Annual exam, PhD Nanosciences, 20 October 2020

Inference and communication with quantum systems

Marco Fanizza

Topics

Quantum statistical inference

Learn properties of quantum states measuring the least amount of copies of a state

Communication in the presence of noise

Study the fundamental limits on communication rates of classical and quantum information

Quantum learning

Quantum tomography requires $O(d^2)$ copies of a state in the worst case, giving a full classical description of the state. If one is interested in partial properties of a state, less copies may be required.

- Quantum learning machines: minimize the error probability of a universal machine that distinguishes $\rho^{\otimes n} \otimes \sigma^{\otimes n} \otimes \rho$ from $\rho^{\otimes n} \otimes \sigma^{\otimes n} \otimes \sigma$

  Fanizza, Mari, Giovannetti, IEEE Transactions on Information Theory 65(9), (2019)

- Evaluate the minimum mean square error for an estimator of the overlap $|\langle\psi|\phi\rangle|$ of two unknown states $|\psi\rangle$ and $|\phi\rangle$, given m copies of $|\psi\rangle$ and n copies of $|\phi\rangle$.

  Fanizza, Rosati, Skotiniotis, Calsamiglia, Giovannetti, Physical Review Letters 124.6 (2020)

Optimal universal measurements can be obtained using symmetry principles.

Quantum learning: current projects

- Estimate distances between mixed states: quantum Jensen-Shannon divergence $QJS(\rho,\sigma) = S(\frac{\rho+\sigma}{2}) - \frac{1}{2}S(\rho) - \frac{1}{2}S(\sigma)$, achievable with $\tilde{O}(r^2/\epsilon^2)$ copies, trace distance $D(\rho,\sigma) = \frac{||\rho-\sigma||_1}{2}$ currently open
- Testing identity of a collection of m states, scaling in m of the sample complexity
- Testing membership to a set of m states, scaling in m of the sample complexity.

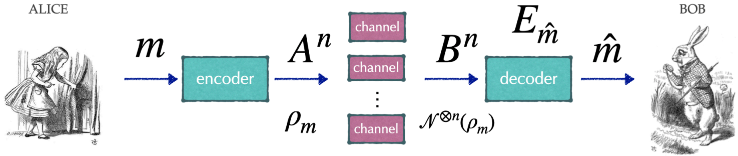
New tool: regularized least square measurements

A quantum generalization of the union bound $P(\bigcup_i E_i) \leq \sum_i P(E_i)$, with a regulator granting a better bound if events are not linearly independent.

Classical communication over quantum channels

Alice wants to send a message to Bob, in the presence of noise. Alice and Bob want to establish a protocol such that despite of the noise, Bob is able to recover Alice's message exactly, with high probability.

- Noise model $\mathcal{N}_{A\to B} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$
- Alice encodes: $m \to \mathcal{C}(m) = \rho \in \mathfrak{S}(\mathcal{H}_A^{\otimes n})$, M total messages
- Bob receives $\mathcal{N}^{\otimes n}(\rho)$
- Bob decodes with a POVM $\{E_{\hat{m}}\}_{\hat{m}\in|M|}$, $p(\hat{m}|m) = \mathrm{tr}[\mathcal{N}^{\otimes n}(\rho_m)E_{\hat{m}}]$
- Rate $R = \frac{\log_2 M}{n}$, classical capacity $C(\mathcal{N})$ supremum of achievable rates

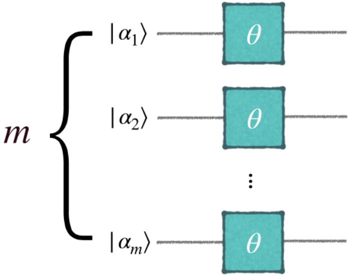


Loss of phase reference

- Alice's and Bob's frame differ by a phase shift: $e^{-i\theta a^\dagger a}$
- Loss of any phase reference after m uses of the transmission line
- In each sequence of m consecutive uses the phase shift is uniformly random

$$\Phi_m(\rho) = \int \frac{d\theta}{2\pi} e^{-i\theta\hat{n}} \rho e^{i\theta\hat{n}} \qquad (1)$$

where $\hat{n} = \sum_{i=1}^{m} \hat{a}_i^\dagger \hat{a}_i$ is the total-photon-number operator



