Nets vs hierarchies for hard optimization problems

Aram Harrow
arXiv:1509.05065 (with Fernando Brandão)
in preparation (with Anand Natarajan and Xiaodi Wu)

QuiCS QMA(2)
1.8.2016

outline

1. separable states and operator norms
2. approximating the set of separable states
3. approximating general operator norms
4. the simple case of the simplex

entanglement and optimization

Definition: ρ is separable (i.e. not entangled) if it can be written as

$$\rho = \Sigma_i \; p_i \; |v_i\rangle\langle v_i| \otimes |w_i\rangle\langle w_i|$$

probability distribution

unit vectors

$$\text{Sep} = \text{conv}\{|v_i\rangle\langle v_i| \otimes |w_i\rangle\langle w_i|\}$$
$$= \text{conv}\{\rho\otimes\sigma\}$$
$$=$$

Weak membership problem: Given ρ and the promise that $\rho \in \text{Sep}$ or ρ is far from Sep, determine which is the case.

Optimization: $h_{\text{Sep}}(M) := \max \{ \text{tr}[M\rho] : \rho \in \text{Sep} \}$

operator norms

X:A->B

$\|X\|_{A\text{->}B} = \sup \|Xa\|_B \,/\, \|a\|_A$

operator norm

Examples

$l_2 \to l_2$	largest singular value
$l_\infty \to l_1$	MAX-CUT = $\max\{\langle \mathrm{vec}(X), a\otimes b\rangle : \|a\|_\infty, \|b\|_\infty \le 1\}$
$l_1 \to l_\infty$	$\max_{i,j} \lvert X_{i,j}\rvert = \max\{\langle \mathrm{vec}(X), a\otimes b\rangle : \|a\|_1, \|b\|_1 \le 1\}$
S_1 -> S_1 of $X\otimes \mathrm{id}$	channel distinguishability (cb norm, diamond norm)
S_1 -> S_p	max output p-norm, min output Rènyi-p entropy
$l_2 \to l_4$	hypercontractivity, small-set expansion
$S_1 \to S_\infty$	$h_{Sep} = \max\{ \langle \mathrm{Choi}(X), a\otimes b\rangle : \|a\|_{S_1}, \|b\|_{S_1} \le 1 \}$

complexity of h_{Sep}

$h_{Sep}(M) \pm 0.1\ ||M||_{2\to 2}$ at least as hard as
- planted clique [Brubaker, Vempala '09]
- 3-SAT[$\log^2(n)$ / polyloglog(n)] [H, Montanaro '10]

$h_{Sep}(M) \pm 100\ h_{Sep}(M)$ at least as hard as
- small-set expansion [Barak, Brandão, H, Kelner, Steurer, Zhou '12]

$h_{Sep}(M) \pm ||M||_{2\to 2}$ / poly(n) at least as hard as
- 3-SAT[n] [Gurvits '03], [Le Gall, Nakagawa, Nishimura '12]

complexity of $l_2 \rightarrow l_4$ norm

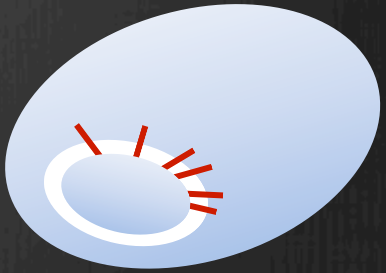
Unique Games (UG):

Given a system of linear equations: $x_i - x_j = a_{ij} \mod k$.
Determine whether $\geq 1-\epsilon$ or $\leq \epsilon$ fraction are satisfiable.

Small-Set Expansion (SSE):

Is the minimum expansion of a set with $\leq \delta n$ vertices $\geq 1-\epsilon$ or $\leq \epsilon$?

