quantum information and convex optimization

Aram Harrow (MIT)

Simons Institute
25.9.2014

outline

1. quantum information, entanglement, tensors
2. optimization problems from quantum mechanics
3. SDP approaches
4. analyzing LPs & SDPs using (quantum) information theory
5. ε-nets

quantum information ≈ noncommutative probability

	probability	quantum
states	$\Delta_n = \{p \in \mathbb{R}^n, p \geq 0, \sum_i p_i = \|p\|_1 = 1\}$	$D_n = \{\rho \in \mathbb{C}^{n \times n}, \rho \geq 0$ $\mathrm{tr}\,\rho = \|\rho\|_1 = 1\}$
measurement	$m \in \mathbb{R}^n$ $0 \leq m_i \leq 1$	$M \in \mathbb{C}^{n \times n}$, $0 \leq M \leq I$
"accept"	$\langle m, p \rangle$	$\langle M, \rho \rangle = \mathrm{tr}[M\rho]$
distance = best bias	$\frac{1}{2}\|p - q\|_1$	$\frac{1}{2}\|\rho - \sigma\|_1$

bipartite states

	probability	quantum
product states (independent)	$(p \otimes q)_{ij} = p_i \, q_j$	$(\rho \otimes \sigma)_{ij,kl} = \rho_{i,k} \, \sigma_{j,l}$
local measurement	$m \otimes 1_n$ or $1_n \otimes m$	$M \otimes I_n$ or $I_n \otimes M$
marginal state	$p_i^{(1)} = \Sigma_j \, p_{ij}$ $p_j^{(2)} = \Sigma_i \, p_{ij}$	$\rho_{i,j}^{(1)} = tr_2 \rho = \Sigma_k \, \rho_{ik,jk}$ $\rho_{i,j}^{(2)} = tr_1 \rho = \Sigma_k \, \rho_{ki,kj}$
separable states (not entangled)	$conv\{p \otimes q\} = \Delta_{n^2}$ (never entangled)	$Sep = conv\{p \otimes \sigma\} \subsetneq D_{n^2}$ (sometimes entangled)

entanglement and optimization

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

$\rho = \Sigma_i\, p_i\, v_i\, v_i^* \otimes w_i\, w_i^*$

probability distribution

unit vectors

$\text{Sep} = \text{conv}\{vv^* \otimes ww^*\}$
$= \text{conv}\{\rho\otimes\sigma\}$
=

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

complexity of h_{Sep}

Equivalent to: [H, Montanaro '10]

- computing $||T||_{inj} := \max_{x,y,z} |\langle T, x \otimes y \otimes z \rangle|$
- computing $||A||_{2\to 4} := \max_x ||Ax||_4 / ||x||_2$
- computing $||T||_{2\to op} := \max_x ||\Sigma_i x_i T_i||_{op}$
- maximizing degree-4 polys over unit sphere
- maximizing degree-O(1) polys over unit sphere

$h_{Sep}(M) \pm 0.1\, ||M||_{op}$ 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]

$h_{Sep}(M) \pm ||M||_{op}$ / poly(n) at least as hard as

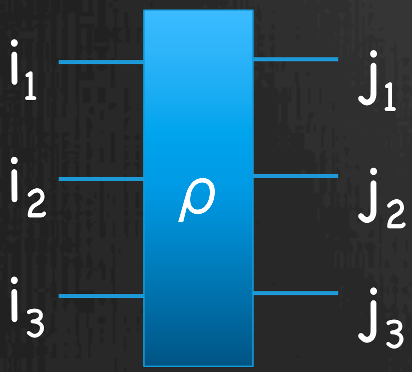
- 3-SAT[n] [Gurvits '03], [Le Gall, Nakagawa, Nishimura '12]