Fanizza, Rosati, Skotiniotis, Calsamiglia, Giovannetti, arXiv preprint arXiv:2006.06522

Classical communication with loss of phase synchronization

The optimal strategy: Fock states maximize the capacity, but they cannot be easily produced

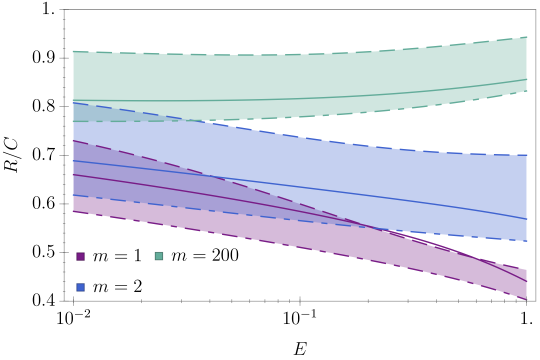
Restricted encodings

What are the optimal rate with more economical states?

- An intuitive strategy: send a phase reference on one mode and coherent states on the remaining $m - 1$ modes.
- Optimal strategy with coherent states: generate a coherent state in one mode and then apply a random interferometer. We compute upper and lower bounds. Better than any phase synchronization strategy at high energy

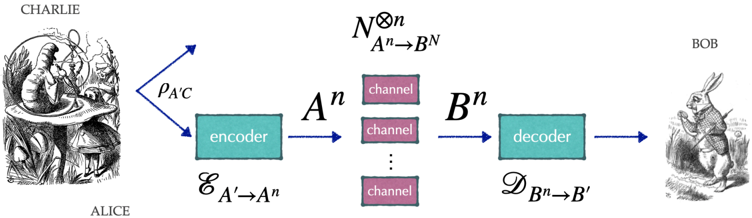
Advantage from squeezed coherent states

- Squeezed coherent states can have lower entropy than Poisson in the photon number distribution
- For $m = 1$ a binary encoding with vacuum and a squeezed coherent state beats coherent states. In stark contrast with phase-insensitive Gaussian channels
- For $m > 1$ lower bound with squeezing are considerably higher. Conjectured to beat coherent encodings at all m



Quantum communication over quantum channels

Alice possesses a share of a quantum state, possibly entangled with Charlie. She wants to send Bob her share of the quantum state, such that the joint state held by Bob and Charlie is close to the original one, independently of the the particular state.



The maximum ratio between the qubits sent and number of uses of the channel gives fundamental limits on error correction in quantum computers and quantum memories.

Bounds from flagged extensions (with F. Kianvash)

For a convex combination of channels $\Lambda(\rho) = \sum_i p_i \Lambda_i(\rho)$ one can define a flagged extension

$$\Lambda^F(\rho) = \sum_i p_i \Lambda_i(\rho) \otimes |\phi_i\rangle\langle\phi_i|$$

- Flagged extensions have higher quantum capacity, but exactly computable under certain sufficient conditions.
- We can compute state-of-the-art upper bounds for the main finite dimensional quantum channels, such as depolarizing channel and generalized amplitude damping channel - fundamental noise models applicable to qubit quantum processors.
- Open problems: better bounds from flagged extensions of $\Lambda^{\otimes n}$, generalization to continuous variable channels

Fanizza, Kianvash, Giovannetti, Physical Review Letters 125.2, 020503, (2020)

Kianvash, Fanizza, Giovannetti, arXiv preprint arXiv:2008.02461

Summary and open questions

Key results

- Characterization of optimal measurements for several quantum learning tasks
- Squeezing helps in classical communication over a practically motivated non-Gaussian channel
- New flexible and effective technique to bound quantum capacities.

Open questions

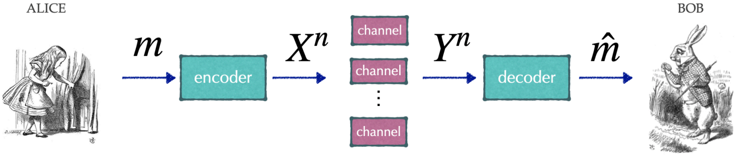
- Membership problem using regularized least square measurement
- Improve and extend bounds from flagged channels

Thank you

Classical communication

Alice wants to send a message to Bob, but anything that Alice sends gets corrupted by some noise. Alice and Bob want to establish a protocol such that despite of the noise, Bob is able to recover Alice's message exactly, with high probability.

