Some applications of hypercontractive inequalities in quantum information theory

Ashley Montanaro

Computer Science Department, University of Bristol, UK

18 June 2013

arXiv:1208.0161

Introduction

In this talk, I will discuss how so-called hypercontractive inequalities can be used to give new(ish) proofs of results in quantum information theory:

- …a bound on the bias of multiplayer XOR games (originally due to [Defant, Popa and Schwarting '10] [Pellegrino and Seoane-Sepúlveda '12]) which implies the first progress on a conjecture about quantum query algorithms;

- …a bound on the bias of local 4-design measurements (originally due to [Lancien and Winter '12]).

Hypercontractive inequalities: a CS perspective

Hypercontractive inequalities have been much used in the quantum field theory literature:

- introduced (in the form of log-Sobolev inequalities) by [Gross '75];
- for detailed reviews see e.g. [Davies, Gross and Simon '92], [Gross '06].

Hypercontractive inequalities: a CS perspective

Hypercontractive inequalities have been much used in the quantum field theory literature:

- introduced (in the form of log-Sobolev inequalities) by [Gross '75];
- for detailed reviews see e.g. [Davies, Gross and Simon '92], [Gross '06].

In the computer science literature, first used by [Kahn, Kalai and Linial '88] in an important paper proving that every boolean function has an influential variable.

The hypercontractive inequality they used is a particularly simple and clean special case due to [Bonami '70], [Gross '75], and often known as the Bonami-Beckner inequality.

Noise

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

Noise

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

- For $\epsilon \in [0, 1]$, define the noise operator T_ϵ as follows:

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

- Here the expectation is over strings $y \in \{\pm 1\}^n$ obtained from x by negating each element of x with independent probability $(1 - \epsilon)/2$.

Noise

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

- For $\epsilon \in [0, 1]$, define the noise operator T_ϵ as follows:

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

- Here the expectation is over strings $y \in \{\pm 1\}^n$ obtained from x by negating each element of x with independent probability $(1 - \epsilon)/2$. So...
  - If $\epsilon = 1$, $T_\epsilon f = f$;
  - If $\epsilon = 0$, $T_\epsilon f$ is constant.

Noise

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

- For $\epsilon \in [0, 1]$, define the noise operator T_ϵ as follows:

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

- Here the expectation is over strings $y \in \{\pm 1\}^n$ obtained from x by negating each element of x with independent probability $(1 - \epsilon)/2$. So…
  - If $\epsilon = 1$, $T_\epsilon f = f$;
  - If $\epsilon = 0$, $T_\epsilon f$ is constant.
- Fairly easy to show that T_ϵ is a contraction, i.e.

$$\|T_\epsilon f\|_p \leqslant \|f\|_p$$

where $\|f\|_p := \left(\frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} |f(x)|^p \right)^{1/p}$.

Hypercontractivity of T_ϵ

The Bonami-Beckner inequality [Bonami '70] [Gross '75]

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

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

Intuition: usually $\|f\|_p \leqslant \|f\|_q$ for $p \leqslant q$, but applying noise to f smoothes out its peaks and makes the norms comparable.

Hypercontractivity of T_ϵ

The Bonami-Beckner inequality [Bonami '70] [Gross '75]

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

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

Intuition: usually $\|f\|_p \leqslant \|f\|_q$ for $p \leqslant q$, but applying noise to f smoothes out its peaks and makes the norms comparable.

Why should we care about this?

Hypercontractivity of T_ϵ

The Bonami-Beckner inequality [Bonami '70] [Gross '75]

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

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

Intuition: usually $\|f\|_p \leqslant \|f\|_q$ for $p \leqslant q$, but applying noise to f smoothes out its peaks and makes the norms comparable.

Why should we care about this?

Corollary

Let $f : \{\pm 1\}^n \to \mathbb{R}$ be a polynomial of degree d. Then:

- for any $p \leqslant 2$, $\|f\|_p \geqslant (p-1)^{d/2}\|f\|_2$;
- for any $q \geqslant 2$, $\|f\|_q \leqslant (q-1)^{d/2}\|f\|_2$.

Intuition: low-degree polynomials are smooth.

Proof of the corollary

Given a degree d (multilinear) polynomial

$$f(x_1, \dots, x_n) = \sum_{S \subseteq [n], |S| \leqslant d} \hat{f}(S) x_S,$$

where $x_S = \prod_{i \in S} x_i$, write $f^{=k} = \sum_{S, |S|=k} \hat{f}(S) x_S$.

Proof of the corollary

Given a degree d (multilinear) polynomial

$$f(x_1, \ldots, x_n) = \sum_{S \subseteq [n], |S| \leqslant d} \hat{f}(S) x_S,$$

where $x_S = \prod_{i \in S} x_i$, write $f^{=k} = \sum_{S, |S|=k} \hat{f}(S) x_S$. Then

$$\|f\|_q^2 = \left\| \sum_{k=0}^{d} f^{=k} \right\|_q^2$$

Proof of the corollary

Given a degree d (multilinear) polynomial

$$f(x_1, \ldots, x_n) = \sum_{S \subseteq [n], |S| \leqslant d} \hat{f}(S) x_S,$$

where $x_S = \prod_{i \in S} x_i$, write $f^{=k} = \sum_{S,|S|=k} \hat{f}(S) x_S$. Then

$$\|f\|_q^2 = \left\| \sum_{k=0}^{d} f^{=k} \right\|_q^2 = \left\| T_{1/\sqrt{q-1}} \left(\sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right) \right\|_q^2$$

Proof of the corollary

Given a degree d (multilinear) polynomial

$$f(x_1, \ldots, x_n) = \sum_{S \subseteq [n], |S| \leqslant d} \hat{f}(S) x_S,$$

where $x_S = \prod_{i \in S} x_i$, write $f^{=k} = \sum_{S, |S|=k} \hat{f}(S) x_S$. Then

$$\begin{aligned}
\|f\|_q^2 &= \left\| \sum_{k=0}^{d} f^{=k} \right\|_q^2 = \left\| T_{1/\sqrt{q-1}} \left(\sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right) \right\|_q^2 \\
&\leqslant \left\| \sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right\|_2^2
\end{aligned}$$

Proof of the corollary

Given a degree d (multilinear) polynomial

$$f(x_1, \ldots, x_n) = \sum_{S \subseteq [n], |S| \leqslant d} \hat{f}(S) x_S,$$

where $x_S = \prod_{i \in S} x_i$, write $f^{=k} = \sum_{S, |S|=k} \hat{f}(S) x_S$. Then

$$\begin{aligned} \|f\|_q^2 &= \left\| \sum_{k=0}^{d} f^{=k} \right\|_q^2 = \left\| T_{1/\sqrt{q-1}} \left(\sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right) \right\|_q^2 \\ &\leqslant \left\| \sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right\|_2^2 = \sum_{k=0}^{d} (q-1)^k \sum_{S \subseteq [n], |S|=k} \hat{f}(S)^2 \end{aligned}$$