UG ≈ SSE ≤ 2->4



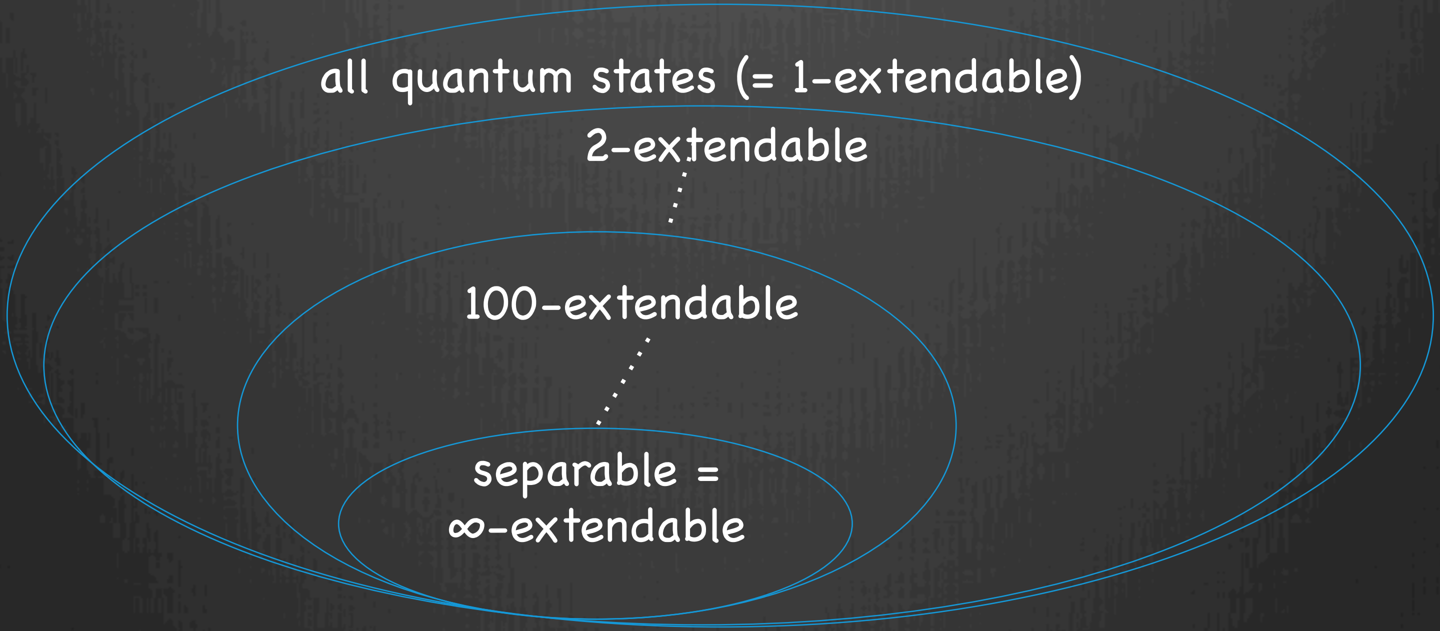
G = normalized adjacency matrix
P_λ = largest projector s.t. $G \geq \lambda P$

Theorem:

All sets of volume $\leq \delta$ have expansion $\geq 1 - \lambda^{O(1)}$
iff
$\|P_\lambda\|_{2\to 4} \leq n^{-1/4}/\delta^{O(1)}$

A hierachy of tests for entanglement

Definition: ρ^{AB} is k-extendable if there exists an extension $\rho^{AB_1 \ldots B_k}$ with $\rho^{AB} = \rho^{AB_i}$ for each i.



Algorithms: Can search/optimize over k-extendable states in time $n^{O(k)}$.

Question: How close are k-extendable states to separable states?

SDP hierarchies for h_{Sep}

$\mathrm{Sep}(n,m) = \mathrm{conv}\{\rho_1 \otimes \dots \otimes \rho_m : \rho_m \in D_n\}$
$\mathrm{SepSym}(n,m) = \mathrm{conv}\{\rho^{\otimes m} : \rho \in D_n\}$

bipartite

doesn't match hardness

Thm: If $M = \sum_i A_i \otimes B_i$ with $\sum_i |B_i| \leq I$, each $|A_i| \leq I$, then
$h_{Sep(n,2)}(M) \leq h_{k\text{-}ext}(M) \leq h_{Sep(n,2)}(M) + c\,(\log(n)/k)^{1/2}$