- Alice encodes: $m \to \mathcal{C}(m) = (x_1, ..., x_n)$
- Channel $x_i \to y_i \sim p(y_i|x_i)$
- Bob receives $(y_1, ..., y_n)$ with probability $p(y_1|x_1)p(y_2|x_2)\cdots p(y_n|x_n)$
- Bob decodes: $\hat{m} = \mathcal{D}(y_1, ..., y_n)$



Transmission rate

Alice and Bob want to operate protocols such that the probability of decoding each word exactly is arbitrarily close to one. They also want to communicate in the most economical way, that is to use the transmission line the smallest number of times possible.

- $\hat{m} \neq m$ with some probability. Largest probability of decoding error $p_{err}(\mathcal{C})$
- Number of possible messages M, equivalent to $\log_2 M$ bits
- Number of uses of the channel n
- Rate $R = \frac{\log_2 M}{n}$, achievable if there exists a sequence of codes $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ with rate R such that $\lim_{n\to\infty} p_{err}(\mathcal{C}_n) = 0$
- Capacity of a channel, $C =$ supremum of achievable rates

Classical capacity and mutual information

Shannon noisy coding theorem

$$C = \max_{p(x)} I(X;Y)$$

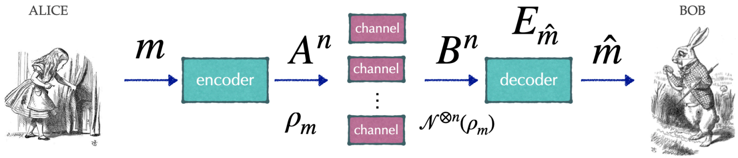
- $p(x,y) = p(y|x)p(x)$
- $I(X,Y) = H(X) + H(Y) - H(X,Y)$, $H(X) = -\sum_i p(x_i) \log_2 p(x_i)$
- Single-letter formula

Remark

In Shannon theory computation is free, communication is expensive. We are neglecting the computational difficulties of encoding and decoding. In this sense, the theorem sets fundamental limits on information transmission, and sets a goal for any proposed practical implementation.

Classical communication over quantum channels

- Noise model: CPTP map $\mathcal{N}_{A\to B} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$
- Alice encodes: $m \to \mathcal{C}(m) = \rho \in \mathfrak{S}(\mathcal{H}_A^{\otimes n})$
- Bob receives $\mathcal{N}^{\otimes n}(\rho)$
- Bob decodes with a POVM $\{E_{\hat{m}}\}_{\hat{m}\in|M|}$, $p(\hat{m}|m) = \mathrm{tr}[\mathcal{N}^{\otimes n}(\rho_m)E_{\hat{m}}]$
- Rate $R = \frac{\log_2 M}{n}$, classical capacity $C(\mathcal{N})$ supremum of achievable rates



Classical capacity of a quantum channel

Holevo-Schumacher-Westmoreland theorem

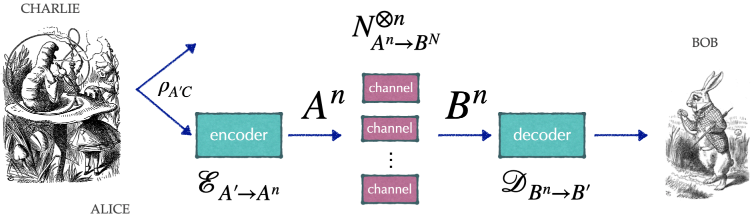
$$C(\mathcal{N}) = \lim_{n\to\infty} \frac{\chi(\mathcal{N}^{\otimes n})}{n} \qquad \chi(\mathcal{N}) := \max_{\{p_i,\rho_i\}} \chi(\mathcal{N},\rho)$$

- $\chi(\mathcal{N},\{p_i,\rho_i\}) = S(\sum_i p_i \mathcal{N}(\rho_i)) - \sum_i p_i S(\mathcal{N}(\rho_i))$
- $S(\rho) = -\operatorname{tr}[\rho \log_2 \rho]$
- Regularized formula, cannot be used to compute the capacity. One could have $\chi(\mathcal{N}^{\otimes k}) > k\chi(\mathcal{N})$ (superadditivity)
- $\chi(\mathcal{N})$ is a lower bound on the capacity and it's the best rate achievable without using entanglement across different uses of the channel

Quantum communication over quantum channels

Alice possesses a share of a quantum state, possibly entangled with Charlie. She wants to send Bob her share of the quantum state, such that the joint state held by Bob and Charlie is close to the original one, independently of the the particular state.