Proof of the corollary

Given a degree d (multilinear) polynomial

$$f(x_1, \ldots, x_n) = \sum_{S \subseteq [n], |S| \leqslant d} \hat{f}(S) x_S,$$

where $x_S = \prod_{i \in S} x_i$, write $f^{=k} = \sum_{S, |S|=k} \hat{f}(S) x_S$. Then

$$\begin{aligned}
\|f\|_q^2 &= \left\| \sum_{k=0}^{d} f^{=k} \right\|_q^2 = \left\| T_{1/\sqrt{q-1}} \left(\sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right) \right\|_q^2 \\
&\leqslant \left\| \sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right\|_2^2 = \sum_{k=0}^{d} (q-1)^k \sum_{S \subseteq [n], |S|=k} \hat{f}(S)^2 \\
&\leqslant (q-1)^d \sum_{S \subseteq [n]} \hat{f}(S)^2
\end{aligned}$$

Proof of the corollary

Given a degree d (multilinear) polynomial

$$f(x_1, \dots, x_n) = \sum_{S \subseteq [n], |S| \leqslant d} \hat{f}(S) x_S,$$

where $x_S = \prod_{i \in S} x_i$, write $f^{=k} = \sum_{S,|S|=k} \hat{f}(S) x_S$. Then

$$\begin{aligned}
\|f\|_q^2 &= \left\| \sum_{k=0}^{d} f^{=k} \right\|_q^2 = \left\| T_{1/\sqrt{q-1}} \left(\sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right) \right\|_q^2 \\
&\leqslant \left\| \sum_{k=0}^{d} (q-1)^{k/2} f^{=k} \right\|_2^2 = \sum_{k=0}^{d} (q-1)^k \sum_{S \subseteq [n], |S|=k} \hat{f}(S)^2 \\
&\leqslant (q-1)^d \sum_{S \subseteq [n]} \hat{f}(S)^2 = (q-1)^d \|f\|_2^2.
\end{aligned}$$

(last step: Parseval's equality)

Applications in quantum computation

The above inequality has recently found some applications in quantum computation:

- Separations between quantum and classical communication complexity [Gavinsky et al '07]
- Limitations on quantum random access codes [Ben-Aroya, Regev and de Wolf '08]
- Bounds on non-local games [Buhrman '11]
- Lower bounds on quantum query complexity [Ambainis and de Wolf '12]
- Many more in classical computer science…

Applications in quantum computation

The above inequality has recently found some applications in quantum computation:

- Separations between quantum and classical communication complexity [Gavinsky et al '07]
- Limitations on quantum random access codes [Ben-Aroya, Regev and de Wolf '08]
- Bounds on non-local games [Buhrman '11]
- Lower bounds on quantum query complexity [Ambainis and de Wolf '12]
- Many more in classical computer science…

Today: two more applications.

Application 1: multiplayer XOR games

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

Application 1: multiplayer XOR games

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

A k-player XOR game is defined as follows:

- Fix a multidimensional array $A \in (\{\pm 1\}^n)^k$.

Application 1: multiplayer XOR games

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

A k-player XOR game is defined as follows:

- Fix a multidimensional array $A \in (\{\pm 1\}^n)^k$.
- The j'th player gets an input $i_j \in \{1, \dots, n\}$, picked according to a known distribution π.

Application 1: multiplayer XOR games

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

A k-player XOR game is defined as follows:

- Fix a multidimensional array $A \in (\{\pm 1\}^n)^k$.
- The j'th player gets an input $i_j \in \{1, \ldots, n\}$, picked according to a known distribution π.
- The j'th player must reply with an output $x^j_{i_j} \in \{\pm 1\}$.

Application 1: multiplayer XOR games

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

A k-player XOR game is defined as follows:

- Fix a multidimensional array $A \in (\{\pm 1\}^n)^k$.
- The j'th player gets an input $i_j \in \{1, \ldots, n\}$, picked according to a known distribution π.
- The j'th player must reply with an output $x^j_{i_j} \in \{\pm 1\}$.
- The players win if the product of their outputs is equal to $A_{i_1,\ldots,i_k}$.

Application 1: multiplayer XOR games

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

A k-player XOR game is defined as follows:

- Fix a multidimensional array $A \in (\{\pm 1\}^n)^k$.
- The j'th player gets an input $i_j \in \{1, \dots, n\}$, picked according to a known distribution π.
- The j'th player must reply with an output $x^j_{i_j} \in \{\pm 1\}$.
- The players win if the product of their outputs is equal to $A_{i_1,\dots,i_k}$.

The players are allowed to communicate before the game starts, to agree a strategy, but cannot communicate during the game.

Multiplayer XOR games

For example, consider the CHSH game:

- Two players, two possible inputs, chosen uniformly ($k = 2$, $n = 2$, π is uniform).
- $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$: the players win if their outputs are the same, unless $i_1 = i_2 = 2$, when they win if their outputs are different.

Multiplayer XOR games

For example, consider the CHSH game:

- Two players, two possible inputs, chosen uniformly ($k = 2$, $n = 2$, π is uniform).
- $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$: the players win if their outputs are the same, unless $i_1 = i_2 = 2$, when they win if their outputs are different.

In general, the maximal bias (i.e. difference between probability of success and failure) achievable by deterministic strategies is

$$\beta(G) := \max_{x^1,\dots,x^k \in \{\pm 1\}^n} \left| \sum_{i_1,\dots,i_k=1}^{n} \pi_{i_1,\dots,i_k} A_{i_1,\dots,i_k} x^1_{i_1} \dots x^k_{i_k} \right|.$$

It's easy to see that shared randomness doesn't help.

Why care about XOR games?

- In some cases (e.g. the CHSH game), if the players are allowed to share entanglement they can beat any possible classical strategy.

Why care about XOR games?

- In some cases (e.g. the CHSH game), if the players are allowed to share entanglement they can beat any possible classical strategy.
- XOR games thus provide a clean, mathematically tractable way of studying the power of entanglement.

Why care about XOR games?

- In some cases (e.g. the CHSH game), if the players are allowed to share entanglement they can beat any possible classical strategy.
- XOR games thus provide a clean, mathematically tractable way of studying the power of entanglement.
- XOR games are also interesting in themselves classically:
  - Applications in communication complexity, e.g. [Ford and Gál '05]
  - Known to be NP-hard to compute bias
  - Connections to combinatorics and coding theory.

Why care about XOR games?

