Nets vs hierarchies for hard optimization problems

Aram Harrow
arXiv:1509.05065 (with Fernando Brandão)
in preparation (with Anand Natarajan and Xiaodi Wu)
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**: \( \rho \) is separable (i.e. not entangled) if it can be written as
\[
\rho = \sum_i p_i |v_i\rangle\langle v_i| \otimes |w_i\rangle\langle w_i|
\]

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

**Weak membership problem**: Given \( \rho \) and the promise that \( \rho \in Sep \) or \( \rho \) is far from \( Sep \), determine which is the case.

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

\[ X: A \to B \]
\[ \|X\|_{A \to B} = \sup \|Xa\|_B / \|a\|_A \]

operator norm

Examples

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

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

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

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

$UG \approx SSE \leq 2 \rightarrow 4$

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

**Theorem:**
All sets of volume $\leq \delta$ have expansion $\geq 1 - \lambda^{O(1)}$
iff
$\|P_\lambda\|_{2 \rightarrow 4} \leq n^{-1/4}/\delta^{O(1)}$
A hierarchy of tests for entanglement

Definition: $\rho^{AB}$ is $k$-extendable if there exists an extension 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 $h_{\text{Sep}}$

$\text{Sep}(n,m) = \text{conv}\{\rho_1 \otimes \cdots \otimes \rho_m : \rho_m \in D_n\}$
$\text{SepSym}(n,m) = \text{conv}\{\rho \circ m : \rho \in D_n\}$

Thm: If $M = \sum_i A_i \otimes B_i$ with $\sum_i |B_i| \leq I$, each $|A_i| \leq I$, then

$h_{\text{Sep}(n,2)}(M) \leq h_{k-\text{ext}}(M) \leq h_{\text{Sep}(n,2)}(M) + c \left(\log(n)/k\right)^{1/2}$

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

Thm:

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

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

[Brandão, H ’12], [Li, Smith ’14]

$\approx$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 $M$.

case 1

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

case 2

$p(a, b_2 | b_1)$ has less mutual information

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

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

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

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

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

$\|x\|_B = \|\sum_i x_i B_i\|_{2\rightarrow 2}$
net-based algorithms

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

\[ \Delta_m \]

\[ \Sigma_i p_i B_i \in B(S_\infty) \]

\[ \|x\|_B = \|\Sigma_i x_i B_i\|_{2 \rightarrow 2} \]

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

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

**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**
- \( k = \log(n)/\varepsilon^2 \), \( m = \text{poly}(n) \)
- run-time
  \[ O(m^k) = \exp(\log^2(n)/\varepsilon^2) \]
nets for Banach spaces

\[ X : A \rightarrow B \]

\[ \|X\|_{A \rightarrow B} = \sup \|Xa\|_B / \|a\|_A \quad \text{operator norm} \]

\[ \|X\|_{A \rightarrow C \rightarrow B} = \min \{\|Z\|_{A \rightarrow C} \|Y\|_{C \rightarrow B} : X = YZ\} \quad \text{factorization norm} \]

Let A, B be arbitrary. \( C = l_1^m \)

Only changes are sparsification (cannot assume \( m \leq \text{poly}(n) \)) and operator Chernoff for B.

**Type-2 constant:** \( T_2(B) \) is smallest \( \lambda \) such that

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

result:

\[ \|X\|_{A \rightarrow B} \pm \epsilon \|X\|_{A \rightarrow \ell_1^m \rightarrow B} \]

estimated in time \( \exp(T_2(B)^2 \log(m)/\epsilon^2) \)
applications

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

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

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

$c = \frac{p}{\varepsilon^2}$ \quad for $p \geq 2$

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

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

low-rank measurements:

$\text{h}_{\text{Sep}}(\sum_i A_i \otimes B_i) \pm \varepsilon$ for

$\sum_i |A_i| = 1$, $\|B_i\|_\infty \leq 1$, $\text{rank } B_i \leq r$

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

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

$\|X\|_{2 \rightarrow p} \pm \varepsilon \|X\|_{2 \rightarrow 2} \|X\|_{2 \rightarrow \infty}$