- Initial state $\rho_{A'C} \in \mathfrak{S}(\mathcal{H}_{A'} \otimes \mathcal{H}_C)$, Alice encodes: $\rho_{A'C} \to \mathcal{E}_{A' \to A^n}(\rho_{A'C})$
- Bob receives $\mathcal{N}^{\otimes n}_{A^n \to B^n} \circ \mathcal{E}_{A' \to A^n}(\rho_{A'C})$
- Bob decodes and obtains $\rho'_{B'C} = \mathcal{D}_{B^n \to B'} \circ \mathcal{N}^{\otimes n}_{A^n \to B^N} \circ \mathcal{E}_{A' \to A^n}(\rho_{A'C})$
- Rate $R = \frac{\log_2 \dim \mathcal{H}_{A'}}{n}$, quantum capacity $Q(\mathcal{N})$ supremum of achievable rates



Quantum capacity of a quantum channel

Lloyd-Devetak-Shor theorem

$$Q(\mathcal{N}) = \lim_{n\to\infty} \frac{I_c(\mathcal{N}^{\otimes n})}{n} \qquad I_c(\mathcal{N}) := \max_{\rho_{AA'}} I_c(\mathcal{N}, \rho_{AA'})$$

- $I_c(\mathcal{N}, \rho_{AA'}) = S(\mathcal{N}(\rho_A)) - S(\mathcal{N} \otimes I(\rho_{AA'}))$
- Regularized formula, cannot be used to compute the capacity. One could have $I_c(\mathcal{N}^{\otimes k}) > kI_c(\mathcal{N})$ (superadditivity)
- $I_c(\mathcal{N})$ is a lower bound on the quantum capacity

Given the difficulties of calculating capacities of quantum channels, finding general bounds or bounds that work for particular channels is considered an important improvement.

Quantum capacity of the depolarizing channel

$$\Lambda(\rho) := (1-p)\rho + p\,\mathrm{Tr}[\rho]\frac{I}{2} = (1-\frac{3p}{4})\rho + \frac{p}{4}X\rho X + \frac{p}{4}Y\rho Y + \frac{p}{4}Z\rho Z \tag{2}$$

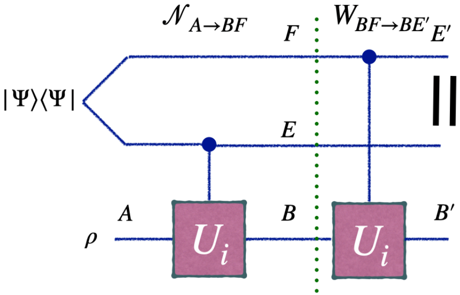
- Most symmetric qubit to qubit channel (generalizable to d-level systems).
- Convex combination of Pauli operations, models random phase-flip and bit-flip errors in qubits, which represent ideal errors to correct to preserve the state of a quantum memory.
- $Q = 0$ for $p \geq 1/4$.
- Additive classical capacity, superadditive coherent information.

After more than 20 years of study, the quantum capacity of this simple channel is still unknown, but a long series of efforts have improved the upper bounds on the quantum capacity.