- In some cases (e.g. the CHSH game), if the players are allowed to share entanglement they can beat any possible classical strategy.
- XOR games thus provide a clean, mathematically tractable way of studying the power of entanglement.
- XOR games are also interesting in themselves classically:
  - Applications in communication complexity, e.g. [Ford and Gál '05]
  - Known to be NP-hard to compute bias
  - Connections to combinatorics and coding theory.

Today's question

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

i.e. what is the game which minimises the maximal bias achievable?

Previously known results

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

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

Previously known results

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

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

A recent and significant 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

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

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

A recent and significant 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, $\beta(G) = \Omega(k^{-c} n^{-(k-1)/2})$.

We will show how this result can be proven using hypercontractivity (as a small step in the proof).

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 $\hat{f} \in \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$.

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 $\hat{f} \in \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$.
Define as before

$$\|f\|_p := \left(\frac{1}{2^{nk}} \sum_{x^1,\ldots,x^k \in \{\pm 1\}^n} |f(x^1, \ldots, x^k)|^p \right)^{1/p}.$$

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 $\hat{f} \in \mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n$.
Define as before

$$\|f\|_p := \left(\frac{1}{2^{nk}} \sum_{x^1,\dots,x^k \in \{\pm 1\}^n} |f(x^1, \dots, x^k)|^p \right)^{1/p}.$$

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

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

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

What we want to prove

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 $O(k^{\log_2 e}) \approx O(k^{1.45})$.

What we want to prove

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 appropriately.

What we want to prove

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 appropriately.

We'll prove the claim by induction on k, for k a power of 2.

What we want to prove

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 appropriately.

We'll prove the claim by induction on k, for k a power of 2.

- As $\|\hat{f}\|_p$ is nonincreasing with p, it suffices to prove the claim for $p = 2k/(k+1)$.

What we want to prove

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 appropriately.

We'll prove the claim by induction on k, for k a power of 2.

- As $\|\hat{f}\|_p$ is nonincreasing with p, it suffices to prove the claim for $p = 2k/(k+1)$.
- The base case $k = 1$ is trivial ($C_1 = 1$). So, assuming the theorem holds for $k/2$, we prove it holds for k.

Proof

We start with a matrix inequality [Defant, Popa and Schwarting '10]:

$$\|\hat{f}\|_{2k/(k+1)} \leqslant \left(\sum_{i_1,\dots,i_{k/2}} \|(\hat{f}_{i_1,\dots,i_k})_{i_{k/2+1},\dots,i_k=1}^{n}\|_2^{2k/(k+2)} \right)^{(k+2)/4k}$$

$$\times \left(\sum_{i_{k/2+1},\dots,i_k} \|(\hat{f}_{i_1,\dots,i_k})_{i_1,\dots,i_{k/2}=1}^{n}\|_2^{2k/(k+2)} \right)^{(k+2)/4k}$$

Proof

We start with a matrix inequality [Defant, Popa and Schwarting '10]:

$$\begin{aligned}\|\hat{f}\|_{2k/(k+1)} &\leqslant \left(\sum_{i_1,\dots,i_{k/2}} \|(\hat{f}_{i_1,\dots,i_k})^n_{i_{k/2+1},\dots,i_k=1}\|_2^{2k/(k+2)}\right)^{(k+2)/4k} \\ &\times \left(\sum_{i_{k/2+1},\dots,i_k} \|(\hat{f}_{i_1,\dots,i_k})^n_{i_1,\dots,i_{k/2}=1}\|_2^{2k/(k+2)}\right)^{(k+2)/4k}\end{aligned}$$

We estimate the second term (the first follows exactly the same procedure).

Proof

For each $i_{k/2+1}, \ldots, i_k \in [n]$, define $f_{i_{k/2+1},\ldots,i_k} : (\mathbb{R}^n)^{k/2} \to \mathbb{R}$ by

$$f_{i_{k/2+1},\ldots,i_k}(x^1, \ldots, x^{k/2}) = \sum_{i_1,\ldots,i_{k/2}} \hat{f}_{i_1,\ldots,i_k} x^1_{i_1} x^2_{i_2} \ldots x^{k/2}_{i_{k/2}}.$$

Proof

For each $i_{k/2+1}, \ldots, i_k \in [n]$, define $f_{i_{k/2+1},\ldots,i_k} : (\mathbb{R}^n)^{k/2} \to \mathbb{R}$ by

$$f_{i_{k/2+1},\ldots,i_k}(x^1, \ldots, x^{k/2}) = \sum_{i_1,\ldots,i_{k/2}} \hat{f}_{i_1,\ldots,i_k} x^1_{i_1} x^2_{i_2} \ldots x^{k/2}_{i_{k/2}}.$$

Also define a "dual" function $f'_{x^1,\ldots,x^{k/2}} : (\mathbb{R}^n)^{k/2} \to \mathbb{R}$ by

$$f'_{x^1,\ldots,x^{k/2}}(x^{k/2+1}, \ldots, x^k) = f(x^1, \ldots, x^k).$$

Proof

For each $i_{k/2+1}, \ldots, i_k \in [n]$, define $f_{i_{k/2+1},\ldots,i_k} : (\mathbb{R}^n)^{k/2} \to \mathbb{R}$ by

$$f_{i_{k/2+1},\ldots,i_k}(x^1, \ldots, x^{k/2}) = \sum_{i_1,\ldots,i_{k/2}} \hat{f}_{i_1,\ldots,i_k} x^1_{i_1} x^2_{i_2} \ldots x^{k/2}_{i_{k/2}}.$$

Also define a “dual” function $f'_{x^1,\ldots,x^{k/2}} : (\mathbb{R}^n)^{k/2} \to \mathbb{R}$ by

$$f'_{x^1,\ldots,x^{k/2}}(x^{k/2+1}, \ldots, x^k) = f(x^1, \ldots, x^k).$$

We have

$$f'_{x^1,\ldots,x^{k/2}}(x^{k/2+1}, \ldots, x^k) = \sum_{i_{k/2+1},\ldots,i_k=1}^{n} f_{i_{k/2+1},\ldots,i_k}(x^1, \ldots, x^{k/2}) x^{k/2+1}_{i_{k/2+1}} \ldots x^k_{i_k};$$

of course $\|f'_{x^1,\ldots,x^{k/2}}\|_\infty \leqslant \|f\|_\infty$.

