraints.

SDPs for Quantum Optimization (continued)

- General form quantum information [B. *et al.* arXiv:1810.12197]:

$$
\begin{aligned}
F(G,\Lambda,\Gamma,X,Y) = \underset{W}{\text{maximize}} \quad & \mathrm{Tr}\big[G_{AB}(W_A \otimes W_B)\big] \\
\text{subject to} \quad & W_A \succeq 0, \quad W_B \succeq 0, \quad \mathrm{Tr}(W_A) = \mathrm{Tr}(W_B) = 1 \\
& \Lambda_{A\to C_A}(W_A) = X_{C_A}, \quad \Gamma_{B\to C_B}(W_B) = Y_{C_B}
\end{aligned}
$$

  where G_{AB} is a matrix, $\Lambda_{A\to C_A}$ and $\Gamma_{B\to C_B}$ are linear maps and X_{C_A} and Y_{C_B} are fixed matrices defining affine subspaces.

- Approximating $F(G,\Lambda,\Gamma,X,Y)$ or already $F(\mathcal{N},k)$ encodes so-called quantum separability problem $\Rightarrow$ strong hardness results [Gharibian '10] [Harrow *et al. '19*] and slow convergence is expected (versus classical setting!).

Ongoing work:

Identify settings in quantum information that allow for faster analytical convergence, e.g., classical-quantum channels — other methods via ε-nets / polynomials needed.

Imperial College London

Certifying Optimality

Rank loop conditions:

We give sufficient condition for exact convergence $F(\mathcal{N}, k) = \mathrm{sdp}_n(\mathcal{N}, k)$ but this requires finding low-rank SDP solutions via rank minimization heuristics [Fazel '02].

- Example numerics: Uniform noise modelled by qubit depolarizing channel

$$\mathcal{N}_{\mathrm{Dep}(p)} : |\psi\rangle\langle\psi| \mapsto p \cdot \mathrm{Uniform}_2 + (1-p) \cdot |\psi\rangle\langle\psi| \quad \text{with } p \in [0, 4/3].$$

$\Rightarrow$ What is the optimal quantum code for storing $k = 1$ qubit in $N = 5$ noisy qubits:

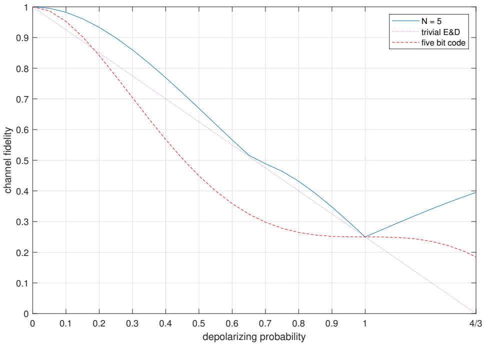
$$p\left(\mathcal{N}_{\mathrm{Dep}(p)}^{\otimes 5}, 2\right) = ?$$

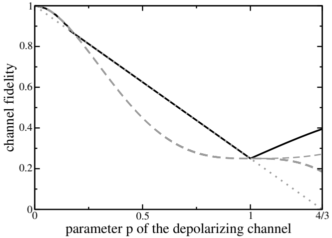
- Analytical [Bennett *et al.* '96] as well as numerical see-saw type [Reimpell & Werner '05] lower bounds available, what about upper bounds? Our work gives

$$\mathrm{sdp}_n\left(\mathcal{N}_{\mathrm{Dep}(p)}^{\otimes 5}, 2\right) \geq p\left(\mathcal{N}_{\mathrm{Dep}(p)}^{\otimes 5}, 2\right).$$

Preliminary Numerics

- Exploiting symmetries for analytical dimension reduction, we calculated the first level $\mathrm{sdp}_n\left(\mathcal{N}_{\mathrm{Dep}(p)}^{\otimes 5}, 2\right)$ as an upper bound on $p\left(\mathcal{N}_{\mathrm{Dep}(p)}^{\otimes 5}, 2\right)$ in MATLAB using CVX and MOSEK:





[Reimpell & Werner '05]

- For $p \in [1, 4/3]$ the codes from [Reimpell & Werner '05] are optimal — for $p \in [0, 0.18]$ there is room to look for improved codes.

Imperial College London

Conclusion

Take home message:

Methods form optimization theory for quantifying the difference between classical and quantum problems in information processing.

- Relevant bilinear programs in general hard but approximated via non-commutative sum-of-squares hierarchies.
- Analytical question: Identify operational settings in quantum information that allow for efficient approximation. For example, free games / quantum adversaries in classical cryptography / classical-quantum channel coding?
- Numerical question: Dimension reduction tools exploiting symmetries needed (group theory) $\Rightarrow$ extensive numerics for quantum channel coding practically relevant for NISQ technology.
- Multilinear / multipartite extensions?

Thanks!

[SIAM Journal on Optimization '16] with Fawzi and Scholz
[arXiv:1810.12197] with Borderi, Fawzi, and Scholz