Applications of hypercontractivity in quantum information

Ashley Montanaro

Department of Computer Science,
University of Bristol

23 February 2015

Some applications

Boolean cube

- Communication complexity separations [Gavinsky et al. '07]
- Bounds on nonlocal games [Buhrman et al. '11] [Defant et al. '10, Pellegrino+Seoane-Sepúlveda '12, AM '12]
- Quantum query complexity bounds [Ambainis+de Wolf '12]

Some applications

Boolean cube

- Communication complexity separations [Gavinsky et al. '07]
- Bounds on nonlocal games [Buhrman et al. '11] [Defant et al. '10, Pellegrino+Seoane-Sepúlveda '12, AM '12]
- Quantum query complexity bounds [Ambainis+de Wolf '12]

Real n-sphere

- Communication complexity separations [Klartag+Regev '11]
- Biases of local measurements [Lancien+Winter '11, AM '12]

Some applications

Boolean cube

- Communication complexity separations [Gavinsky et al. '07]
- Bounds on nonlocal games [Buhrman et al. '11] [Defant et al. '10, Pellegrino+Seoane-Sepúlveda '12, AM '12]
- Quantum query complexity bounds [Ambainis+de Wolf '12]

Real n-sphere

- Communication complexity separations [Klartag+Regev '11]
- Biases of local measurements [Lancien+Winter '11, AM '12]

Noncommutative generalisations

- Limits of quantum random access codes [Ben-Aroya et al. '08]
- Rapid mixing of quantum channels [Kastoryano+Temme '13]

Hypercontractivity on the boolean cube

Consider functions $f : \{0, 1\}^n \to \mathbb{R}$.

- Set $\|f\|_p = \left(\frac{1}{2^n} \sum_x |f(x)|^p\right)^{1/p}$.
- For $\rho \in [0, 1]$, define the noise operator T_ρ as follows:

$$(T_\rho f)(x) = \mathbb{E}_{y \sim_\epsilon x}[f(y)],$$

where the expectation is over strings $y \in \{0, 1\}^n$ obtained from x by flipping each bit of x with independent probability $\epsilon = (1 - \rho)/2$.

Hypercontractivity on the boolean cube

Consider functions $f : \{0,1\}^n \to \mathbb{R}$.

- Set $\|f\|_p = \left(\frac{1}{2^n} \sum_x |f(x)|^p\right)^{1/p}$.
- For $\rho \in [0,1]$, define the noise operator T_ρ as follows:

$$(T_\rho f)(x) = \mathbb{E}_{y \sim_\epsilon x}[f(y)],$$

where the expectation is over strings $y \in \{0,1\}^n$ obtained from x by flipping each bit of x with independent probability $\epsilon = (1-\rho)/2$.

Hypercontractive inequality [Bonami '70] [Gross '75] [Beckner '75] [...]

For any $f : \{0,1\}^n \to \mathbb{R}$, and any p and q such that $1 \leqslant p \leqslant q \leqslant \infty$ and $\rho \leqslant \sqrt{\frac{p-1}{q-1}}$,

$$\|T_\rho f\|_q \leqslant \|f\|_p.$$

One-way communication complexity

- Alice and Bob want to determine some property $f(x, y)$ of their distributed inputs x, y, using the minimal amount of communication.

- All communication goes from Alice to Bob.

One-way communication complexity

- Alice and Bob want to determine some property $f(x, y)$ of their distributed inputs x, y, using the minimal amount of communication.

- All communication goes from Alice to Bob.

Question: Can quantum communication be more efficient than classical communication?

One-way communication complexity

Theorem **[Bar-Yossef, Jayram and Kerenidis '08]**

There is a family of relational problems that can be solved with $O(\log n)$ qubits of quantum communication, but requires $\Omega(\sqrt{n})$ bits of classical communication.

One-way communication complexity

Theorem **[Bar-Yossef, Jayram and Kerenidis '08]**

There is a family of relational problems that can be solved with $O(\log n)$ qubits of quantum communication, but requires $\Omega(\sqrt{n})$ bits of classical communication.

- Original proof used information theory methods.
- [Gavinsky et al. '08] improved this to prove a similar separation for a related partial boolean function. Their proof used hypercontractivity.
- [Buhrman, Regev, Scarpa, de Wolf '11] includes a hypercontractive proof of the (simpler) result above.

The Hidden Matching problem

The problem we consider is defined as follows:

- Alice gets $x \in \{0,1\}^n$.
- Bob gets a perfect matching M on $[n]$, i.e. a partition of $\{1, \ldots, n\}$ into pairs.
- Goal: output (i, j, b) such that $(i, j) \in M$ and $b = x_i \oplus x_j$.

The Hidden Matching problem

The problem we consider is defined as follows:

- Alice gets $x \in \{0,1\}^n$.
- Bob gets a perfect matching M on $[n]$, i.e. a partition of $\{1,\ldots,n\}$ into pairs.
- Goal: output (i,j,b) such that $(i,j) \in M$ and $b = x_i \oplus x_j$.

Claim **[Buhrman, Regev, Scarpa, de Wolf '11]**

If x and M are picked uniformly at random, any classical (wlog deterministic) protocol for Hidden Matching with c bits of communication has

$$\Pr[b = x_i \oplus x_j] \leqslant \frac{1}{2} + O\left(\frac{c}{\sqrt{n}}\right).$$

Proof ingredients

- A typical short message from Alice specifies a large subset $A \subseteq \{0,1\}^n$ of her possible inputs.

Proof ingredients

- A typical short message from Alice specifies a large subset $A \subseteq \{0,1\}^n$ of her possible inputs.
- The best Bob can do to guess $x_i \oplus x_j$ is output the value of this function that occurs most often among $x \in A$.

Proof ingredients

- A typical short message from Alice specifies a large subset $A \subseteq \{0,1\}^n$ of her possible inputs.
- The best Bob can do to guess $x_i \oplus x_j$ is output the value of this function that occurs most often among $x \in A$.
- Set $\beta_{ij} = |\mathbb{E}_{x \in A}[(-1)^{x_i + x_j}]|$: Bob's advantage over guessing.

Proof ingredients

- A typical short message from Alice specifies a large subset $A \subseteq \{0,1\}^n$ of her possible inputs.
- The best Bob can do to guess $x_i \oplus x_j$ is output the value of this function that occurs most often among $x \in A$.
- Set $\beta_{ij} = |\mathbb{E}_{x \in A}[(-1)^{x_i + x_j}]|$: Bob's advantage over guessing.

Claim **[Talagrand '96] [Gavinsky et al. '07]**

$$\sum_{i<j} \beta_{ij}^2 = O\left(\left(\log \frac{2^n}{|A|}\right)^2\right).$$

Proof ingredients

- A typical short message from Alice specifies a large subset $A \subseteq \{0,1\}^n$ of her possible inputs.
- The best Bob can do to guess $x_i \oplus x_j$ is output the value of this function that occurs most often among $x \in A$.
- Set $\beta_{ij} = |\mathbb{E}_{x \in A}[(-1)^{x_i + x_j}]|$: Bob's advantage over guessing.

Claim [Talagrand '96] [Gavinsky et al. '07]

$$\sum_{i<j} \beta_{ij}^2 = O\left(\left(\log \frac{2^n}{|A|}\right)^2\right).$$

Proof sketch of claim:

- $\beta_{ij} = |\mathbb{E}_{x \in A}[\chi_{\{i,j\}}(x)]| = (2^n/|A|)\,|\hat{f}(\{i,j\})|$.

Proof ingredients

- A typical short message from Alice specifies a large subset $A \subseteq \{0,1\}^n$ of her possible inputs.
- The best Bob can do to guess $x_i \oplus x_j$ is output the value of this function that occurs most often among $x \in A$.
- Set $\beta_{ij} = |\mathbb{E}_{x \in A}[(-1)^{x_i + x_j}]|$: Bob's advantage over guessing.