Proof

For each tuple $i_{k/2+1}, \ldots, i_k$ we have by Parseval's equality

$$\|(\hat{f}_{i_1,\ldots,i_k})_{i_1,\ldots,i_{k/2}=1}^{n}\|_2 = \left(\sum_{i_1,\ldots,i_{k/2}=1}^{n} \hat{f}_{i_1,\ldots,i_k}^2\right)^{1/2} = \|f_{i_{k/2+1},\ldots,i_k}\|_2.$$

Proof

For each tuple $i_{k/2+1}, \ldots, i_k$ we have by Parseval's equality

$$\|(\hat{f}_{i_1,\ldots,i_k})_{i_1,\ldots,i_{k/2}=1}^n\|_2 = \left(\sum_{i_1,\ldots,i_{k/2}=1}^n \hat{f}_{i_1,\ldots,i_k}^2\right)^{1/2} = \|f_{i_{k/2+1},\ldots,i_k}\|_2.$$

By hypercontractivity,

$$\|f_{i_{k/2+1},\ldots,i_k}\|_2^{2k/(k+2)} \leqslant \left(\frac{k+2}{k-2}\right)^{\frac{k^2}{2(k+2)}} \|f_{i_{k/2+1},\ldots,i_k}\|_{2k/(k+2)}^{2k/(k+2)}.$$

Proof

For each tuple $i_{k/2+1}, \ldots, i_k$ we have by Parseval's equality

$$\|(\hat{f}_{i_1,\ldots,i_k})_{i_1,\ldots,i_{k/2}=1}^{n}\|_2 = \left(\sum_{i_1,\ldots,i_{k/2}=1}^{n} \hat{f}_{i_1,\ldots,i_k}^2\right)^{1/2} = \|f_{i_{k/2+1},\ldots,i_k}\|_2.$$

By hypercontractivity,

$$\|f_{i_{k/2+1},\ldots,i_k}\|_2^{2k/(k+2)} \leqslant \left(\frac{k+2}{k-2}\right)^{\frac{k^2}{2(k+2)}} \|f_{i_{k/2+1},\ldots,i_k}\|_{2k/(k+2)}^{2k/(k+2)}.$$

We now observe that, for any $p \geqslant 1$,

$$\begin{aligned}\sum_{i_{k/2+1},\ldots,i_k} \|f_{i_{k/2+1},\ldots,i_k}\|_p^p &= \mathbb{E}_{x^1,\ldots,x^{k/2}}\left[\sum_{i_{k/2+1},\ldots,i_k} |f_{i_{k/2+1},\ldots,i_k}(x^1,\ldots,x^{k/2})|^p\right] \\ &= \mathbb{E}_{x^1,\ldots,x^{k/2}}\left[\|\hat{f}'_{x^1,\ldots,x^{k/2}}\|_p^p\right].\end{aligned}$$

Proof

Hence, taking $p = 2k/(k+2) = 2(k/2)/(k/2+1)$, we have

$$\sum_{i_{k/2+1},\ldots,i_k} \|(\hat{f}_{i_1,\ldots,i_k})_{i_1,\ldots,i_{k/2}=1}^{n}\|_2^{2k/(k+2)}$$

$$\leqslant \left(\frac{k+2}{k-2}\right)^{\frac{k^2}{2(k+2)}} \mathbb{E}_{x^1,\ldots,x^{k/2}}\left[\|\hat{f}'_{x^1,\ldots,x^{k/2}}\|_{2k/(k+2)}^{2k/(k+2)}\right]$$

Proof

Hence, taking $p = 2k/(k+2) = 2(k/2)/(k/2+1)$, we have

$$\sum_{i_{k/2+1},\ldots,i_k} \|(\hat{f}_{i_1,\ldots,i_k})_{i_1,\ldots,i_{k/2}=1}^n\|_2^{2k/(k+2)}$$

$$\leqslant \left(\frac{k+2}{k-2}\right)^{\frac{k^2}{2(k+2)}} \mathbb{E}_{x^1,\ldots,x^{k/2}} \left[\|\hat{f}'_{x^1,\ldots,x^{k/2}}\|_{2k/(k+2)}^{2k/(k+2)}\right]$$

$$\leqslant \left(\frac{k+2}{k-2}\right)^{\frac{k^2}{2(k+2)}} C_{k/2}^{2k/(k+2)} \|f\|_\infty^{2k/(k+2)}$$

by the inductive hypothesis.

Proof

Combining both terms in the first inequality,

$$\left(\sum_{i_1,\dots,i_k} |\hat{f}_{i_1,\dots,i_k}|^{2k/(k+1)}\right)^{(k+1)/(2k)} \leqslant \left(\frac{k+2}{k-2}\right)^{k/4} C_{k/2}\|f\|_\infty.$$

Proof

Combining both terms in the first inequality,

$$\left(\sum_{i_1,\ldots,i_k} |\hat{f}_{i_1,\ldots,i_k}|^{2k/(k+1)}\right)^{(k+1)/(2k)} \leqslant \left(\frac{k+2}{k-2}\right)^{k/4} C_{k/2}\|f\|_\infty.$$

Thus

$$\boxed{C_k \leqslant \left(1+\frac{4}{k-2}\right)^{k/4} C_{k/2}.}$$

Proof

Combining both terms in the first inequality,

$$\left(\sum_{i_1,\ldots,i_k} |\hat{f}_{i_1,\ldots,i_k}|^{2k/(k+1)}\right)^{(k+1)/(2k)} \leqslant \left(\frac{k+2}{k-2}\right)^{k/4} C_{k/2}\|f\|_\infty.$$

Thus

$$\boxed{C_k \leqslant \left(1+\frac{4}{k-2}\right)^{k/4} C_{k/2}.}$$

Observing that $(1+4/(k-2))^{k/4} \leqslant (1+O(1/k))e$, we have $C_k = O(k^{\log_2 e})$ as claimed.

A conjecture of Aaronson and Ambainis

The following beautiful conjecture is currently open:

Conjecture [Aaronson and Ambainis '11]

Every bounded low-degree polynomial on the boolean cube has an influential variable.

A conjecture of Aaronson and Ambainis

The following beautiful conjecture is currently open:

Conjecture [Aaronson and Ambainis '11]

Every bounded low-degree polynomial on the boolean cube has an influential variable.

- Generalises a prior result showing this for decision trees [O'Donnell et al '05].
- One reason this conjecture is interesting: it would imply that every quantum query algorithm can be approximated by a classical algorithm on "most" inputs.
- One special case known: when f is symmetric, i.e. $f(x)$ depends only on $\sum_i x_i$ [Backurs '12].

A conjecture of Aaronson and Ambainis

A more formal version of the conjecture:

Conjecture **[Aaronson and Ambainis '11]**

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

A conjecture of Aaronson and Ambainis

A more formal version of the conjecture:

Conjecture [Aaronson and Ambainis '11]

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

What does this mean?

- Write $\mathbb{E}[f] = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} f(x)$. Then the ($\ell_2$) variance of f is

$$\mathsf{Var}(f) = \mathbb{E}[(f - \mathbb{E}[f])^2]$$

- Define the influence of the j'th variable on f as

$$I_j(f) = \frac{1}{2^{n+2}} \sum_{x \in \{\pm 1\}^n} (f(x) - f(x^j))^2,$$

where x^j is x with the j'th variable negated.