[Brandão, Christandl, Yard '10], [Yang '06], [Brandão, H '12], [Li, Winter '12]

multipartite

$$M = \sum_{i_1,\dots,i_m} c_{i_1,\dots,i_m} A^{(1)}_{i_1} \otimes \dots \otimes A^{(m)}_{i_m} \qquad \sum_i |A^{(j)}_i| \leq I \qquad |c_{i_1,\dots,i_m}| \leq 1$$

Thm:
ε-approx to $h_{SepSym(n,m)}(M)$ in time $\exp(m^2 \log^2(n)/\varepsilon^2)$.
ε-approx to $h_{Sep(n,m)}(M)$ in time $\exp(m^3 \log^2(n)/\varepsilon^2)$.

[Brandão, H '12], [Li, Smith '14]

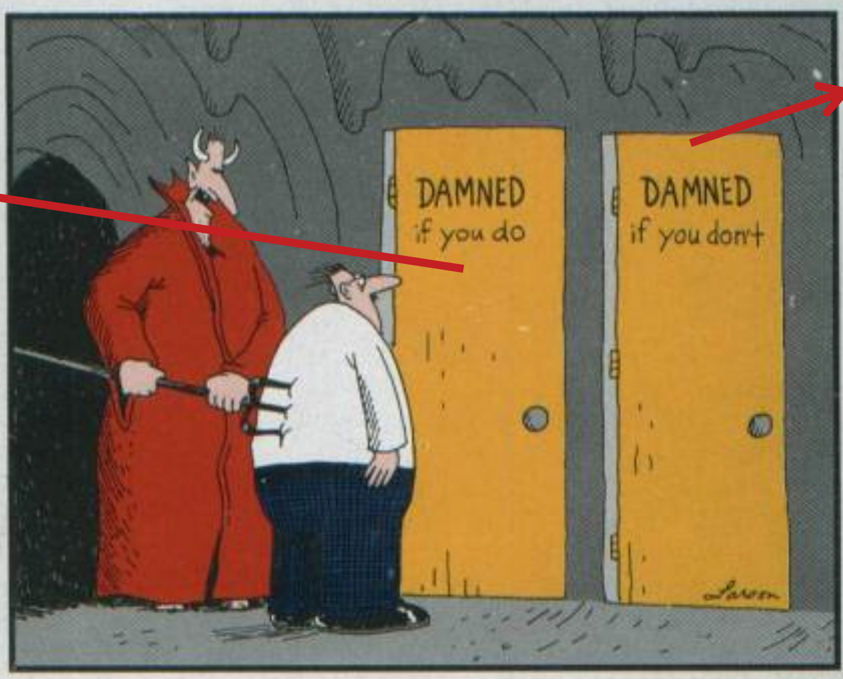
≈matches Chen-Drucker hardness

proof intuition

Measure extended state and get outcomes $p(a,b_1,\ldots,b_k)$.
Possible because of 1-LOCC form of M.

<u>case 1</u>

$p(a,b_1) \approx p(a) \cdot p(b_1)$



"C'mon, c'mon — it's either one or the other."

<u>case 2</u>

$p(a, b_2 \mid b_1)$

has less mutual information

questions

- Run-time $\exp(c \log^2(n) / \varepsilon^2)$ appears in both
    - Algorithm for M in 1-LOCC
    - Hardness for M in SEP.

Why? Can we bridge the gap?