Claim [Talagrand '96] [Gavinsky et al. '07]

$$\sum_{i<j} \beta_{ij}^2 = O\left(\left(\log \frac{2^n}{|A|}\right)^2\right).$$

Proof sketch of claim:

- $\beta_{ij} = |\mathbb{E}_{x \in A}[\chi_{\{i,j\}}(x)]| = (2^n/|A|)\,|\hat{f}(\{i,j\})|$.

$$\sum_{i<j} \beta_{ij}^2 = \frac{2^{2n}}{|A|^2} \sum_{i<j} \hat{f}(\{i,j\})^2$$

Proof ingredients

- A typical short message from Alice specifies a large subset $A \subseteq \{0,1\}^n$ of her possible inputs.
- The best Bob can do to guess $x_i \oplus x_j$ is output the value of this function that occurs most often among $x \in A$.
- Set $\beta_{ij} = |\mathbb{E}_{x \in A}[(-1)^{x_i + x_j}]|$: Bob's advantage over guessing.

Claim [Talagrand '96] [Gavinsky et al. '07]

$$\sum_{i<j} \beta_{ij}^2 = O\left(\left(\log \frac{2^n}{|A|}\right)^2\right).$$

Proof sketch of claim:

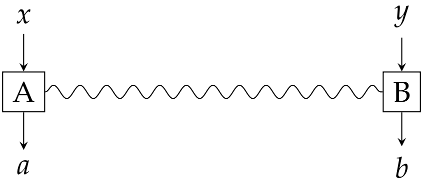
- $\beta_{ij} = |\mathbb{E}_{x \in A}[\chi_{\{i,j\}}(x)]| = (2^n/|A|)\,|\hat{f}(\{i,j\})|$.

$$\sum_{i<j} \beta_{ij}^2 = \frac{2^{2n}}{|A|^2} \sum_{i<j} \hat{f}(\{i,j\})^2 \leqslant \frac{2^{2n}}{\delta^2 |A|^2} \left(\frac{|A|}{2^n}\right)^{2/(1+\delta)}$$

for any $0 \leqslant \delta \leqslant 1$, using KKL. Then minimise over δ.

Nonlocal games

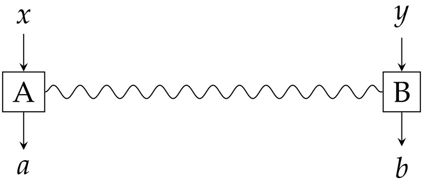
A simple and natural way of exploring the power of quantum correlations is via nonlocal games.



- Alice and Bob get inputs x, y, respectively, drawn from some known distribution π.
- They win the game if their outputs a, b satisfy a known predicate $V(x, y, a, b)$.

Nonlocal games

A simple and natural way of exploring the power of quantum correlations is via nonlocal games.



- Alice and Bob get inputs x, y, respectively, drawn from some known distribution π.
- They win the game if their outputs a, b satisfy a known predicate $V(x, y, a, b)$.
- The players are allowed to communicate before the game starts, to agree a strategy, but cannot communicate during the game.

Nonlocal games

Let the optimal probability of winning G be denoted by:

- $\omega(G)$, if the players are classical;
- $\omega^*(G)$, if the players are allowed to share entanglement.

Nonlocal games

Let the optimal probability of winning G be denoted by:

- $\omega(G)$, if the players are classical;
- $\omega^*(G)$, if the players are allowed to share entanglement.

The CHSH game shows that, for some games, $\omega^*(G) > \omega(G)$.

- Inputs x, y are chosen uniformly from $\{0, 1\}$.
- The players win if their outputs $a, b \in \{0, 1\}$ satisfy $a \oplus b = xy$.

$\omega(\mathrm{CHSH}) = 3/4$, but $\omega^*(\mathrm{CHSH}) = \cos^2(\pi/8) \approx 0.85$.

Nonlocal games

Let the optimal probability of winning G be denoted by:

- $\omega(G)$, if the players are classical;
- $\omega^*(G)$, if the players are allowed to share entanglement.

The CHSH game shows that, for some games, $\omega^*(G) > \omega(G)$.

- Inputs x, y are chosen uniformly from $\{0, 1\}$.
- The players win if their outputs $a, b \in \{0, 1\}$ satisfy $a \oplus b = xy$.

$\omega(\text{CHSH}) = 3/4$, but $\omega^*(\text{CHSH}) = \cos^2(\pi/8) \approx 0.85$.

Question

How large can the gap between $\omega^*(G)$ and $\omega(G)$ be?

Nonlocal games

Theorem **[Buhrman, Regev, Scarpa, de Wolf '11]**

Let n be an integer power of 2. Then there are two nonlocal games HM and KV such that:

- $\omega(\mathrm{HM}) = 1/2 + O((\log n)/\sqrt{n})$, and $\omega^*(\mathrm{HM}) = 1$.

Nonlocal games

Theorem [Buhrman, Regev, Scarpa, de Wolf '11]

Let n be an integer power of 2. Then there are two nonlocal games HM and KV such that:

- $\omega(\mathrm{HM}) = 1/2 + O((\log n)/\sqrt{n})$, and $\omega^*(\mathrm{HM}) = 1$.
- $\omega(\mathrm{KV}) = O(1/n^{1-o(1)})$, and $\omega^*(\mathrm{KV}) \geqslant 4/\log^2 n$.

Nonlocal games

Theorem **[Buhrman, Regev, Scarpa, de Wolf '11]**

Let n be an integer power of 2. Then there are two nonlocal games HM and KV such that:

- $\omega(\mathrm{HM}) = 1/2 + O((\log n)/\sqrt{n})$, and $\omega^*(\mathrm{HM}) = 1$.
- $\omega(\mathrm{KV}) = O(1/n^{1-o(1)})$, and $\omega^*(\mathrm{KV}) \geqslant 4/\log^2 n$.

- The quantum protocols use entangled states on $\mathbb{C}^n \otimes \mathbb{C}^n$.
- These separations are close to optimal.

Nonlocal games

Theorem [Buhrman, Regev, Scarpa, de Wolf '11]

Let n be an integer power of 2. Then there are two nonlocal games HM and KV such that:

- $\omega(\mathrm{HM}) = 1/2 + O((\log n)/\sqrt{n})$, and $\omega^*(\mathrm{HM}) = 1$.
- $\omega(\mathrm{KV}) = O(1/n^{1-o(1)})$, and $\omega^*(\mathrm{KV}) \geqslant 4/\log^2 n$.

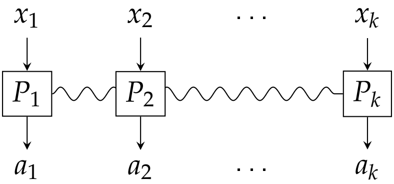
- The quantum protocols use entangled states on $\mathbb{C}^n \otimes \mathbb{C}^n$.
- These separations are close to optimal.

The proofs of the classical lower bounds both use hypercontractivity:

- The HM game is a translation of Hidden Matching to the setting of nonlocal games.
- The KV game is based on work of [Khot and Vishnoi '05] on the unique games conjecture.

Multiplayer nonlocal games

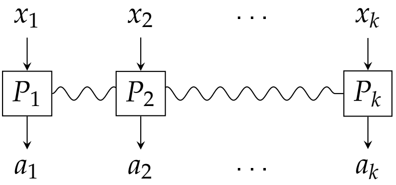
We can generalise the framework of nonlocal games to $k > 2$ players, each receiving an input from $\{1, \ldots, n\}$.



A particularly interesting such class of games is XOR games: games where each output a_i is a single bit, and whether the players win depends only on $a_1 \oplus a_2 \oplus \cdots \oplus a_k$.