A conjecture of Aaronson and Ambainis

Using the above strengthening of the BH inequality, it is easy to prove a very special case of the Aaronson-Ambainis conjecture. Let

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

where $\hat{f}_{i_1, \ldots, i_k} = \pm\alpha$ for some α.

A conjecture of Aaronson and Ambainis

Using the above strengthening of the BH inequality, it is easy to prove a very special case of the Aaronson-Ambainis conjecture. Let

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

where $\hat{f}_{i_1, \ldots, i_k} = \pm\alpha$ for some α.

- f depends on nk variables x^j_ℓ, $1 \leqslant j \leqslant k$ and $1 \leqslant \ell \leqslant n$.
- The influence of variable (j, ℓ) on f is

$$\mathrm{Inf}_{(j,\ell)}(f) = \sum_{i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_k} \hat{f}^2_{i_1, \ldots, i_{j-1}, \ell, i_{j+1}, \ldots, i_k} = n^{k-1}\alpha^2.$$

A conjecture of Aaronson and Ambainis

Using the above strengthening of the BH inequality, it is easy to prove a very special case of the Aaronson-Ambainis conjecture. Let

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

where $\hat{f}_{i_1,\ldots,i_k} = \pm\alpha$ for some α.

- f depends on nk variables x^j_ℓ, $1 \leqslant j \leqslant k$ and $1 \leqslant \ell \leqslant n$.
- The influence of variable (j, ℓ) on f is

$$\mathrm{Inf}_{(j,\ell)}(f) = \sum_{i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_k} \hat{f}^2_{i_1,\ldots,i_{j-1},\ell,i_{j+1},\ldots,i_k} = n^{k-1}\alpha^2.$$

Corollary

If f is a multilinear form such that $\|f\|_\infty \leqslant 1$ and $\hat{f}_{i_1,\ldots,i_k} = \pm\alpha$ for some α, then $I_{(j,\ell)}(f) = \Omega(\mathsf{Var}(f)^2/k^3)$ for all (j, ℓ).

Application 2: The bias of local 4-designs

Given a quantum state which is promised to be either ρ (with probability p) or σ (with probability $1-p$), we want to determine which is the case via a measurement.

- The most general kind of quantum measurement is known as a POVM, i.e. a partition of the identity into positive operators.

- The optimal measurement achieves success probability

$$\frac{1}{2}\left(1 + \|p\rho - (1-p)\sigma\|_1\right),$$

  where $\|M\|_1 = \operatorname{tr}|M|$ is the usual trace norm.

- Setting $\Delta = p\rho - (1-p)\sigma$, the optimal bias is just $\|\Delta\|_1$.

The bias of local measurements

What if we are not allowed to perform an arbitrary measurement, but can only perform a single fixed quantum measurement, followed by arbitrary classical postprocessing?

The bias of local measurements

What if we are not allowed to perform an arbitrary measurement, but can only perform a single fixed quantum measurement, followed by arbitrary classical postprocessing?

- Given a POVM $M = (M_i)$, let ρ^M, σ^M be the probability distributions on measurement outcomes induced by performing M on ρ, σ.
- The optimal bias one can achieve by performing M is then equal to

$$\|\Delta\|_M \quad := \quad \|p\rho^M - (1-p)\sigma^M\|_1$$

The bias of local measurements

What if we are not allowed to perform an arbitrary measurement, but can only perform a single fixed quantum measurement, followed by arbitrary classical postprocessing?

- Given a POVM $M = (M_i)$, let ρ^M, σ^M be the probability distributions on measurement outcomes induced by performing M on ρ, σ.

- The optimal bias one can achieve by performing M is then equal to

$$\begin{aligned} \|\Delta\|_M &:= \|p\rho^M - (1-p)\sigma^M\|_1 \\ &= \sum_i |p \operatorname{tr} M_i \rho - (1-p) \operatorname{tr} M_i \sigma| \end{aligned}$$

The bias of local measurements

What if we are not allowed to perform an arbitrary measurement, but can only perform a single fixed quantum measurement, followed by arbitrary classical postprocessing?

- Given a POVM $M = (M_i)$, let ρ^M, σ^M be the probability distributions on measurement outcomes induced by performing M on ρ, σ.

- The optimal bias one can achieve by performing M is then equal to

$$\begin{aligned} \|\Delta\|_M &:= \|p\rho^M - (1-p)\sigma^M\|_1 \\ &= \sum_i |p \operatorname{tr} M_i\rho - (1-p) \operatorname{tr} M_i\sigma| \\ &= \sum_i |\operatorname{tr} M_i\Delta|. \end{aligned}$$

The uniform POVM

One particularly natural (and optimal!) fixed measurement is the uniform POVM in n dimensions.

The uniform POVM

One particularly natural (and optimal!) fixed measurement is the uniform POVM in n dimensions.

- This is a continuous POVM with a measurement operator $M_\psi \propto |\psi\rangle\langle\psi|$ for every n-dimensional pure state $|\psi\rangle$, normalised such that $\int d\psi M_\psi = I$.

The uniform POVM

One particularly natural (and optimal!) fixed measurement is the uniform POVM in n dimensions.

- This is a continuous POVM with a measurement operator $M_\psi \propto |\psi\rangle\langle\psi|$ for every n-dimensional pure state $|\psi\rangle$, normalised such that $\int d\psi M_\psi = I$.
- We can't actually perform this physically, but can approximate it using t-designs.

The uniform POVM

One particularly natural (and optimal!) fixed measurement is the uniform POVM in n dimensions.

- This is a continuous POVM with a measurement operator $M_\psi \propto |\psi\rangle\langle\psi|$ for every n-dimensional pure state $|\psi\rangle$, normalised such that $\int d\psi M_\psi = I$.
- We can't actually perform this physically, but can approximate it using t-designs.
- A rank-one POVM $M = (M_i)$ in n dimensions is called a t-design if

$$\sum_i p_i P_i^{\otimes t} = \int d\psi |\psi\rangle\langle\psi|^{\otimes t},$$

  where $p_i = \frac{1}{n} \operatorname{tr} M_i$ and $P_i = \frac{1}{\operatorname{tr} M_i} M_i$.

The uniform POVM

One particularly natural (and optimal!) fixed measurement is the uniform POVM in n dimensions.

