Hypercontractivity, XOR games and the Aaronson-Ambainis conjecture

Ashley Montanaro

Computer Science Department, University of Bristol, UK

20 August 2013

arXiv:1208.0161

University of BRISTOL

UNIVERSITY OF CAMBRIDGE

Introduction

In this talk, I will discuss how so-called hypercontractive inequalities can be used to give a new(ish) proof of a bound on the bias of multiplayer XOR games, which implies a (very) special case of a conjecture about quantum query algorithms.

Introduction

In this talk, I will discuss how so-called hypercontractive inequalities can be used to give a new(ish) proof of a bound on the bias of multiplayer XOR games, which implies a (very) special case of a conjecture about quantum query algorithms.

Outline:

- Introduction to hypercontractivity
- XOR games
- The Bohnenblust-Hille inequality and its proof
- The Aaronson-Ambainis conjecture.

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

Noise and polynomials

- Any function $f : \{\pm 1\}^n \to \mathbb{R}$ can be expanded as a multilinear polynomial:

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

where $x_S = \prod_{i \in S} x_i$.

Noise and polynomials

- Any function $f : \{\pm 1\}^n \to \mathbb{R}$ can be expanded as a multilinear polynomial:

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

  where $x_S = \prod_{i \in S} x_i$.
- Parseval's equality: $\|f\|_2^2 = \sum_{S \subseteq [n]} \hat{f}(S)^2$.

Noise and polynomials

- Any function $f : \{\pm 1\}^n \to \mathbb{R}$ can be expanded as a multilinear polynomial:

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

where $x_S = \prod_{i \in S} x_i$.

- Parseval's equality: $\|f\|_2^2 = \sum_{S \subseteq [n]} \hat{f}(S)^2$.
- The noise operator has a nice "Fourier-side" description in terms of polynomials: for $g(x) = x_S$,

$$(T_\epsilon g)(x) = \epsilon^{|S|} x_S,$$

and by linearity, for any $f : \{\pm 1\}^n \to \mathbb{R}$,

$$(T_\epsilon f)(x) = \sum_{S \subseteq [n]} \epsilon^{|S|} \hat{f}(S) x_S.$$

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

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

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

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

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

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

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

(last two equalities: Parseval's equality)

Applications in quantum computation

The above inequality has recently found a number of 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 a number of 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: one more application.

Application: multiplayer XOR games

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

Application: 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: 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: 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: 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: 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

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

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

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 $\hat{f} \in \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$.
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 $p \geqslant 2k/(k+1)$,

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

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

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 $p \geqslant 2k/(k+1)$,

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

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.

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 $p = 2k/(k+1)$.

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]:

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

Proof

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

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

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

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

Proof

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

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

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

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

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

$$\sum_{i_{k/2+1},\ldots,i_k} \left(\sum_{i_1,\ldots,i_{k/2}} \hat{f}^2_{i_1,\ldots,i_k} \right)^{k/(k+2)}$$

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

Proof

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

$$\sum_{i_{k/2+1},\dots,i_k} \left(\sum_{i_1,\dots,i_{k/2}} \hat{f}^2_{i_1,\dots,i_k} \right)^{k/(k+2)}$$

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

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

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 $\sum_i x_i$ [Bačkurs '12].
- There are "L1" and "L2" versions of the conjecture [Bačkurs and Bavarian '13]; both are open. Here: the L2 version.

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

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

where $\hat{f}_{i_1,\dots,i_k} = \pm\alpha$ 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

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

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

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

where $\hat{f}_{i_1,\dots,i_k} = \pm\alpha$ 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

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

Corollary

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

Summary

We have:

- …used hypercontractivity to prove the Bohnenblust-Hille inequality;
- …and hence give strong bounds on the worst-case classical bias in XOR games;
- …and also prove a very special case of the Aaronson-Ambainis conjecture.

Open problems:

- Prove the Aaronson-Ambainis conjecture (using hypercontractivity!).

Summary

On a more concrete level:

- Can one generalise the Bohnenblust-Hille inequality to polynomials? i.e. prove that for any degree d multilinear polynomial $f : \{\pm 1\}^n \to \mathbb{R}$, and any $p \geqslant 2d/(d+1)$,

$$\|\hat{f}\|_p := \left(\sum_{S \subseteq [n]} |\hat{f}(S)|^p \right)^{1/p} \leqslant C_d \|f\|_\infty,$$

  where $C_d = \mathrm{poly}(d)$.

- This inequality holds for $C_d = 2^{O(d)}$ (Andreas Defant, personal communication).

- Would this imply the Aaronson-Ambainis conjecture?

Summary

On a more concrete level:

- Can one generalise the Bohnenblust-Hille inequality to polynomials? i.e. prove that for any degree d multilinear polynomial $f : \{\pm 1\}^n \to \mathbb{R}$, and any $p \geqslant 2d/(d+1)$,

$$\|\hat{f}\|_p := \left(\sum_{S \subseteq [n]} |\hat{f}(S)|^p \right)^{1/p} \leqslant C_d \|f\|_\infty,$$

  where $C_d = \text{poly}(d)$.

- This inequality holds for $C_d = 2^{O(d)}$ (Andreas Defant, personal communication).

- Would this imply the Aaronson-Ambainis conjecture?

Thanks!