Multiplayer nonlocal games

We can generalise the framework of nonlocal games to $k > 2$ players, each receiving an input from $\{1, \ldots, n\}$.



A particularly interesting such class of games is XOR games: games where each output a_i is a single bit, and whether the players win depends only on $a_1 \oplus a_2 \oplus \cdots \oplus a_k$.

Question

What is the hardest k-player XOR game for classical players?

Previously known results

Define the (classical) bias $\beta(G) = \omega(G) - \frac{1}{2}$.

Until recently, there was a big gap between lower and upper bounds on $\min_G \beta(G)$:

- There exists an XOR game G for which $\beta(G) \leqslant n^{-(k-1)/2}$ [Ford and Gál '05].
- Any XOR game G has $\beta(G) \geqslant 2^{-O(k)} n^{-(k-1)/2}$ [Bohnenblust and Hille '31].

Previously known results

Define the (classical) bias $\beta(G) = \omega(G) - \frac{1}{2}$.

Until recently, there was a big gap between lower and upper bounds on $\min_G \beta(G)$:

- There exists an XOR game G for which $\beta(G) \leqslant n^{-(k-1)/2}$ [Ford and Gál '05].
- Any XOR game G has $\beta(G) \geqslant 2^{-O(k)} n^{-(k-1)/2}$ [Bohnenblust and Hille '31].

A recent and substantial improvement:

Theorem [Defant, Popa and Schwarting '10] [Pellegrino and Seoane-Sepúlveda '12]

There exists a universal constant $c > 0$ such that, for any XOR game G as above, $\beta(G) = \Omega(k^{-c} n^{-(k-1)/2})$.

Previously known results

Define the (classical) bias $\beta(G) = \omega(G) - \frac{1}{2}$.

Until recently, there was a big gap between lower and upper bounds on $\min_G \beta(G)$:

- There exists an XOR game G for which $\beta(G) \leqslant n^{-(k-1)/2}$ [Ford and Gál '05].
- Any XOR game G has $\beta(G) \geqslant 2^{-O(k)} n^{-(k-1)/2}$ [Bohnenblust and Hille '31].

A recent and substantial improvement:

Theorem [Defant, Popa and Schwarting '10] [Pellegrino and Seoane-Sepúlveda '12]

There exists a universal constant $c > 0$ such that, for any XOR game G as above, $\beta(G) = \Omega(k^{-c} n^{-(k-1)/2})$.

This result can be proven using hypercontractivity.

XOR games and multilinear forms

A homogeneous polynomial $f : (\mathbb{R}^n)^k \to \mathbb{R}$ is said to be a multilinear form if it can be written as

$$f(x^1, \dots, x^k) = \sum_{i_1,\dots,i_k} \hat{f}_{i_1,\dots,i_k} x^1_{i_1} x^2_{i_2} \dots x^k_{i_k}$$

for some multidimensional array $\hat{f} \in \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$.

XOR games and multilinear forms

A homogeneous polynomial $f : (\mathbb{R}^n)^k \to \mathbb{R}$ is said to be a multilinear form if it can be written as

$$f(x^1, \ldots, x^k) = \sum_{i_1,\ldots,i_k} \hat{f}_{i_1,\ldots,i_k} x^1_{i_1} x^2_{i_2} \ldots x^k_{i_k}$$

for some multidimensional array $\hat{f} \in \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$.

Any XOR game $G = (\pi, V)$ corresponds to a multilinear form f:

$$f(x^1, \ldots, x^k) = \sum_{i_1,\ldots,i_k} \pi_{i_1,\ldots,i_k} V'_{i_1,\ldots,i_k} x^1_{i_1} x^2_{i_2} \ldots x^k_{i_k}.$$

- $x^j_\ell \in \{\pm 1\}$: what the j'th player outputs given input ℓ.
- $V'_{i_1,\ldots,i_k}$: $+1$ or -1 depending on the input.

The bias $\beta(G)$ is precisely $\|f\|_\infty := \max_{x \in \{\pm 1\}^n} |f(x)|$.

A powerful inequality

Bohnenblust-Hille inequality **[BH '31, DPS '10, PS '12]**

For any multilinear form $f : (\mathbb{R}^n)^k \to \mathbb{R}$, and any $p \geqslant 2k/(k+1)$,

$$\|\hat{f}\|_p := \left(\sum_{i_1,\dots,i_k} |\hat{f}_{i_1,\dots,i_k}|^p \right)^{1/p} \leqslant C_k \|f\|_\infty,$$

where C_k may be taken to be $O(k^{\log_2 e}) \approx O(k^{1.45})$.

A powerful inequality

Bohnenblust-Hille inequality [BH '31, DPS '10, PS '12]

For any multilinear form $f : (\mathbb{R}^n)^k \to \mathbb{R}$, and any $p \geqslant 2k/(k+1)$,

$$\|\hat{f}\|_p := \left(\sum_{i_1,\ldots,i_k} |\hat{f}_{i_1,\ldots,i_k}|^p \right)^{1/p} \leqslant C_k \|f\|_\infty,$$

where C_k may be taken to be $O(k^{\log_2 e}) \approx O(k^{1.45})$.

Implies $\beta(G) = \Omega(C_k^{-1} n^{-(k-1)/2})$ by choosing $p = 2k/(k+1)$.

A powerful inequality

Bohnenblust-Hille inequality [BH '31, DPS '10, PS '12]

For any multilinear form $f : (\mathbb{R}^n)^k \to \mathbb{R}$, and any $p \geqslant 2k/(k+1)$,

$$\|\hat{f}\|_p := \left(\sum_{i_1,\dots,i_k} |\hat{f}_{i_1,\dots,i_k}|^p \right)^{1/p} \leqslant C_k \|f\|_\infty,$$

where C_k may be taken to be $O(k^{\log_2 e}) \approx O(k^{1.45})$.

Implies $\beta(G) = \Omega(C_k^{-1} n^{-(k-1)/2})$ by choosing $p = 2k/(k+1)$.

Proof is by a delicate induction on k, for k a power of 2.

- Inductive step goes from $k \to k/2$ via Hölder's inequality, relating $\|\hat{f}\|_{2k/(k+1)}$ to ℓ_2 norms of restricted versions of f.
- Hypercontractivity lets us relate ℓ_2 norms to $\frac{2k}{k+2}$-norms.

Moving to the real n-sphere

- Let $S^n := \{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$ be the real n-sphere.

Moving to the real n-sphere

- Let $S^n := \{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$ be the real n-sphere.
- Any smooth function $f : S^n \to \mathbb{R}$ can be expanded in terms of spherical harmonics: $f = \sum_k Y_k$, for degree k polynomials Y_k such that

$$\int Y_j(x) Y_k(x) dx = 0$$

for $j \neq k$.

Moving to the real n-sphere

- Let $S^n := \{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$ be the real n-sphere.

- Any smooth function $f : S^n \to \mathbb{R}$ can be expanded in terms of spherical harmonics: $f = \sum_k Y_k$, for degree k polynomials Y_k such that

$$\int Y_j(x) Y_k(x) dx = 0$$

  for $j \neq k$.

- Set $\|f\|_p = \left(\int |f(x)|^p dx\right)^{1/p}$.

Moving to the real n-sphere

- Let $S^n := \{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$ be the real n-sphere.

- Any smooth function $f : S^n \to \mathbb{R}$ can be expanded in terms of spherical harmonics: $f = \sum_k Y_k$, for degree k polynomials Y_k such that

$$\int Y_j(x) Y_k(x) dx = 0$$

  for $j \neq k$.

- Set $\|f\|_p = \left(\int |f(x)|^p dx\right)^{1/p}$.

- Parseval's equality: $\|f\|_2^2 = \sum_k \|Y_k\|_2^2$.