multipartite states

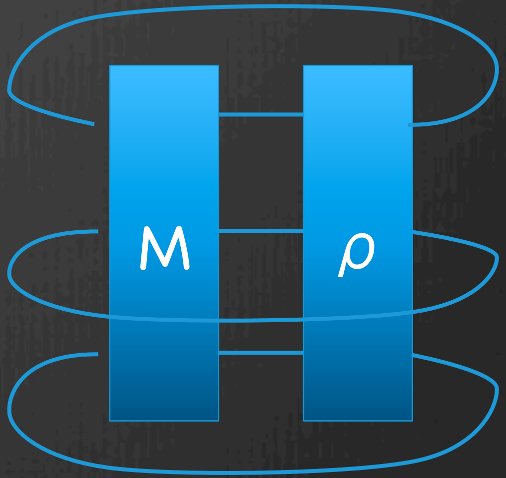
n d-dimensional systems → d^n dimensions
This explains:

- power of quantum computers
- difficulty of classically simulating q mechanics

Can also interpret as 2n-index tensors.



$\mathrm{tr}_3\,\rho$



$\mathrm{tr}[M\rho]$

local Hamiltonians

Definition: k-local operators are linear combinations of $\{A_1 \otimes A_2 \otimes \ldots \otimes A_n$: at most k positions have $A_i \neq I.\}$

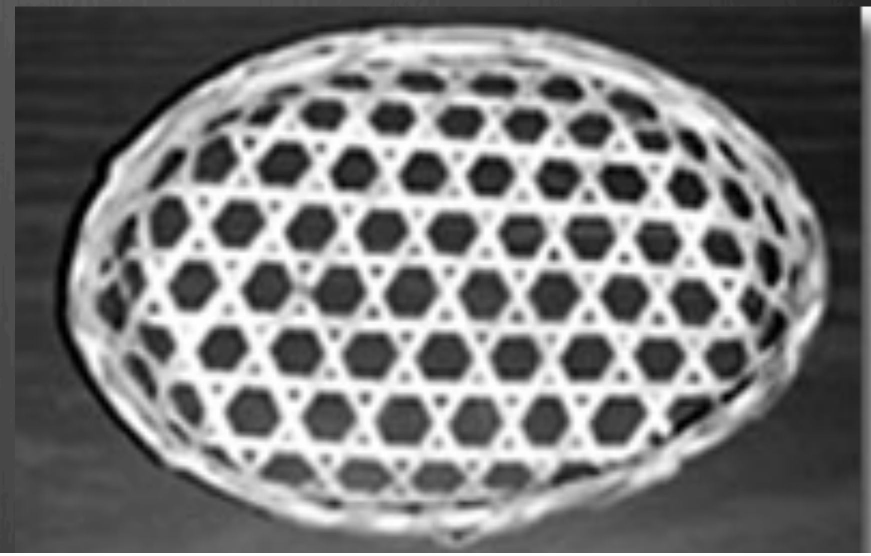
intuition: Diagonal case = k-CSPs = degree-k polys

Local Hamiltonian problem:
Given k-local H, find $\lambda_{min}(H) = \min_\rho \text{tr}[H\rho]$.

QMA-complete to estimate to accuracy $||H|| / \text{poly}(n)$.
qPCP conjecture: ... or with error $\varepsilon ||H||$

QMA vs NP:
do low-energy states have good classical descriptions?

kagome antiferromagnet



$ZnCu_3(OH)_6Cl_2$

Herbertsmithite

Spin-Liquid Ground State of the $S = 1/2$ Kagome Heisenberg Antiferromagnet

Simeng Yan,[1] David A. Huse,[2,3] Steven R. White[1]*

We use the density matrix renormalization group to perform accurate calculations of the ground state of the nearest-neighbor quantum spin $S = 1/2$ Heisenberg antiferromagnet on the kagome lattice. We study this model on numerous long cylinders with circumferences up to 12 lattice spacings. Through a combination of very-low-energy and small finite-size effects, our results provide strong evidence that, for the infinite two-dimensional system, the ground state of this model is a fully gapped spin liquid.