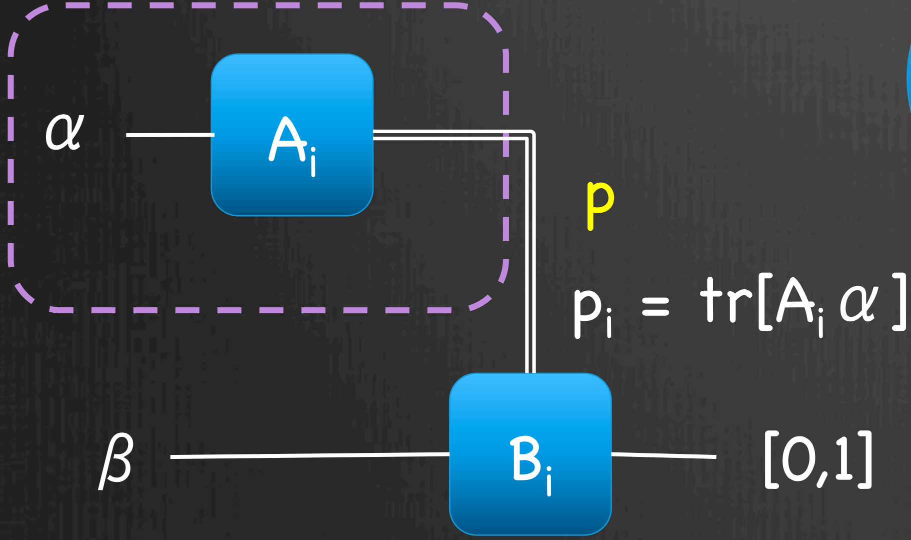
- Can we find multiplicative approximations, or otherwise use these approaches for SSE?

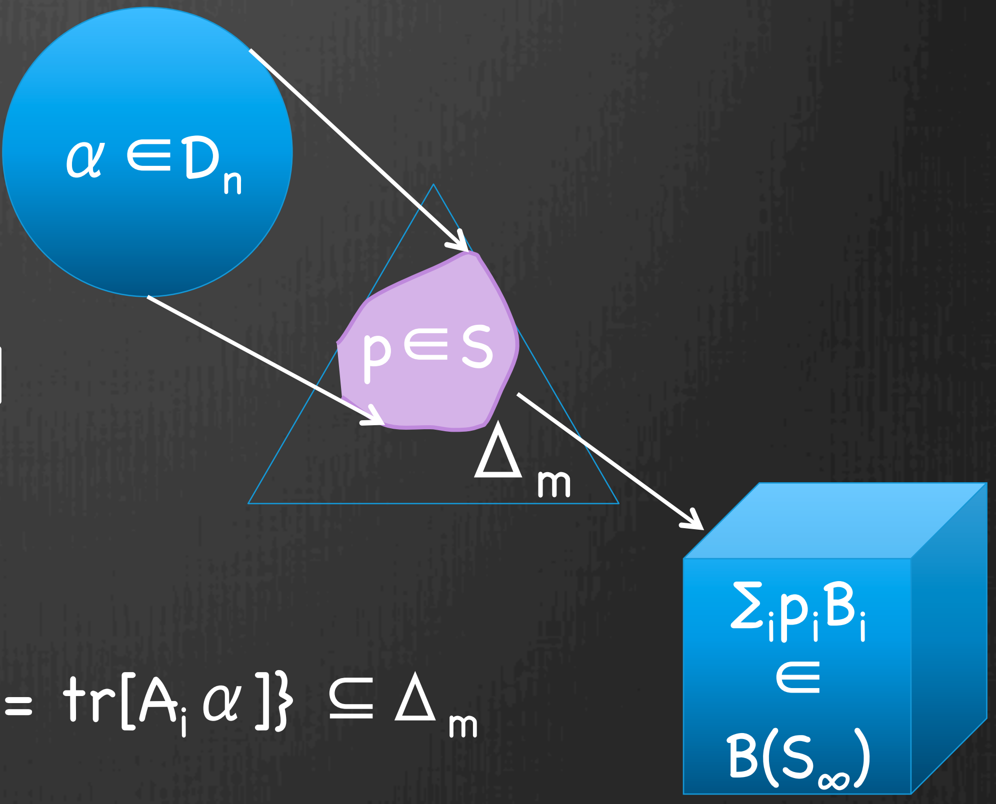
net-based algorithms

$M = \Sigma_{i\in[m]} A_i \otimes B_i$ with $\Sigma_i A_i \leq I$, each $|B_i| \leq I$, $A_i \geq 0$

Hierarchies estimate $h_{Sep}(M) \pm \varepsilon$ in time $\exp(\log^2(n)/\varepsilon^2)$

$$h_{Sep}(M) = \max_{\alpha,\beta} \mathrm{tr}[M(\alpha \otimes \beta)] = \max_{p\in S} \|p\|_B$$