Hypercontractivity on the real n-sphere

- For $\rho \in [0,1]$, define the Poisson semigroup P_ρ as follows:

$$(P_\rho f)(x) = \sum_{k \geqslant 0} \rho^k Y_k(x).$$

Hypercontractivity on the real n-sphere

- For $\rho \in [0, 1]$, define the Poisson semigroup P_ρ as follows:

$$(P_\rho f)(x) = \sum_{k \geqslant 0} \rho^k Y_k(x).$$

- Alternatively:

$$(P_\rho f)(x) = (1 - \rho^2) \int |x - \rho y|^{-(n+1)} f(y) dy$$

Hypercontractivity on the real n-sphere

- For $\rho \in [0, 1]$, define the Poisson semigroup P_ρ as follows:

$$(P_\rho f)(x) = \sum_{k \geqslant 0} \rho^k Y_k(x).$$

- Alternatively:

$$(P_\rho f)(x) = (1 - \rho^2) \int |x - \rho y|^{-(n+1)} f(y) dy$$

Hypercontractive inequality [Beckner '92]

For any $f : S^n \to \mathbb{R}$, and any p and q such that $1 \leqslant p \leqslant q \leqslant \infty$ and $\rho \leqslant \sqrt{\frac{p-1}{q-1}}$,

$$\|P_\rho f\|_q \leqslant \|f\|_p.$$

Hypercontractivity on the real n-sphere

As this framework is so similar to the case of the boolean cube, many corollaries carry across without change. For example:

Corollary

For any degree d polynomial $f : S^n \to \mathbb{R}$, and any $q \geqslant 2$,

$$\|f\|_q \leqslant (q-1)^{d/2} \|f\|_2.$$

Proof is exactly the same as on the boolean cube.

Communication complexity separation

- We have seen that one-way quantum communication is more powerful than one-way classical communication.
- What about one-way quantum vs. two-way classical?

Communication complexity separation

- We have seen that one-way quantum communication is more powerful than one-way classical communication.
- What about one-way quantum vs. two-way classical?

Theorem [Klartag+Regev '11]

There is a partial function which can be computed with an $O(\log n)$-qubit message from Alice to Bob, but for which every classical two-way protocol requires $\Omega(n^{1/3})$ bits of communication.

Communication complexity separation

- We have seen that one-way quantum communication is more powerful than one-way classical communication.
- What about one-way quantum vs. two-way classical?

Theorem [Klartag+Regev '11]

There is a partial function which can be computed with an $O(\log n)$-qubit message from Alice to Bob, but for which every classical two-way protocol requires $\Omega(n^{1/3})$ bits of communication.

The problem:

- Alice gets a unit vector $v \in S^{n-1}$, Bob gets a subspace $H \subset \mathbb{R}^n$ of dimension $n/2$.
- Promise: either $v \in H$ or $v \in H^\perp$.
- Task: determine which is the case.

Classical communication lower bound

Many technical steps…

Classical communication lower bound

Many technical steps... one key lemma:

Lemma (variant of [Klartag+Regev '11])

Assume $f : S^{n-1} \to \mathbb{R}$ has $\|f\|_1 = 1$, $\|f\|_\infty = M$. Expand $f = \sum_k Y_k$. Then

$$\|Y_k\|_2 \leqslant \left(\frac{2e \ln M}{k}\right)^{k/2}.$$

Classical communication lower bound

Many technical steps… one key lemma:

Lemma (variant of [Klartag+Regev '11])

Assume $f : S^{n-1} \to \mathbb{R}$ has $\|f\|_1 = 1$, $\|f\|_\infty = M$. Expand $f = \sum_k Y_k$. Then

$$\|Y_k\|_2 \leqslant \left(\frac{2e \ln M}{k}\right)^{k/2}.$$

Proof:

$$\|Y_k\|_2 = \|T_\rho^{-1} T_\rho Y_k\|_2 = \rho^{-k} \|T_\rho Y_k\|_2 \leqslant \rho^{-k} \|T_\rho f\|_2 \leqslant \rho^{-k} \|f\|_p$$

for $p = 1 + \rho^2$. Observing $\|f\|_p \leqslant M^{p-1}$ and optimising over p gives the claimed result.

Classical communication lower bound

Many technical steps... one key lemma:

Lemma (variant of [Klartag+Regev '11])

Assume $f : S^{n-1} \to \mathbb{R}$ has $\|f\|_1 = 1$, $\|f\|_\infty = M$. Expand $f = \sum_k Y_k$. Then

$$\|Y_k\|_2 \leqslant \left(\frac{2e \ln M}{k}\right)^{k/2}.$$

Proof:

$$\|Y_k\|_2 = \|T_\rho^{-1} T_\rho Y_k\|_2 = \rho^{-k} \|T_\rho Y_k\|_2 \leqslant \rho^{-k} \|T_\rho f\|_2 \leqslant \rho^{-k} \|f\|_p$$

for $p = 1 + \rho^2$. Observing $\|f\|_p \leqslant M^{p-1}$ and optimising over p gives the claimed result.

- [Klartag+Regev '11] used a different noise operator and a different hypercontractive inequality, but the eventual result is essentially the same.

Biases of local measurements

- Imagine we are given a quantum state promised to be either ρ or σ, with equal probability of each.
- We want to determine which state we have, but are forced to use just one fixed measurement for all ρ, σ.

Biases of local measurements

- Imagine we are given a quantum state promised to be either ρ or σ, with equal probability of each.
- We want to determine which state we have, but are forced to use just one fixed measurement for all ρ, σ.
- We use the uniform POVM U putting equal weight on each state $|\psi\rangle \in \mathbb{C}^n$.

Biases of local measurements

- Imagine we are given a quantum state promised to be either ρ or σ, with equal probability of each.
- We want to determine which state we have, but are forced to use just one fixed measurement for all ρ, σ.
- We use the uniform POVM U putting equal weight on each state $|\psi\rangle \in \mathbb{C}^n$.
- Set $\Delta = (\rho - \sigma)/2$. Then the optimal success probability is

$$\frac{1}{2}\left(1 + n \int |\langle\psi|\Delta|\psi\rangle|\, d\psi\right) =: \frac{1}{2}\left(1 + \|\Delta\|_U\right).$$

Biases of local measurements

- Imagine we are given a quantum state promised to be either ρ or σ, with equal probability of each.
- We want to determine which state we have, but are forced to use just one fixed measurement for all ρ, σ.
- We use the uniform POVM U putting equal weight on each state $|\psi\rangle \in \mathbb{C}^n$.
- Set $\Delta = (\rho - \sigma)/2$. Then the optimal success probability is

$$\frac{1}{2}\left(1 + n\int |\langle\psi|\Delta|\psi\rangle|\, d\psi\right) =: \frac{1}{2}\left(1 + \|\Delta\|_U\right).$$

Theorem [Ambainis+Emerson '07, Matthews et al. '09]

There is a universal constant C such that

$$\|\Delta\|_U \geqslant C\sqrt{\operatorname{tr}\Delta^2}.$$

Proving this using hypercontractivity

The proof is based on the "fourth moment method":

$$\|\Delta\|_U = n\int |\langle\psi|\Delta|\psi\rangle|\, d\psi \geqslant n\frac{\left(\int\langle\psi|\Delta|\psi\rangle^2 d\psi\right)^{3/2}}{\left(\int\langle\psi|\Delta|\psi\rangle^4 d\psi\right)^{1/2}}.$$

Proving this using hypercontractivity

The proof is based on the "fourth moment method":

$$\|\Delta\|_U = n\int |\langle\psi|\Delta|\psi\rangle|\, d\psi \geqslant n\frac{\left(\int \langle\psi|\Delta|\psi\rangle^2 d\psi\right)^{3/2}}{\left(\int \langle\psi|\Delta|\psi\rangle^4 d\psi\right)^{1/2}}.$$

- It's easy to compute

$$\int \langle\psi|\Delta|\psi\rangle^2 d\psi = \operatorname{tr}\left(\int d\psi |\psi\rangle\langle\psi|^{\otimes 2}\right)\Delta^{\otimes 2} = \frac{\operatorname{tr}\Delta^2}{n(n+1)}.$$

Proving this using hypercontractivity

The proof is based on the "fourth moment method":

$$\|\Delta\|_U = n \int |\langle\psi|\Delta|\psi\rangle| \, d\psi \geqslant n \frac{\left(\int \langle\psi|\Delta|\psi\rangle^2 d\psi\right)^{3/2}}{\left(\int \langle\psi|\Delta|\psi\rangle^4 d\psi\right)^{1/2}}.$$

- It's easy to compute

$$\int \langle\psi|\Delta|\psi\rangle^2 d\psi = \operatorname{tr}\left(\int d\psi |\psi\rangle\langle\psi|^{\otimes 2}\right) \Delta^{\otimes 2} = \frac{\operatorname{tr} \Delta^2}{n(n+1)}.$$

- To bound the denominator, we use hypercontractivity.