We consider the quantum spin $S = 1/2$ kagome Heisenberg antiferromagnet (KHA) with only nearest-neighbor isotropic exchange interactions (Hamiltonian $H = \Sigma\vec{S}_i \cdot \vec{S}_j$, where $\vec{S}_i$ and $\vec{S}_j$ are the spin operators for sites i and j, respectively) on a kagome lattice (Fig. 1A). This frustrated spin system has long been thought to be an ideal candidate for a simple, physically realistic model that shows a spin-liquid ground state (*1*–*3*). A spin liquid is a magnetic system that has "melted" in its ground state because of quantum fluctuations, so it has no spontaneously broken symmetries (*4*). A key problem in searching for spin liquids in two-dimensional (2D) models is that there are no exact or nearly exact analytical or computational methods to solve infinite 2D quantum lattice systems. For 1D systems, the density matrix renormalization group (DMRG) (*5*, *6*), the method we use here, serves in this capacity. In addition to its interest as an important topic in quantum magnetism, the search for spin liquids thus serves as a test-bed for the development of accurate and widely applicable computational methods for 2D many-body quantum systems.

[1]Department of Physics and Astronomy, University of California, Irvine, CA 92617, USA. [2]Department of Physics, Princeton University, Princeton, NJ 08544, USA. [3]Institute for Advanced Study, Princeton, NJ 08540, USA.

*To whom correspondence should be addressed. E-mail: srwhite@uci.edu

quantum marginal problem

Local Hamiltonian problem:
Given l-local H, find $\lambda_{min}(H) = \min_\rho \mathrm{tr}[H\rho]$.

Write $H = \Sigma_{|S|\leq l} H_S$ with H_S acting on systems S.
Then $\mathrm{tr}[H\rho] = \Sigma_{|S|\leq l} \mathrm{tr}[H_S \rho^{(S)}]$.

$O(n^l)$-dim convex optimization:
$\min \Sigma_{|S|\leq l} \mathrm{tr}[H_S \rho^{(S)}]$
such that $\{\rho^{(S)}\}_{|S|\leq l}$ are compatible.

QMA-complete to check

$O(n^k)$-dim relaxation: $(k \geq l)$
$\min \Sigma_{|S|\leq l} \mathrm{tr}[H_S \rho^{(S)}]$
such that $\{\rho^{(S)}\}_{|S|\leq k}$ are locally compatible.

Other Hamiltonian problems

Properties of ground state:
i.e. estimate $\mathrm{tr}[A\rho]$ for $\rho = \mathrm{argmin}\ \mathrm{tr}[H\rho]$

reduces to estimating $\lambda_{min}(H + \mu A)$

Non-zero temperature:
Estimate $\log \mathrm{tr}\, e^{-H}$ and derivatives

#P-complete, but some special cases are easier

(Noiseless) time evolution:
Estimate matrix elements of e^{iH}

BQP-complete

SOS hierarchies for q info

1. Goal: approximate Sep
   Relaxation: k-extendable + PPT (positive partial transpose)

2. Goal: λ_{min} for Hamiltonian on n qudits
   Relaxation: L : k-local observables → $\mathbb{R}$
   such that L[X*X] ≥ 0 for all k/2-local X.