- This is a continuous POVM with a measurement operator $M_\psi \propto |\psi\rangle\langle\psi|$ for every n-dimensional pure state $|\psi\rangle$, normalised such that $\int d\psi M_\psi = I$.
- We can't actually perform this physically, but can approximate it using t-designs.
- A rank-one POVM $M = (M_i)$ in n dimensions is called a t-design if
$$\sum_i p_i P_i^{\otimes t} = \int d\psi |\psi\rangle\langle\psi|^{\otimes t},$$
where $p_i = \frac{1}{n} \operatorname{tr} M_i$ and $P_i = \frac{1}{\operatorname{tr} M_i} M_i$.
- As t increases, t-designs become better and better approximations to the uniform POVM.

The bias of 4-design measurements

Theorem **[Ambainis and Emerson '07], [Matthews, Wehner and Winter '09]**

Let M be a 4-design and set $\Delta = (\rho - \sigma)/2$. Then

$$\|\Delta\|_M \geqslant C\sqrt{\operatorname{tr}\Delta^2},$$

for some universal constant $C > 0$.

The bias of 4-design measurements

Theorem [Ambainis and Emerson '07], [Matthews, Wehner and Winter '09]

Let M be a 4-design and set $\Delta = (\rho - \sigma)/2$. Then

$$\|\Delta\|_M \geqslant C\sqrt{\operatorname{tr}\Delta^2},$$

for some universal constant $C > 0$.

One can generalise this to a setting where locality comes into play by making M into a tensor product of 4-designs. That is:

- Each operator is of the form $M_{i_1,\ldots,i_k} = M_{i_1} \otimes M_{i_2} \otimes \ldots M_{i_k}$.
- Each individual measurement (M_j) is a 4-design.

This is interesting because it allows us to explore the power of local vs. global measurements.

Local 4-designs

Theorem [Lancien and Winter '12]

Let M be a k-partite measurement which is a product of local 4-designs and set $\Delta = p\rho - (1-p)\sigma$. Then

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

for some universal constant $D > 0$.

- Previously known for $k = 2$ [Matthews, Wehner and Winter '09].

Local 4-designs

Theorem [Lancien and Winter '12]

Let M be a k-partite measurement which is a product of local 4-designs and set $\Delta = p\rho - (1-p)\sigma$. Then

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

for some universal constant $D > 0$.

- Previously known for $k = 2$ [Matthews, Wehner and Winter '09].

We give a new proof using hypercontractivity.

The $k = 1$ case

We use the "fourth moment method" [Littlewood '30] [Berger '97]:

$$\|\Delta\|_M \quad = \quad \sum_i |\operatorname{tr} M_i \Delta|$$

The $k = 1$ case

We use the "fourth moment method" [Littlewood '30] [Berger '97]:

$$\|\Delta\|_M = \sum_i |\operatorname{tr} M_i \Delta| = n \sum_i p_i |\operatorname{tr} P_i \Delta|$$

The $k = 1$ case

We use the "fourth moment method" [Littlewood '30] [Berger '97]:

$$\|\Delta\|_M = \sum_i |\operatorname{tr} M_i \Delta| = n \sum_i p_i |\operatorname{tr} P_i \Delta| \geqslant n \frac{\left(\sum_i p_i (\operatorname{tr} P_i \Delta)^2\right)^{3/2}}{\left(\sum_i p_i (\operatorname{tr} P_i \Delta)^4\right)^{1/2}}$$

The $k = 1$ case

We use the "fourth moment method" [Littlewood '30] [Berger '97]:

$$\begin{aligned}
\|\Delta\|_M &= \sum_i |\operatorname{tr} M_i \Delta| = n \sum_i p_i |\operatorname{tr} P_i \Delta| \geqslant n \frac{\left(\sum_i p_i (\operatorname{tr} P_i \Delta)^2\right)^{3/2}}{\left(\sum_i p_i (\operatorname{tr} P_i \Delta)^4\right)^{1/2}} \\
&= n \frac{\left(\operatorname{tr}\left(\sum_i p_i P_i^{\otimes 2}\right) \Delta^{\otimes 2}\right)^{3/2}}{\left(\operatorname{tr}\left(\sum_i p_i P_i^{\otimes 4}\right) \Delta^{\otimes 4}\right)^{1/2}}.
\end{aligned}$$

The $k = 1$ case

We use the "fourth moment method" [Littlewood '30] [Berger '97]:

$$\begin{aligned} \|\Delta\|_M &= \sum_i |\operatorname{tr} M_i \Delta| = n \sum_i p_i |\operatorname{tr} P_i \Delta| \geqslant n \frac{\left(\sum_i p_i (\operatorname{tr} P_i \Delta)^2\right)^{3/2}}{\left(\sum_i p_i (\operatorname{tr} P_i \Delta)^4\right)^{1/2}} \\ &= n \frac{\left(\operatorname{tr}\left(\sum_i p_i P_i^{\otimes 2}\right) \Delta^{\otimes 2}\right)^{3/2}}{\left(\operatorname{tr}\left(\sum_i p_i P_i^{\otimes 4}\right) \Delta^{\otimes 4}\right)^{1/2}}. \end{aligned}$$

As M is a 4-design,

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

The $k = 1$ case

We use the "fourth moment method" [Littlewood '30] [Berger '97]:

$$\begin{aligned}\|\Delta\|_M &= \sum_i |\operatorname{tr} M_i\Delta| = n\sum_i p_i|\operatorname{tr} P_i\Delta| \geqslant n\frac{\left(\sum_i p_i(\operatorname{tr} P_i\Delta)^2\right)^{3/2}}{\left(\sum_i p_i(\operatorname{tr} P_i\Delta)^4\right)^{1/2}}\\ &= n\frac{\left(\operatorname{tr}\left(\sum_i p_i P_i^{\otimes 2}\right)\Delta^{\otimes 2}\right)^{3/2}}{\left(\operatorname{tr}\left(\sum_i p_i P_i^{\otimes 4}\right)\Delta^{\otimes 4}\right)^{1/2}}.\end{aligned}$$

As M is a 4-design,

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

So, if we can upper bound $\int(\operatorname{tr}\Delta|\psi\rangle\langle\psi|)^4 d\psi$ in terms of $\int(\operatorname{tr}\Delta|\psi\rangle\langle\psi|)^2 d\psi$, this will give a lower bound on $\|\Delta\|_M$.

Functions on the sphere

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

Functions on the sphere

- Let S^n be the real n-sphere, i.e. $\{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$.
- For $f : S^n \to \mathbb{R}$ define the $L^p(S^n)$ norms as

$$\|f\|_{L^p(S^n)} := \left(\int |f(\xi)|^p d\xi \right)^{1/p},$$

where we integrate with respect to the uniform measure on S^n, normalised so that $\int d\xi = 1$.

Functions on the sphere

- Let S^n be the real n-sphere, i.e. $\{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$.
- For $f : S^n \to \mathbb{R}$ define the $L^p(S^n)$ norms as

$$\|f\|_{L^p(S^n)} := \left(\int |f(\xi)|^p d\xi \right)^{1/p},$$

  where we integrate with respect to the uniform measure on S^n, normalised so that $\int d\xi = 1$.