Proving this using hypercontractivity

We go from the complex to the real unit sphere:

- Associate $|\psi\rangle$ with $\xi \in S^{2n-1}$.

Proving this using hypercontractivity

We go from the complex to the real unit sphere:

- Associate $|\psi\rangle$ with $\xi \in S^{2n-1}$.
- Claim: $f(\xi) = \langle\psi|\Delta|\psi\rangle$ is a degree-2 polynomial in ξ.

Proving this using hypercontractivity

We go from the complex to the real unit sphere:

- Associate $|\psi\rangle$ with $\xi \in S^{2n-1}$.
- Claim: $f(\xi) = \langle\psi|\Delta|\psi\rangle$ is a degree-2 polynomial in ξ.

So, by hypercontractivity,

$$\|f\|_p = \left(\int |\langle\psi|\Delta|\psi\rangle|^p\right)^{1/p} \leqslant (p-1)\left(\int \langle\psi|\Delta|\psi\rangle^2\right)^{1/2}.$$

Proving this using hypercontractivity

We go from the complex to the real unit sphere:

- Associate $|\psi\rangle$ with $\xi \in S^{2n-1}$.
- Claim: $f(\xi) = \langle\psi|\Delta|\psi\rangle$ is a degree-2 polynomial in ξ.

So, by hypercontractivity,

$$\|f\|_p = \left(\int |\langle\psi|\Delta|\psi\rangle|^p\right)^{1/p} \leqslant (p-1)\left(\int \langle\psi|\Delta|\psi\rangle^2\right)^{1/2}.$$

Taking $p = 4$ and substituting in gives an overall bound

$$\|\Delta\|_U \geqslant \left(\frac{1}{9} - o(1)\right)\sqrt{\operatorname{tr}\Delta^2}.$$

The multipartite case

What about if ρ, σ are multipartite states on $(\mathbb{C}^n)^{\otimes k}$, and we use as our measurement the uniform POVM on each party separately?

The multipartite case

What about if ρ, σ are multipartite states on $(\mathbb{C}^n)^{\otimes k}$, and we use as our measurement the uniform POVM on each party separately?

Theorem **[Matthews et al. '09, Lancien+Winter '13]**

$$\|\Delta\|_U \geqslant C^{k/2} \left(\sum_{S \subseteq [k]} \mathrm{tr}[(\mathrm{tr}_S \, \Delta)^2] \right)^{1/2} .$$

The multipartite case

What about if ρ, σ are multipartite states on $(\mathbb{C}^n)^{\otimes k}$, and we use as our measurement the uniform POVM on each party separately?

Theorem [Matthews et al. '09, Lancien+Winter '13]

$$\|\Delta\|_U \geqslant C^{k/2} \left(\sum_{S \subseteq [k]} \operatorname{tr}[(\operatorname{tr}_S \Delta)^2] \right)^{1/2}.$$

Claim: hypercontractivity gives us this result for free using multiplicativity of the $L_p \to L_q$ norm!

The multipartite case

What about if ρ, σ are multipartite states on $(\mathbb{C}^n)^{\otimes k}$, and we use as our measurement the uniform POVM on each party separately?

Theorem **[Matthews et al. '09, Lancien+Winter '13]**

$$\|\Delta\|_U \geqslant C^{k/2} \left(\sum_{S \subseteq [k]} \operatorname{tr}[(\operatorname{tr}_S \Delta)^2] \right)^{1/2}.$$

Claim: hypercontractivity gives us this result for free using multiplicativity of the $L_p \to L_q$ norm!

Compare the original proof…

In the particular case of all the seven permutations in $\mathfrak{A}$, $\sigma_{\mathcal{A}} = \mathrm{id}$, $\sigma_{\mathcal{B}} = (14)$, $\sigma_{\mathcal{C}} = (23)$, $\sigma_{\mathcal{D}} = (1234)$, $\sigma_{\mathcal{E}} = (1432)$, $\sigma_{\mathcal{F}} = (12)(34)$ and $\sigma_{\mathcal{G}} = (14)(23)$, this becomes

$$\begin{aligned}
\operatorname{Tr} \Delta^{\otimes 4}(U_{\sigma_{\mathcal{A}}} \otimes \cdots \otimes U_{\sigma_{\mathcal{G}}}) &= \sum_{\substack{a_1,\dots,g_1\\ a_2,\dots,g_2\\ a_3,\dots,g_3\\ a_4,\dots,g_4}} \Delta^{a_1b_4c_1d_2e_4f_2g_4}_{a_1b_1c_1d_1e_1f_1g_1} \Delta^{a_2b_2c_3d_3e_1f_1g_3}_{a_2b_2c_2d_2e_2f_2g_2} \Delta^{a_3b_3c_2d_4e_2f_4g_2}_{a_3b_3c_3d_3e_3f_3g_3} \Delta^{a_4b_1c_4d_1e_3f_3g_1}_{a_4b_4c_4d_4e_4f_4g_4} \\
&= \sum_{\substack{b_1,d_1,\dots,g_1\\ c_2,\dots,g_2\\ c_3,\dots,g_3\\ b_4,d_4,\dots,g_4}} \left[(\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{C}} \Delta)^{\Gamma_{\mathcal{E}}}\right]^{b_4d_2e_1f_2g_4}_{b_1d_1e_4f_1g_1} \left[(\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{B}} \Delta)^{\Gamma_{\mathcal{E}}}\right]^{c_3d_3e_2f_1g_3}_{c_2d_2e_1f_2g_2} \\
&\qquad \times \left[(\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{B}} \Delta)^{\Gamma_{\mathcal{E}}}\right]^{c_2d_4e_3f_4g_2}_{c_3d_3e_2f_3g_3} \left[(\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{C}} \Delta)^{\Gamma_{\mathcal{E}}}\right]^{b_1d_1e_4f_3g_1}_{b_4d_4e_3f_4g_4},
\end{aligned}$$

where $\Gamma_{\mathcal{E}}$ denotes the partial transposition on $\mathcal{E}$.