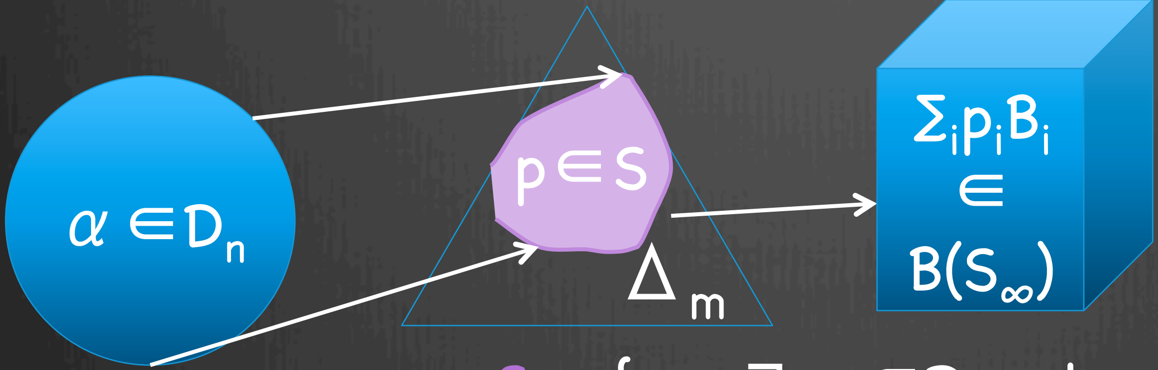


$S = \{p : \exists\, \alpha \in D_n \text{ s.t. } p_i = \mathrm{tr}[A_i \alpha]\} \subseteq \Delta_m$

$\|x\|_B = \|\Sigma_i x_i B_i\|_{2\text{->}2}$

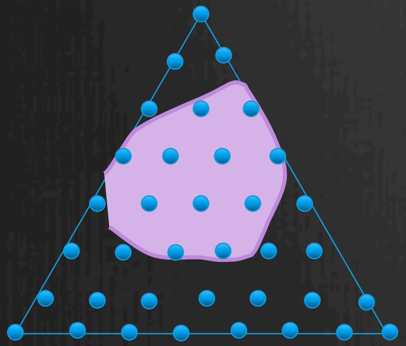
net-based algorithms

$h_{Sep}(M) = \max_{\alpha,\beta} \mathrm{tr}[M(\alpha \otimes \beta)] = \max_{p \in S} \|p\|_B$



$\|x\|_B = \|\Sigma_i\ x_i\ B_i\|_{2\text{->}2}$

$S = \{p : \exists\ \alpha \in D_n \text{ s.t. } p_i = \mathrm{tr}[A_i \alpha]\}$



Lemma: $\forall p \in \Delta_m$ $\exists q$ k-sparse (i.e. $\in \mathbb{Z}^m/k$) s.t. $\|p-q\|_B \leq c(\log(n)/k)^{1/2}$.

Pf: matrix Chernoff [Ahlswede-Winter]

Algorithm:

Enumerate over k-sparse q

- check whether $\exists p \in S$, $\|p-q\|_B \leq \varepsilon$
- if so, compute $\|q\|_B$

Performance

$k \simeq \log(n)/\varepsilon^2$, m=poly(n)

run-time

$O(m^k) = \exp(\log^2(n)/\varepsilon^2)$

nets for Banach spaces

X:A->B

$||X||_{A\to B}$ = sup $||Xa||_B$ / $||a||_A$ **operator norm**

$||X||_{A\to C\to B}$ = min $\{||Z||_{A\to C}\ ||Y||_{C\to B}$: X=YZ$\}$ **factorization norm**

Let A,B be arbitrary. $C = l_1^m$

Only changes are sparsification (cannot assume m≤poly(n))

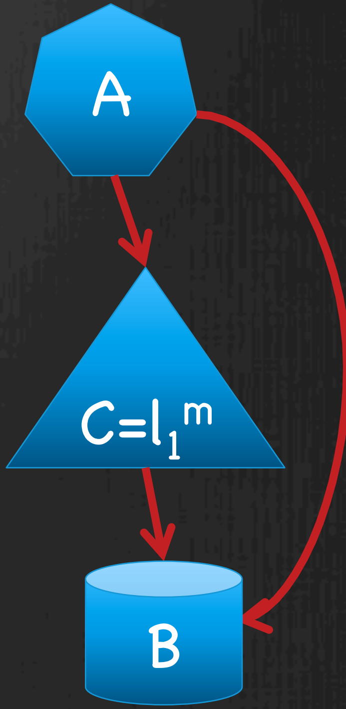
and operator Chernoff for B.