- Identify each n-dimensional quantum state $|\psi\rangle$ (element of the unit sphere in $\mathbb{C}^n$) with a real vector $\xi \in S^{2n-1}$ by taking real and imaginary parts.

Functions on the sphere

- Let S^n be the real n-sphere, i.e. $\{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$.
- For $f : S^n \to \mathbb{R}$ define the $L^p(S^n)$ norms as

$$\|f\|_{L^p(S^n)} := \left(\int |f(\xi)|^p d\xi \right)^{1/p},$$

  where we integrate with respect to the uniform measure on S^n, normalised so that $\int d\xi = 1$.
- Identify each n-dimensional quantum state $|\psi\rangle$ (element of the unit sphere in $\mathbb{C}^n$) with a real vector $\xi \in S^{2n-1}$ by taking real and imaginary parts.
- Now consider the function $f(\xi) = \operatorname{tr} \Delta |\psi\rangle\langle\psi|$.

Functions on the sphere

- Let S^n be the real n-sphere, i.e. $\{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$.
- For $f : S^n \to \mathbb{R}$ define the $L^p(S^n)$ norms as

$$\|f\|_{L^p(S^n)} := \left(\int |f(\xi)|^p d\xi\right)^{1/p},$$

  where we integrate with respect to the uniform measure on S^n, normalised so that $\int d\xi = 1$.

- Identify each n-dimensional quantum state $|\psi\rangle$ (element of the unit sphere in $\mathbb{C}^n$) with a real vector $\xi \in S^{2n-1}$ by taking real and imaginary parts.
- Now consider the function $f(\xi) = \operatorname{tr} \Delta |\psi\rangle\langle\psi|$.
- We want to upper bound $\|f\|_{L^4(S^n)}$ in terms of $\|f\|_{L^2(S^n)}$.

Hypercontractivity to the rescue?

Claim

f is a degree 2 polynomial in the components of ξ.

Suggests that we could relate $\|f\|_{L^4(S^n)}$ to $\|f\|_{L^2(S^n)}$ using some form of hypercontractivity…

Hypercontractivity to the rescue?

Claim

f is a degree 2 polynomial in the components of ξ.

Suggests that we could relate $\|f\|_{L^4(S^n)}$ to $\|f\|_{L^2(S^n)}$ using some form of hypercontractivity…

We need to understand hypercontractivity for functions on the sphere, and some basic ideas from the theory of spherical harmonics.

Spherical harmonics

- The restriction of every degree d polynomial $f : \mathbb{R}^{n+1} \to \mathbb{R}$ to the sphere S^n can be written as

$$f(x) = \sum_{k=0}^{d} Y_k(x),$$

where $Y_k : S^n \to \mathbb{R}$ is called a spherical harmonic, and is the restriction of a degree k polynomial to the sphere, satisfying $\int Y_j(\xi) Y_k(\xi) d\xi = 0$ for $j \neq k$.

Spherical harmonics

- The restriction of every degree d polynomial $f : \mathbb{R}^{n+1} \to \mathbb{R}$ to the sphere S^n can be written as

$$f(x) = \sum_{k=0}^{d} Y_k(x),$$

where $Y_k : S^n \to \mathbb{R}$ is called a spherical harmonic, and is the restriction of a degree k polynomial to the sphere, satisfying $\int Y_j(\xi) Y_k(\xi) d\xi = 0$ for $j \neq k$.

- The Poisson semigroup (which can be thought of as a "noise operator" for the sphere) is defined by

$$(P_\epsilon f)(x) = \sum_k \epsilon^k Y_k(x).$$

Hypercontractivity on the sphere

Crucially, it is known that the Poisson semigroup is indeed hypercontractive.

Theorem [Beckner '92]

If $1 \leqslant p \leqslant q \leqslant \infty$ and $\epsilon \leqslant \sqrt{\frac{p-1}{q-1}}$, then

$$\|P_\epsilon f\|_{L^q(S^n)} \leqslant \|f\|_{L^p(S^n)}.$$

Hypercontractivity on the sphere

Crucially, it is known that the Poisson semigroup is indeed hypercontractive.

Theorem [Beckner '92]

If $1 \leqslant p \leqslant q \leqslant \infty$ and $\epsilon \leqslant \sqrt{\frac{p-1}{q-1}}$, then

$$\|P_\epsilon f\|_{L^q(S^n)} \leqslant \|f\|_{L^p(S^n)}.$$

Just as in the setting of the cube $\{\pm 1\}^n$, this implies the following corollary.

Corollary

Let $f : \mathbb{R}^{n+1} \to \mathbb{R}$ be a degree d polynomial. Then, for $q \geqslant 2$,

$$\|f\|_{L^q(S^n)} \leqslant (q-1)^{d/2} \|f\|_{L^2(S^n)}.$$

Declare victory

Taking $q = 4$, we see that

$$\left(\int (\operatorname{tr} \Delta|\psi\rangle\langle\psi|)^4 d\psi\right)^{1/4} \leqslant 3 \left(\int (\operatorname{tr} \Delta|\psi\rangle\langle\psi|)^2 d\psi\right)^{1/2},$$

so we get

$$\|\Delta\|_M \geqslant \frac{n}{9} \left(\int (\operatorname{tr} \Delta|\psi\rangle\langle\psi|)^2 d\psi\right)^{1/2};$$

the RHS can be explicitly evaluated to give

$$\boxed{\|\Delta\|_M \geqslant \frac{1}{9(1+1/n)^{1/2}} \sqrt{\operatorname{tr} \Delta^2}.}$$

Declare victory

Taking $q = 4$, we see that

$$\left(\int (\operatorname{tr} \Delta|\psi\rangle\langle\psi|)^4 d\psi\right)^{1/4} \leqslant 3 \left(\int (\operatorname{tr} \Delta|\psi\rangle\langle\psi|)^2 d\psi\right)^{1/2},$$

so we get

$$\|\Delta\|_M \geqslant \frac{n}{9} \left(\int (\operatorname{tr} \Delta|\psi\rangle\langle\psi|)^2 d\psi\right)^{1/2};$$

the RHS can be explicitly evaluated to give

$$\boxed{\|\Delta\|_M \geqslant \frac{1}{9(1+1/n)^{1/2}} \sqrt{\operatorname{tr} \Delta^2}.}$$

So we've solved the case $k = 1 \ldots$ what about higher k?