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

Multipartite versions of 1-LOCC norm too [cf. Li-Smith ’14]
### $\epsilon$-nets vs. SoS

<table>
<thead>
<tr>
<th>Problem</th>
<th>$\epsilon$-nets</th>
<th>SoS/info theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\max_p \in \Delta \ p^T A p$</td>
<td>BK '02, KLP '06</td>
<td>DF '80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BK '02, KLP '06</td>
</tr>
<tr>
<td>approx Nash</td>
<td>LMM '03</td>
<td>HNW '16</td>
</tr>
<tr>
<td>free games</td>
<td>AIM '14</td>
<td>BH '13</td>
</tr>
<tr>
<td>unique games</td>
<td>ABS '10</td>
<td>BRS '11</td>
</tr>
<tr>
<td>small-set expansion</td>
<td>ABS '10</td>
<td>BBHKSZ '12</td>
</tr>
<tr>
<td>$h_{\text{Sep}}$</td>
<td>SW '11 BH '15</td>
<td>BCY '10 BH '12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BKS '13</td>
</tr>
</tbody>
</table>
simplest version: polynomial optimization over the simplex

\[ \Delta_n = \{ p \in \mathbb{R}^n : p \geq 0, \sum_i p_i = 1 \} \]

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

**NP-complete**: given graph \( G \) with clique number \( \alpha \),
\[
\max_p p^T A p = 1 - \frac{1}{\alpha}.
\]  
[Motzkin-Strauss, '65]

**Approximation algorithms**
- **Net**: Enumerate over all points in \( \Delta_n(k) := \Delta_n \cap \mathbb{Z}^n/k \).
- **Hierarchy**: \( \min \lambda \text{ s.t. } (\sum_i p_i)^k (\lambda (\sum_i p_i)^d - f(p)) \) 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:
\[ g_1(x) \geq 0 \]
\[ \vdots \]
\[ g_m(x) \geq 0 \]

\[ \text{derive} \quad f(x) \leq \lambda \]

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 \( r(x) \)

“Positivstellensatz” [Stengel ’74]
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 $\text{vec}(f) \in (\mathbb{R}^n)^{\otimes d}$ to be the symmetric tensor such that $f(x) = \langle \text{vec}(f), x^{\otimes d} \rangle$.

Then $\max_p f(p) = h_K(\text{vec}(f))$ for

$K = \text{conv}\{p^{\otimes d} : p \in \Delta_n\}$

$h_K(y) := \max_{x \in K} \langle x, y \rangle$

**relaxation:**

$q \in \Delta_{nd+k}$ symmetric (aka “exchangeable”)

$\pi = q^{(1,2,\ldots,d)}$

**convergence:** [Diaconis, Freedman ’80]

$\text{dist}(\pi, \text{conv}\{p^{\otimes d}\}) \leq O(d^2/k)$

$\Rightarrow$ error $\|\text{vec}(f)\|_\infty / k$ 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
\( \varepsilon \)-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 \( \text{poly}(m,n) \) time with LP.
- Nash equilibrium = correlated equilibrium 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 $\varepsilon$-approx Nash in time $\exp(\log(m)\log(n)/\varepsilon^2)$ based on enumerating over nets for $\Delta_m$, $\Delta_n$. Planted clique and 3-SAT[$\log^2(n)$] reduce to optimizing over $\varepsilon$-approx Nash.

[New result [HNW16]: Another algorithm for finding $\varepsilon$-approximate Nash with the same run-time. (uses k-extendable distributions)]
algorithm for approx Nash

Search over \( p^{AB_1\ldots B_k} \in \Delta_{mn^k} \)
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 \( \approx \) product.

Proof:
\[
\log(m) \geq H(A) \geq I(A:B_1\ldots B_k) = \sum_{1 \leq i \leq k} I(A:B_i|B_{<i})
\]
\[
\mathbb{E}_i I(A:B_i|B_{<i}) \leq \log(m)/k =: \varepsilon^2
\]
\[
\therefore k = \log(m)/\varepsilon^2 \text{ suffices.}
\]
open questions

• Application to unique games, small-set expansion, etc. Which norms are the right ones here?

• Tight hardness results, e.g. for $h_{\text{sep}}$.

• Explain the coincidences!