Type-2 constant: $T_2(B)$ is smallest λ such that

$$\mathbb{E}_{\epsilon_1,\ldots,\epsilon_n\in\{\pm1\}}\left\|\sum_{1=1}^{n}\epsilon_i Z_i\right\|_B^2 \leq \lambda^2\sum_{1=1}^{n}\|Z_i\|_B^2$$

result: $\|X\|_{A\to B} \pm \epsilon\|X\|_{A\to\ell_1^m\to B}$

estimated in time $\exp(T_2(B)^2 \log(m)/ \varepsilon^2)$



applications

$S_1 \rightarrow S_p$ norms of entanglement-breaking channels

$N(\rho) = \Sigma_i \operatorname{tr}[A_i \rho] B_i$, where $\Sigma_i A_i = I$, $\|B_i\|_1 = 1$.

Can estimate $\|N\|_{1\rightarrow p} \pm \varepsilon$ in time $n^{O(c)}$ where

$c = p/\varepsilon^2$ for $p \geq 2$

$c = (p/\varepsilon^p)^{1/(p-1)}$ for $1<p<2$

(uses bounds on $T_2(S_p)$ from [Ball-Carlen-Lieb '94]

low-rank measurements:

$h_{Sep}(\Sigma_i A_i \otimes B_i) \pm \varepsilon$ for

$\Sigma_i |A_i| = 1$, $\|B_i\|_\infty \leq 1$, rank $B_i \leq r$

in time $n^{O(r/\varepsilon^2)}$

$l_2 \rightarrow l_p$ for even $p \geq 4$

$$\|X\|_{2\to p}^p \pm \epsilon \|X\|_{2\to 2}^2 \|X\|_{2\to\infty}^{p-2}$$

in time $n^{O(p/\varepsilon^2)}$

Multipartite versions of 1-LOCC norm too [cf. Li-Smith '14]

ε -nets vs. SoS

Problem	ε -nets	SoS/info theory
$\max_{p\in\Delta} p^TAp$	BK ’02, KLP ‘06	DF ’80 BK ’02, KLP ‘06
approx Nash	LMM ’03	HNW ‘16
free games	AIM ‘14	BH ‘13
unique games	ABS ’10	BRS ‘11
small-set expansion	ABS ’10	BBHKSZ ’12
h_{Sep}	SW ‘11 BH ‘15	BCY ‘10 BH ’12 BKS ‘13

simplest version: polynomial optimization over the simplex

$\Delta_n = \{p \in \mathbb{R}^n : p \geq 0, \Sigma_i\, p_i = 1\}$

Given homogenous degree-d poly $f(p_1, \ldots, p_n)$, find $\max_p f(p)$.

NP-complete: given graph G with clique number α, $\max_p p^T A p = 1 - 1/\alpha$. [Motzkin-Strauss, '65]

Approximation algorithms

- Net: Enumerate over all points in $\Delta_n(k) := \Delta_n \cap \mathbb{Z}^n/k$.
- Hierarchy: min λ s.t. $(\Sigma_i\, p_i)^k\, (\lambda\, (\Sigma_i\, p_i)^d - f(p))$ has all nonnegative coefficients.

Thm: Each gives error $\leq (\max_p f(p) - \min_p f(p)) \exp(d) / k$ in time $n^{O(k)}$. [de Klerk, Laurent, Parrilo, '06]

sum-of-squares (SoS) proofs

Axioms:

$g_1(x) \geq 0$
$\vdots$
$g_m(x) \geq 0$ $\xrightarrow{\text{derive}}$ $f(x) \leq \lambda$

Rules:

1. polynomial operations
2. intermediate polys have deg $\leq k$
3. [optional: changes LP to SDP]

$r(x)^2 \geq 0$ for any polynomial $r(x)$

"Positivestellensatz" [Stengel '74]

hierarchies & SoS proofs

Given axioms: $\Sigma_i\ p_i = 1$ and $p_i \geq 0$
prove that $\lambda - f(p) \geq 0$.