Arbitrary k

We start the proof in the same way: As M is a tensor product of local 4-designs,

$$\|\Delta\|_M \geqslant n^k \frac{\left(\int \dots \int d\psi_1 \dots d\psi_k (\operatorname{tr} \Delta(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|))^2\right)^{3/2}}{\left(\int \dots \int d\psi_1 \dots d\psi_k (\operatorname{tr} \Delta(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|))^4\right)^{1/2}}$$

Arbitrary k

We start the proof in the same way: As M is a tensor product of local 4-designs,

$$\begin{aligned}
\|\Delta\|_M &\geqslant n^k \frac{\left(\int \ldots \int d\psi_1 \ldots d\psi_k (\operatorname{tr} \Delta(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|))^2\right)^{3/2}}{\left(\int \ldots \int d\psi_1 \ldots d\psi_k (\operatorname{tr} \Delta(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|))^4\right)^{1/2}} \\
&= n^k \frac{\|f\|^3_{L^2((S^{2n-1})^k)}}{\|f\|^2_{L^4((S^{2n-1})^k)}},
\end{aligned}$$

where we define the function $f : (S^{2n-1})^k \to \mathbb{R}$ by

$$f(\xi_1, \ldots, \xi_k) = \operatorname{tr} \Delta(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|),$$

where $|\psi_i\rangle$ is the n-dimensional complex unit vector whose real and imaginary parts are given by $\xi_i \in S^{2n-1}$ in the obvious way.

Arbitrary k

We start the proof in the same way: As M is a tensor product of local 4-designs,

$$\begin{aligned} \|\Delta\|_M &\geqslant n^k \frac{\left(\int \dots \int d\psi_1 \dots d\psi_k (\operatorname{tr} \Delta(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|))^2\right)^{3/2}}{\left(\int \dots \int d\psi_1 \dots d\psi_k (\operatorname{tr} \Delta(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|))^4\right)^{1/2}} \\ &= n^k \frac{\|f\|^3_{L^2((S^{2n-1})^k)}}{\|f\|^2_{L^4((S^{2n-1})^k)}}, \end{aligned}$$

where we define the function $f : (S^{2n-1})^k \to \mathbb{R}$ by

$$f(\xi_1, \dots, \xi_k) = \operatorname{tr} \Delta(|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|),$$

where $|\psi_i\rangle$ is the n-dimensional complex unit vector whose real and imaginary parts are given by $\xi_i \in S^{2n-1}$ in the obvious way.

As before, we want to relate $\|f\|_{L^4((S^{2n-1})^k)}$ to $\|f\|_{L^2((S^{2n-1})^k)}$.

Arbitrary k

Here's where the magic happens: the $L^p \to L^q$ norm is multiplicative, so as a corollary of Beckner's result...

Corollary

Let $f : (S^n)^k \to \mathbb{R}$. If $1 \leqslant p \leqslant q \leqslant \infty$ and $\epsilon \leqslant \sqrt{\frac{p-1}{q-1}}$, then

$$\|P_\epsilon^{\otimes k} f\|_{L^q((S^n)^k)} \leqslant \|f\|_{L^p((S^n)^k)}.$$

Arbitrary k

Here's where the magic happens: the $L^p \to L^q$ norm is multiplicative, so as a corollary of Beckner's result...

Corollary

Let $f : (S^n)^k \to \mathbb{R}$. If $1 \leqslant p \leqslant q \leqslant \infty$ and $\epsilon \leqslant \sqrt{\frac{p-1}{q-1}}$, then

$$\|P_\epsilon^{\otimes k} f\|_{L^q((S^n)^k)} \leqslant \|f\|_{L^p((S^n)^k)}.$$

Also, the same corollary goes through!

Corollary

Let $f : (\mathbb{R}^{n+1})^k \to \mathbb{R}$ be a degree d polynomial in the components of each $x^1, \ldots, x^k \in \mathbb{R}^{n+1}$. Then, for any $q \geqslant 2$,

$$\|f\|_{L^q((S^n)^k)} \leqslant (q-1)^{dk/2} \|f\|_{L^2((S^n)^k)}.$$

Completing the proof

We have

$$\|\Delta\|_M \geqslant n^k \frac{\|f\|^3_{L^2((S^{2n-1})^k)}}{\|f\|^2_{L^4((S^{2n-1})^k)}} \geqslant \left(\frac{n}{9}\right)^k \|f\|_{L^2((S^{2n-1})^k)}.$$

Completing the proof

We have

$$\|\Delta\|_M \geqslant n^k \frac{\|f\|^3_{L^2((S^{2n-1})^k)}}{\|f\|^2_{L^4((S^{2n-1})^k)}} \geqslant \left(\frac{n}{9}\right)^k \|f\|_{L^2((S^{2n-1})^k)}.$$

All that remains is to explicitly calculate

$$\begin{aligned}\|f\|^2_{L^2((S^{2n-1})^k)} &= \operatorname{tr}\left(\int \dots \int d\psi_1 \dots d\psi_k |\psi_1\rangle\langle\psi_1|^{\otimes 2} \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|^{\otimes 2}\right) \Delta^{\otimes 2} \\ &= \operatorname{tr}\left(\frac{I+F}{n(n+1)}\right)^{\otimes k} \Delta^{\otimes 2} \\ &= \frac{1}{n^k(n+1)^k} \sum_{S \subseteq [k]} \operatorname{tr}\left[(\operatorname{tr}_S \Delta)^2\right].\end{aligned}$$

Comparison to previous work

The approach of [Lancien and Winter '12] has definite advantages:

- Better constants
- Based only on clever use of "elementary" techniques (e.g. Cauchy-Schwarz)
- More "concrete".

Comparison to previous work

The approach of [Lancien and Winter '12] has definite advantages:

- Better constants
- Based only on clever use of "elementary" techniques (e.g. Cauchy-Schwarz)
- More "concrete".

But the hypercontractive approach has good points too:

- Extension to arbitrary k is essentially immediate
- Can be extended to t-designs for $t > 4$ with little effort
- Gives an intuitive explanation of the exponential prefactor
- More "natural" (if one already knows hypercontractivity!)

Summary

- Hypercontractive inequalities seem to be a powerful tool for proving results in quantum information theory.
- The proofs given here were of previously known results: in both cases the results appear somewhat less technical, at the expense of being less concrete (and giving worse constants).

Open problems:

- Prove the Aaronson-Ambainis conjecture (using hypercontractivity!).
- Explore connections between the results of [Lancien and Winter '12] and testing product states.

Summary

- Hypercontractive inequalities seem to be a powerful tool for proving results in quantum information theory.
- The proofs given here were of previously known results: in both cases the results appear somewhat less technical, at the expense of being less concrete (and giving worse constants).

Open problems:

- Prove the Aaronson-Ambainis conjecture (using hypercontractivity!).
- Explore connections between the results of [Lancien and Winter '12] and testing product states.

Thanks!