Approaches for upper bounds on the quantum capacity

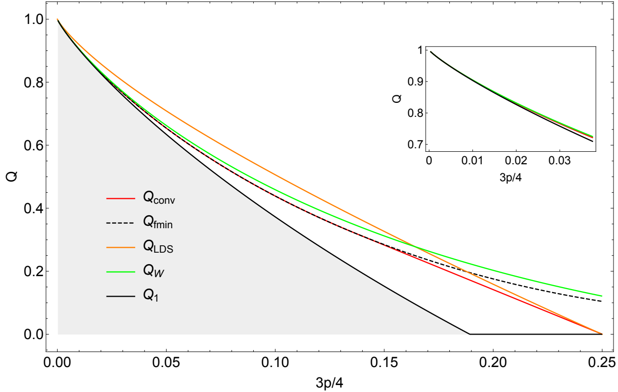
- Find a degradable channel Λ^{ext} that extends $\Lambda = \Xi \circ \Lambda^{ext}$ for some CPTP Ξ. (Smith-Smolin 2008, Ouyang 2014, Leditzky et al. 2018, MF et al. 2020)
- Approximate degradability: minimize $\epsilon = ||\tilde{\Lambda} - \Xi \circ \Lambda||_\diamond$ over maps Ξ. Continuity of the capacity gives bounds on the quantum capacity, as $\epsilon = 0$ for degradable channels. (Sutter et al. 2017, Leditzky et al. 2018)
- Combination of both techniques: approximate degradability for extensions (Wang, arXiv:1912.00931).

Sufficient conditions for degradability of flagged extensions



- $|\Psi\rangle = \sum_i \sqrt{p_i}\,|i\rangle\,|\phi_i\rangle$, with controlled unitary gives a flagged extension $\sum_i p_i \Lambda_i \otimes |\phi_i\rangle\langle\phi_i|$
- If the state after $W_{BF\to BE'}$ is symmetric under exchange $E \leftrightarrow E'$ the channel is degradable
- For two pure flags we get an overlap $c = \sqrt{\frac{1-2p}{1-p}}$
- Generalizable to many flags

Bounds on the quantum capacity of the depolarizing channel



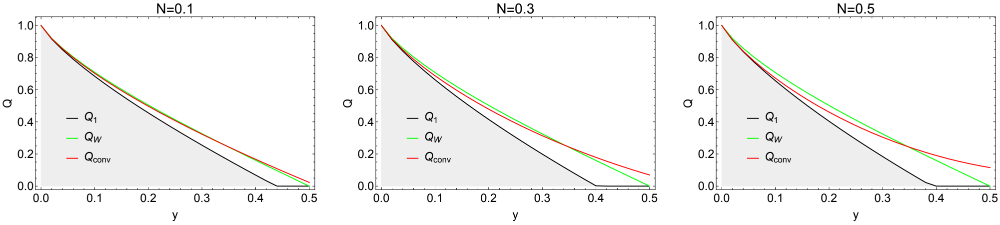
Q_{conv} is the convex hull of the available upper bounds from degradable extensions, Q_{fmin} is the new upper bound, and Q_1 is the lower bound given by the coherent information of one use of the channel. Q_{LDS} is the bound from Leditzky et al. 2018 and Q_W is the bound from Wang using flagged extensions.

Generalized amplitude damping channel (GAD)

$$\mathcal{A}_{y,N}(\rho) = N\,\mathcal{A}_y(\rho) + (1-N)\,X \circ \mathcal{A}_y \circ X(\rho),$$

- Realistic model of errors on superconducting qubits. Physically it represents an interaction with another environmental qubit, such that the states $|00\rangle$ and $|11\rangle$ are invariant but a rotation in the space of the states $|01\rangle$ and $|10\rangle$ occurs.
- The environment is in a thermal state, this channel models energy relaxation to the state of the environment. N is related to the temperature and y to the coupling of the interaction.
- If the environmental qubit is in the ground state, the channel is degradable, but as long as the temperature is not zero the quantum capacity cannot be evaluated anymore.

Bounds on the quantum capacity of GAD



Bounds on the quantum capacity of GAD from different flagged extensions. Q_1 is the lower bound given by the coherent information of one use of the GAD. Q_{conv} is the new upper bound, Q_W is the upper bound obtained by Wang.

Loss of phase reference

- Decoherence between total photon number subspaces

$$\Phi_m(\rho) = \int \frac{d\theta}{2\pi} e^{-i\theta\hat{n}} \rho e^{i\theta\hat{n}} = \sum_{n=0}^{\infty} \hat{\Pi}_n \rho \hat{\Pi}_n, = \sum_{n=0}^{\infty} p_n \rho_n, \tag{3}$$

  where $\hat{\Pi}_n$ is the projector on the subspace with total photon number n, $p_n := \mathrm{tr}\left[\hat{\Pi}_n \rho\right]$ and $\rho_n := \hat{\Pi}_n \rho \hat{\Pi}_n / p_n$.
- Fock states are invariant $\Phi_m(|\vec{n}\rangle \langle \vec{n}|) = |\vec{n}\rangle \langle \vec{n}|$
- The case $m = 1$ is equivalent to measuring photon number of each mode, For $m > 1$ some coherence is preserved.