We can rewrite this using the maximally entangled $\Phi_{\mathcal{F}\otimes\mathcal{F}} = \sum_{ff'} |ff\rangle\langle f'f'|$:

Letting $\mathcal{J} := \mathcal{C} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{G}$, $P := (\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{B}} \Delta)^{\Gamma_{\mathcal{E}}}$ and $R := (P \otimes \mathbb{1}_{\mathcal{F}})(\mathbb{1}_{\mathcal{J}} \otimes \Phi_{\mathcal{F}\otimes\mathcal{F}})(P \otimes \mathbb{1}_{\mathcal{F}})$, we notice that, for all $j, j', f, f', \widetilde{f}, \widetilde{f}'$:

$$R^{j',f',\widetilde{f}'}_{j,f,\widetilde{f}} = \sum_{\substack{j'',j'''\\ f'',f'''\\ \widetilde{f}'',\widetilde{f}'''}} \left(P^{j'',f''}_{j,f} \delta_{\widetilde{f}''=\widetilde{f}}\right) \left(\delta_{j'''=j''} \delta_{\widetilde{f}''=f'',\widetilde{f}'''=f'''}\right) \left(P^{j',f'}_{j''',f'''} \delta_{\widetilde{f}'''=\widetilde{f}'}\right) = \sum_{j''} P^{j'',\widetilde{f}}_{j,f} P^{j',f'}_{j'',\widetilde{f}'}.$$

Likewise, letting $\mathcal{K} := \mathcal{B}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{G}$, $Q := (\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{C}} \Delta)^{\Gamma_{\mathcal{E}}}$ and $S := (Q\otimes\mathbb{1}_{\mathcal{F}})(\mathbb{1}_{\mathcal{K}}\otimes\Phi_{\mathcal{F}\otimes\mathcal{F}})(Q\otimes\mathbb{1}_{\mathcal{F}})$, we have for all $k, k', f, f', \widetilde{f}, \widetilde{f}'$:

$$S^{k',f,\widetilde{f}}_{k,f',\widetilde{f}'} = \sum_{k''} Q^{k'',\widetilde{f}'}_{k,f'} Q^{k',f}_{k'',\widetilde{f}}.$$

We now just have to make the following identifications:

- $j := (c_2, d_2, e_1, g_2)$, $j' := (c_2, d_4, e_3, g_2)$, $j'' := (c_3, d_3, e_2, g_3)$,
- $k := (b_4, d_4, e_3, g_4)$, $k' := (b_4, d_2, e_1, g_4)$, $k'' := (b_1, d_1, e_4, g_1)$,
- $f := f_2$, $f' := f_4$, $\widetilde{f} := f_1$, $\widetilde{f}' := f_3$,

and to notice that we can actually sum over j'' and k'' independently. We thus get:

$$\begin{aligned}
\operatorname{Tr} \Delta^{\otimes 4}(U_{\sigma_{\mathcal{A}}} \otimes \cdots \otimes U_{\sigma_{\mathcal{G}}}) &= \sum_{\substack{e_1,f_1\\ c_2,d_2,f_2,g_2\\ e_3,f_3\\ b_4,d_4,f_4,g_4}} R^{c_2,d_4,e_3,g_2,f_4,f_3}_{c_2,d_2,e_1,g_2,f_2,f_1} S^{b_4,d_2,e_1,g_4,f_2,f_1}_{b_4,d_4,e_3,g_4,f_4,f_3} \\
&= \sum_{\substack{e_1,f_1\\ d_2,f_2\\ e_3,f_3\\ d_4,f_4}} (\operatorname{Tr}_{\mathcal{C}\otimes\mathcal{G}} R)^{d_4,e_3,f_4,f_3}_{d_2,e_1,f_2,f_1} (\operatorname{Tr}_{\mathcal{B}\otimes\mathcal{G}} S)^{d_2,e_1,f_2,f_1}_{d_4,e_3,f_4,f_3} \\
&= \operatorname{Tr}_{\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{F}} (\operatorname{Tr}_{\mathcal{C}\otimes\mathcal{G}} R) (\operatorname{Tr}_{\mathcal{B}\otimes\mathcal{G}} S).
\end{aligned}$$

Defining $\widetilde{P} := (P\otimes\mathbb{1}_{\mathcal{F}})(\mathbb{1}_{\mathcal{J}}\otimes\sum_f |ff\rangle)$ and $\widetilde{Q} := (Q\otimes\mathbb{1}_{\mathcal{F}})(\mathbb{1}_{\mathcal{J}}\otimes\sum_f |ff\rangle)$, we see that $R = \widetilde{P}\widetilde{P}^\dagger$ and $S = \widetilde{Q}\widetilde{Q}^\dagger$. Hence R and S are positive semidefinite, and so are $\operatorname{Tr}_{\mathcal{C}\otimes\mathcal{G}} R$ and $\operatorname{Tr}_{\mathcal{B}\otimes\mathcal{G}} S$. Thus, using the fact that, for positive semidefinite V and W, $\operatorname{Tr} VW \le (\operatorname{Tr} V)(\operatorname{Tr} W)$, we obtain

$$\operatorname{Tr}_{\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{F}} \left[(\operatorname{Tr}_{\mathcal{C}\otimes\mathcal{G}} R)(\operatorname{Tr}_{\mathcal{B}\otimes\mathcal{G}} S)\right] \le (\operatorname{Tr}_{\mathcal{C}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{F}\otimes\mathcal{G}} R)(\operatorname{Tr}_{\mathcal{B}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{F}\otimes\mathcal{G}} S).$$

On right hand side,

$$\begin{aligned}
\operatorname{Tr} R &= \operatorname{Tr}_{\mathcal{C}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}} P^2 \\
&= \operatorname{Tr}_{\mathcal{C}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}} \left((\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{B}} \Delta)^{\Gamma_{\mathcal{E}}}\right)^2 \\
&= \operatorname{Tr}_{\mathcal{C}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}} (\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{B}} \Delta)^2,
\end{aligned}$$

and likewise, $\operatorname{Tr} S = \operatorname{Tr}_{\mathcal{B}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}} (\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{C}} \Delta)^2$. So, we eventually arrive at

$$\operatorname{Tr} \Delta^{\otimes 4}(U_{\sigma_{\mathcal{A}}} \otimes \cdots \otimes U_{\sigma_{\mathcal{G}}}) \le \left[\operatorname{Tr}_{\mathcal{C}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}} (\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{B}} \Delta)^2\right] \left[\operatorname{Tr}_{\mathcal{B}\otimes\mathcal{D}\otimes\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}} (\operatorname{Tr}_{\mathcal{A}\otimes\mathcal{C}} \Delta)^2\right]. \tag{A2}$$

With this inequality as a tool, we can now return to our initial problem: For all $\underline{\pi} \in \mathfrak{A}^K = \{\mathrm{id}, (14), (23), (1234), (1432), (12)(34), (14)(23)\}^K$, we can define the following factors of the global Hilbert space $\mathcal{H}$:

$$\begin{aligned}
&\mathcal{A}(\underline{\pi}) := \bigotimes_{j \text{ s.t. } \pi_j = \mathrm{id}} \mathcal{H}_j, \quad \mathcal{B}(\underline{\pi}) := \bigotimes_{j \text{ s.t. } \pi_j = (14)} \mathcal{H}_j, \quad \mathcal{C}(\underline{\pi}) := \bigotimes_{j \text{ s.t. } \pi_j = (23)} \mathcal{H}_j, \\
&\mathcal{D}(\underline{\pi}) := \bigotimes_{j \text{ s.t. } \pi_j = (1234)} \mathcal{H}_j, \quad \mathcal{E}(\underline{\pi}) := \bigotimes_{j \text{ s.t. } \pi_j = (1432)} \mathcal{H}_j, \\
&\mathcal{F}(\underline{\pi}) := \bigotimes_{j \text{ s.t. } \pi_j = (12)(34)} \mathcal{H}_j, \quad \mathcal{G}(\underline{\pi}) := \bigotimes_{j \text{ s.t. } \pi_j = (14)(23)} \mathcal{H}_j,
\end{aligned}$$

so that clearly, $\mathcal{H} = \mathcal{A}(\underline{\pi}) \otimes \mathcal{B}(\underline{\pi}) \otimes \mathcal{C}(\underline{\pi}) \otimes \mathcal{D}(\underline{\pi}) \otimes \mathcal{E}(\underline{\pi}) \otimes \mathcal{F}(\underline{\pi}) \otimes \mathcal{G}(\underline{\pi})$. Hence, using successively the two inequalities (A1) and (A2), we have:

$$\begin{aligned}
\sum_{\underline{\sigma}\in\mathfrak{S}_4^K} \operatorname{Tr}\left(\Delta^{\otimes 4} U_{\underline{\sigma}}\right) &\le \sum_{\underline{\sigma}\in\mathfrak{S}_4^K} \left\{\frac{1}{2} \operatorname{Tr}\left(\Delta^{\otimes 4} U_{\underline{\sigma}^L}\right) + \frac{1}{2} \operatorname{Tr}\left(\Delta^{\otimes 4} U_{\underline{\sigma}^R}\right)\right\} \\
&\le \sum_{\underline{\sigma}\in\mathfrak{S}_4^K} \left\{\frac{1}{2}\left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^L)\otimes\mathcal{B}(\underline{\sigma}^L)} \Delta\right)^2\right] \left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^L)\otimes\mathcal{C}(\underline{\sigma}^L)} \Delta\right)^2\right]\right. \\
&\qquad \left. + \frac{1}{2}\left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^R)\otimes\mathcal{B}(\underline{\sigma}^R)} \Delta\right)^2\right] \left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^R)\otimes\mathcal{C}(\underline{\sigma}^R)} \Delta\right)^2\right]\right\} \\
&= \sum_{\underline{\sigma}\in\mathfrak{S}_4^K} \left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^L)\otimes\mathcal{B}(\underline{\sigma}^L)} \Delta\right)^2\right] \left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^L)\otimes\mathcal{C}(\underline{\sigma}^L)} \Delta\right)^2\right] \\
&\le \sum_{\underline{\sigma}\in\mathfrak{S}_4^K} \left\{\frac{1}{2}\left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^L)\otimes\mathcal{B}(\underline{\sigma}^L)} \Delta\right)^2\right]^2 + \frac{1}{2}\left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^L)\otimes\mathcal{C}(\underline{\sigma}^L)} \Delta\right)^2\right]^2\right\} \\
&= \sum_{\underline{\sigma}\in\mathfrak{S}_4^K} \left[\operatorname{Tr}\left(\operatorname{Tr}_{\mathcal{A}(\underline{\sigma}^L)\otimes\mathcal{B}(\underline{\sigma}^L)} \Delta\right)^2\right]^2,
\end{aligned}$$

where in the last lines we have made use of the symmetry between $\underline{\sigma}^L$ and $\underline{\sigma}^R$ on the one hand, and that between $\mathcal{B}(\underline{\sigma}^L)$ and $\mathcal{C}(\underline{\sigma}^L)$ on the other, when $\underline{\sigma}$ ranges over $\mathfrak{S}_4^K$.

Noncommutative generalisations

There are at least two sensible ways in which one could generalise the hypercontractive inequality on the boolean cube to a noncommutative setting:

- Matrix-valued functions on the boolean cube:

$$f : \{0,1\}^n \to M_d$$

- Linear operators on $(\mathbb{C}^2)^{\otimes n}$ (the space of n qubits).

Both of these ideas work and lead to interesting consequences.

Matrix-valued functions

The hypercontractive inequality when $q = 2$:

$$\sum_{S \subseteq [n]} (p-1)^{|S|} \hat{f}(S)^2 \leqslant \left(\frac{1}{2^n} \sum_{x \in \{0,1\}^n} |f(x)|^p \right)^{2/p}$$

for any $1 \leqslant p \leqslant 2$.

Matrix-valued functions

The hypercontractive inequality when $q = 2$:

$$\sum_{S \subseteq [n]} (p-1)^{|S|} \hat{f}(S)^2 \leqslant \left(\frac{1}{2^n} \sum_{x \in \{0,1\}^n} |f(x)|^p \right)^{2/p}$$

for any $1 \leqslant p \leqslant 2$. In the matrix-valued case we have:

Theorem **[Ben-Aroya, Regev and de Wolf '08]**

$$\sum_{S \subseteq [n]} (p-1)^{|S|} \|\hat{f}(S)\|_p^2 \leqslant \left(\frac{1}{2^n} \sum_{x \in \{0,1\}^n} \|f(x)\|_p^p \right)^{2/p}$$

for any $1 \leqslant p \leqslant 2$, where $\| \cdot \|_p$ is the Schatten p-norm and

$$\hat{f}(S) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x) \chi_S(x)$$

are now matrices.

Applications

One example: proving limitations on quantum random access codes [Ben-Aroya, Regev, de Wolf '08].

- We want to encode $x \in \{0,1\}^n$ in a state $\rho \in M_{2^m}$ such that we can recover any k of the n bits with high probability.

Applications

One example: proving limitations on quantum random access codes [Ben-Aroya, Regev, de Wolf '08].

- We want to encode $x \in \{0,1\}^n$ in a state $\rho \in M_{2^m}$ such that we can recover any k of the n bits with high probability.
- Claim: even predicting $\bigoplus_{i \in S} x_i$, for an arbitrary k-subset S, is difficult on average.

Applications

One example: proving limitations on quantum random access codes [Ben-Aroya, Regev, de Wolf '08].

- We want to encode $x \in \{0,1\}^n$ in a state $\rho \in M_{2^m}$ such that we can recover any k of the n bits with high probability.
- Claim: even predicting $\bigoplus_{i \in S} x_i$, for an arbitrary k-subset S, is difficult on average.
- If $f : \{0,1\}^n \to M_{2^m}$ is our encoding function, the success probability is controlled by

$$\|\mathbb{E}_{x, \bigoplus_{i \in S} x_i = 0}[M_x] - \mathbb{E}_{x, \bigoplus_{i \in S} x_i = 1}[M_x]\|_1 = \|\hat{f}(S)\|_1.$$

Applications

One example: proving limitations on quantum random access codes [Ben-Aroya, Regev, de Wolf '08].

- We want to encode $x \in \{0,1\}^n$ in a state $\rho \in M_{2^m}$ such that we can recover any k of the n bits with high probability.
- Claim: even predicting $\bigoplus_{i \in S} x_i$, for an arbitrary k-subset S, is difficult on average.
- If $f : \{0,1\}^n \to M_{2^m}$ is our encoding function, the success probability is controlled by

$$\|\mathbb{E}_{x, \bigoplus_{i \in S} x_i = 0}[M_x] - \mathbb{E}_{x, \bigoplus_{i \in S} x_i = 1}[M_x]\|_1 = \|\hat{f}(S)\|_1.$$

- Claim:

$$\mathbb{E}_{S \sim \binom{[n]}{k}} \left[\|\hat{f}(S)\|_1 \right] \leqslant C \left(\frac{C' m}{n} \right)^{k/2}.$$

Proof: use hypercontractive inequality with carefully chosen p.

A different notion of noncommutativity

- Instead of functions $f : \{0,1\}^n \to \mathbb{R}$, we consider Hermitian operators on the space of n qubits.
- Then a natural generalisation of the noise operator on one bit is the qubit depolarising channel:

$$\mathcal{D}_\rho(M) = (1-\rho)(\operatorname{tr} M)\frac{I}{2} + \rho M.$$

A different notion of noncommutativity

- Instead of functions $f : \{0,1\}^n \to \mathbb{R}$, we consider Hermitian operators on the space of n qubits.
- Then a natural generalisation of the noise operator on one bit is the qubit depolarising channel:

$$\mathcal{D}_\rho(M) = (1-\rho)(\operatorname{tr} M)\frac{I}{2} + \rho M.$$

Hypercontractive inequality [King '12] [AM+Osborne '10]

For any Hermitian operator $M \in \mathcal{B}((\mathbb{C}^2)^{\otimes n})$, and any p and q such that $1 \leqslant p \leqslant q \leqslant \infty$ and $\rho \leqslant \sqrt{\frac{p-1}{q-1}}$,

$$\|\mathcal{D}_\rho^{\otimes n} M\|_q \leqslant \|M\|_p.$$

A different notion of noncommutativity

- Instead of functions $f : \{0,1\}^n \to \mathbb{R}$, we consider Hermitian operators on the space of n qubits.
- Then a natural generalisation of the noise operator on one bit is the qubit depolarising channel:

$$\mathcal{D}_\rho(M) = (1-\rho)(\operatorname{tr} M)\frac{I}{2} + \rho M.$$

Hypercontractive inequality [King '12] [AM+Osborne '10]

For any Hermitian operator $M \in \mathcal{B}((\mathbb{C}^2)^{\otimes n})$, and any p and q such that $1 \leqslant p \leqslant q \leqslant \infty$ and $\rho \leqslant \sqrt{\frac{p-1}{q-1}}$,

$$\|\mathcal{D}_\rho^{\otimes n} M\|_q \leqslant \|M\|_p.$$

([King '12] actually proves hypercontractivity for all semigroups of unital qubit channels)

Application: norm and tail bounds

- Many (though not all!) of the corollaries of hypercontractivity on the boolean cube go through immediately.