Previous strategy:
$\lambda\,(\Sigma_i\ p_i)^d - f(p) = (\Sigma_i\ p_i)^k\ (\lambda\,(\Sigma_i\ p_i)^d - f(p) \geq 0$

difference is divisible
by $1 - \Sigma_i\ p_i$

LHS is nonnegative sum
of products of p_i

Dual is equivalent to net enumeration for modified objective function.
[Bomze, de Klerk '02] [de Klerk, Laurent, Sun '14]

k-extendable hierarchy

For a deg-d homogenous poly f(p), define $\mathrm{vec}(f) \in (R^n)^{\otimes d}$ to be the symmetric tensor such that $f(x) = \langle \mathrm{vec}(f), x^{\otimes d} \rangle$.

Then $\max_p f(p) = h_K(\mathrm{vec}(f))$ for
$K = \mathrm{conv}\{p^{\otimes d} : p \in \Delta_n\}$
$h_K(y) := \max_{x \in K} \langle x, y \rangle$

relaxation:
$q \in \Delta_{n^{d+k}}$ symmetric (aka "exchangeable")
$\pi = q^{(1,2,..,d)}$

convergence: [Diaconis, Freedman '80]
$\mathrm{dist}(\pi, \mathrm{conv}\{p^{\otimes d}\}) \leq O(d^2/k)$
$\rightarrow$ error $\|\mathrm{vec}(f)\|_\infty / k$ in time $n^{O(k)}$

Nash equilibria

Non-cooperative games:

Players choose strategies $p^A \in \Delta_m$, $p^B \in \Delta_n$.

Receive values $\langle V_A, p^A \otimes p^B \rangle$ and $\langle V_B, p^A \otimes p^B \rangle$.

Nash equilibrium: neither player can improve own value

ε -approximate Nash: cannot improve value by $> \varepsilon$

Correlated equilibria:

Players follow joint strategy $p^{AB} \in \Delta_{mn}$.

Receive values $\langle V_A, p^{AB} \rangle$ and $\langle V_B, p^{AB} \rangle$.

Cannot improve value by unilateral change.

- Can find in poly(m,n) time with LP.
- Nash equilibrium = correlated equilibrum with $p = p^A \otimes p^B$

finding (approximate) Nash eq

Known complexity:
Finding exact Nash eq. is PPAD complete.
Optimizing over exact Nash eq is NP-complete.

Algorithm for ε-approx Nash in time $\exp(\log(m)\log(n)/\varepsilon^2)$ based on enumerating over nets for Δ_m, Δ_n.
Planted clique and 3-SAT[$\log^2(n)$] reduce to optimizing over ε-approx Nash.

[Lipton, Markakis, Mehta '03], [Hazan-Krauthgamer '11], [Braverman, Ko, Weinstein '14]

New result [HNW16]: Another algorithm for finding ε-approximate Nash with the same run-time.

(uses k-extendable distributions)

algorithm for approx Nash

Search over $p^{AB_1\dots B_k} \in \Delta_{mn^k}$
such that the A:B_i marginal is a correlated equilibrium conditioned on any values for B_1, ..., B_{i-1}.

LP, so runs in time poly(mn^k)

Claim: Most conditional distributions are ≈ product.

Proof:

$\log(m) \geq H(A) \geq I(A:B_1\dots B_k) = \Sigma_{1\leq i\leq k}\ I(A:B_i|B_{<i})$

$\mathbb{E}_i\ I(A:B_i|B_{<i}) \leq \log(m)/k =: \ \varepsilon^2$

$\therefore\ k = \log(m)/\varepsilon^2$ suffices.

open questions

- Application to unique games, small-set expansion, etc. Which norms are the right ones here?
- Tight hardness results, e.g. for h_{Sep}.
- Explain the coincidences!