Classical communication with loss of phase synchronization

- Any ensemble of states $\mathcal{E} = \{q_x, \rho_x\}$ gives an achievable rate $\chi(\Phi, \mathcal{E})$
- Energy constraint $\sum_x q_x \operatorname{tr}\{\hat{n}\rho_x\} \leq E$
- The optimal strategy: Fock states maximize the capacity, but they cannot be easily produced

$$C(\Phi_m, E) = m\, g(\frac{E}{m}), \qquad g(E) = -E \log E + (1+E)\log(1+E) \tag{4}$$

Restricted encodings

What are the optimal rate with more economical states? An intuitive strategy: send a phase reference on one mode and coherent states on the remaining $m-1$ modes. The phase can be estimated with error $\Delta\phi \geq 1/\Delta\hat{n}$

Optimality of covariant encodings

Any ensemble of states gives a higher or equal rate if scrambled with a random interferometer

$$\mathcal{E} \to \mathcal{E}^{Haar} = \{p_i dU, U\rho_i U^\dagger\}$$

$$\chi(\Phi_m, \mathcal{E}) \leq \chi(\Phi_m, \mathcal{E}^{Haar})$$

Rate of covariant encodings

- The rate for covariant encodings with pure states is

$$\chi(\Phi_m, \mathcal{E}^{Haar}) = \sum_{n=0}^{\infty} \sum_x q_x p_n(x) \log \binom{n+m-1}{m-1} \tag{5}$$

$$+ H\left(\sum_x q_x p(x)\right) - \sum_x q_x H(p(x)) \tag{6}$$

$$= mg(\frac{E}{m}) - D(\sum_x q_x p(x) || p^{Th}) - \sum_x q_x H(p(x)), \tag{7}$$

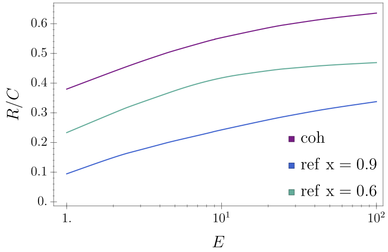
- Only the total photon number distribution is relevant
- States with low entropy in the photon number distribution seem preferred. This goes against phase synchronization schemes.

Covariant coherent states encoding

- Generate a coherent state in one mode and then apply a random unitary
- For $m = 1$ we reduce to the Poisson channel: equivalent to measuring photon number. Capacity still open
- Upper and lower bounds available
- Low energy: next to leading order in E attainable with binary encodings $|0\rangle, |\alpha\rangle$ and photodetection on each mode separately
- High energy: $R = (m - \frac{1}{2}) \log E + O(1)$, while $C(\Phi_m, E) = m \log E + O(1)$

Phase synchronization strategy

- Prepare a fixed reference state with energy $(1-x)E$. The rate is at most $(m-1)g(\frac{xE}{m-1})$, by monotonicity. At high energies this upper bound is $(m-1)\log E$, which is less than what obtained with a thermal ensemble of coherent states: $(m-\frac{1}{2})\log E$.
- The plot illustrates the rates for $m=2$ with a truncated phase state as reference $|\psi\rangle = [2xE+1]^{-1/2}\sum_{n=0}^{2xE}|n\rangle$



All in all, synchronization with nonclassical light seems not only suboptimal, but also detrimental with respect to naive coherent state encodings

Quantum learning: current projects

Testing problems

Devise a quantum measurement with binary outcome (YES/NO) such that the measurement applied to n copies of ρ

- If ρ has a property, returns YES with probability $> 1/3$
- If ρ is ϵ far from having the property, returns NO with probability $> 1/3$

The sample complexity is the minimum n such that this text exists.

Examples:

- Testing identity of a collection of m states, scaling in m of the sample complexity
- Testing membership to a set of m states, scaling in m of the sample complexity. Use a regularized least square measurement!