Application: norm and tail bounds

- Many (though not all!) of the corollaries of hypercontractivity on the boolean cube go through immediately.
- The right analogue of degree d polynomials turns out to be d-local operators on $(\mathbb{C}^2)^{\otimes n}$.

Norm and tail bounds

Let M be a d-local Hermitian operator on n qubits such that $\|M\|_2 = 1$. Then:

- $\|M\|_q \leqslant (q-1)^{d/2}$ for all $q \geqslant 2$.

Application: norm and tail bounds

- Many (though not all!) of the corollaries of hypercontractivity on the boolean cube go through immediately.
- The right analogue of degree d polynomials turns out to be d-local operators on $(\mathbb{C}^2)^{\otimes n}$.

Norm and tail bounds

Let M be a d-local Hermitian operator on n qubits such that $\|M\|_2 = 1$. Then:

- $\|M\|_q \leqslant (q-1)^{d/2}$ for all $q \geqslant 2$.
- $$\frac{|\{i : |\lambda_i| \geqslant t\}|}{2^n} \leqslant \exp(-dt^{2/d}/(2e)).$$

A weaker (but much simpler) version of quantum central limit theorems, e.g. [Hartmann et al. '04].

Application: norm and tail bounds

- Many (though not all!) of the corollaries of hypercontractivity on the boolean cube go through immediately.
- The right analogue of degree d polynomials turns out to be d-local operators on $(\mathbb{C}^2)^{\otimes n}$.

Norm and tail bounds

Let M be a d-local Hermitian operator on n qubits such that $\|M\|_2 = 1$. Then:

- $\|M\|_q \leqslant (q-1)^{d/2}$ for all $q \geqslant 2$.
- $$\frac{|\{i : |\lambda_i| \geqslant t\}|}{2^n} \leqslant \exp(-dt^{2/d}/(2e)).$$

A weaker (but much simpler) version of quantum central limit theorems, e.g. [Hartmann et al. '04].

- Question: is there a quantum version of the KKL theorem?

Application: rapid mixing

- A quantum Markov process is a family of channels of the form
$$\mathcal{E}_t(\rho) = e^{t\mathcal{L}}.$$

- We want to find the mixing time of $\mathcal{E}$: the minimum t such that
$$\|\mathcal{E}_t(\rho) - \sigma\|_1 \leqslant \epsilon$$
for all ρ, where $\sigma = \lim_{t\to\infty} \mathcal{E}_t(\rho)$.

Application: rapid mixing

- A quantum Markov process is a family of channels of the form
$$\mathcal{E}_t(\rho) = e^{t\mathcal{L}}.$$

- We want to find the mixing time of $\mathcal{E}$: the minimum t such that
$$\|\mathcal{E}_t(\rho) - \sigma\|_1 \leqslant \epsilon$$
for all ρ, where $\sigma = \lim_{t\to\infty} \mathcal{E}_t(\rho)$.

[Kastoryano+Temme '13]: hypercontractive ($\equiv$ log-Sobolev) inequalities imply significantly improved mixing time bounds.

- e.g. an exponential improvement over a more naïve bound for the d-dimensional depolarising channel.

Conclusions

A little bit of noise can be very powerful…

Conclusions

A little bit of noise can be very powerful…

Further reading

AM, Some applications of hypercontractive inequalities in quantum information theory
JMP, vol. 53, 122206, 2012
arXiv:1208.0161

and references therein.

Conclusions

A little bit of noise can be very powerful…

Further reading

AM, Some applications of hypercontractive inequalities in quantum information theory
JMP, vol. 53, 122206, 2012
arXiv:1208.0161

and references therein.

Thanks!

A special case of a conjecture

The following beautiful conjecture (a generalisation of KKL) would imply efficient simulations of quantum query algorithms by classical algorithms on most inputs:

Conjecture [Aaronson and Ambainis '11]

For all degree d polynomials $f : \{\pm 1\}^n \to [-1, 1]$, there exists j such that $I_j(f) \geqslant \mathrm{poly}(\mathsf{Var}(f)/d)$.

- The above result proves the special case of this conjecture where f is a multilinear form whose coefficients are all equal (in absolute value).
- Few other special cases known. One example: symmetric functions f [Bačkurs '12].

The Khot-Vishnoi game

- Parametrised by $N = 2^n$ and $\eta \in [0, 1/2]$.
- Let H be subgroup of $\mathbb{Z}_2^N$ containing Hadamard codewords (strings x such that $x_z = z \oplus s$ for some $s \in \{0,1\}^n$).
- Alice gets uniformly random coset of H defined by a bit-string x.
- Bob gets coset defined by $y = x \oplus e$, where $e_i = 1$ with independent probability η.
- Alice outputs $a \in H \oplus x$, Bob outputs $b \in H \oplus y$ such that $a \oplus b = e$.
- The number of possible inputs to each player is N/n and the number of possible outputs for each player is n.

Communication complexity separation

An $O(\log n)$-qubit quantum protocol is easy; the difficult part is proving the classical lower bound.
The key technical component:

Lemma (informal) [Klartag+Regev '11]

Let $A \subseteq S^{n-1}$ have measure $\sigma(A) \geqslant e^{-n^{1/3}}$. Pick an $(n-1)$-dimensional subspace H uniformly at random. Then $\sigma_H(A \cap H) \approx \sigma(A)$ with high probability.

- Via an inductive argument, this is used to show that for any subsets A, B such that $\sigma(A), \sigma(B) \geqslant e^{-Cn^{1/3}}$,

$$\sigma((A \times B) \cap \mathcal{I}) \geqslant C'\sigma(A)\sigma(B).$$

- $A \times B$ is a rectangle representing the inputs identified by Alice and Bob's communication so far; $\mathcal{I}$ is the set of inputs for which they should output "yes".

A key technical lemma

Lemma [Klartag+Regev '11]

Let f, g satisfy $\int f(x)dx = \int g(x)dx = 1$. Then

$$\int f(x)g(y)d_\perp(x,y) = 1 + O\left(\frac{\log\|f\|_\infty \log\|g\|_\infty}{n}\right)$$

where the integral is taken over orthonormal vectors x, y.

Expand f and g in terms of spherical harmonics Y_k, Y'_k, then

$$\int f(x)g(y)d_\perp(x,y) = \sum_{k\geqslant 0} \mu_k \int Y_k(x)Y'_k(x)dx$$

for some $\{\mu_k\}$ such that $\mu_0 = 1$, $|\mu_k| \leqslant \left(C\frac{k}{n}\right)^{k/2}$. So

$$\left|\int f(x)g(y)d_\perp(x,y) - 1\right| \leqslant \sum_{k\geqslant 0} |\mu_k|\|Y_k\|_2\|Y'_k\|_2.$$