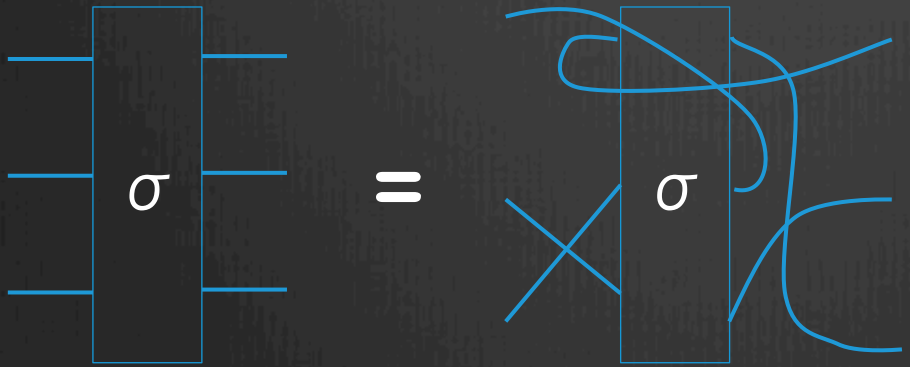
3. Goal: $\sup_{\rho,\{A\},\{B\} \text{ s.t. } \ldots} \sum_{xy} c_{xy} \langle \rho , A_x \otimes B_y \rangle$
   Relaxation: L : products of ≤k operators → $\mathbb{R}$
   such that L[p*p] ≥ 0 ∀ noncommutative poly p of deg ≤ k/2,
   and operators on different parties commute.

Non-commutative positivstellensatz [Helton-McCullough '04]

1. SOS hierarchies for Sep

SepProdR = $\mathrm{conv}\{xx^T \otimes xx^T : \|x\|_2=1, x \in \mathbb{R}^n\}$

relaxation [Doherty, Parrilo, Spedalieri '03]
$\sigma \in D_{nk}$ is a fully symmetric tensor
$\rho = \mathrm{tr}_{3\ldots k}[\sigma]$



Other versions use less symmetry.
e.g. k-ext + PPT

2. SOS hierarchies for λ_{min}

exact convex optimization: (hard)
min $\Sigma_{|S|\leq k}$ tr[$H_S \rho^{(S)}$]
such that $\{\rho^{(S)}\}_{|S|\leq k}$ are compatible.

equivalent:
min $\Sigma_{|S|\leq k}$ L[H_S] s.t.
$\exists\ \rho\ \ \forall$ k-local X, L[X] = tr[ρ X]

relaxation:
min $\Sigma_{|S|\leq k}$ L[H_S] s.t.
L[X*X] $\geq$ 0 for all k/2-local X
L[I]=1

classical analogue of Sep

quadratic optimization over simplex

$\max \{\langle Q, p\otimes p\rangle : p \in \Delta_n\} = h_{\mathrm{conv}\{p\otimes p\}}(Q)$

If Q=A, then max = 1 – 1 / clique#.

relaxation:

$q \in \Delta_{nk}$ symmetric (aka “exchangeable”)

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

convergence: [Diaconis, Freedman `80], [de Klerk, Laurent, Parrilo `06]

$\mathrm{dist}(\pi, \mathrm{conv}\{p\otimes p\}) \leq O(1/k)$

$\rightarrow$ error $\|Q\|_\infty / k$ 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: Another algorithm for finding ε-approximate Nash with the same run-time.

(uses k-extendable distributions)

algorithm for approx Nash

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

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

Claim: Most conditional distributions are ≈ product.

Proof:

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

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

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

SOS results for h_{Sep}

$\mathrm{Sep}(n,m) = \mathrm{conv}\{\rho_1 \otimes \ldots \otimes \rho_m : \rho_m \in D_n\}$

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

bipartite

doesn't match hardness

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

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

multipartite

$$M = \sum_{i_1,\ldots,i_m} c_{i_1,\ldots,i_m} A^{(1)}_{i_1} \otimes \cdots \otimes A^{(m)}_{i_m} \quad \sum_i |A^{(j)}_i| \le I \quad |c_{i_1,\ldots,i_m}| \le 1$$

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

SOS results for λ_{min}

$H = \mathbb{E}_{(i,j)\in E} H_{i,j}$ acts on $(\mathbb{C}^d)^{\otimes n}$ such that

- each $||H_{i,j}|| \leq 1$
- $|V| = n$
- (V,E) is regular
- adjacency matrix has $\leq r$ eigenvalues $\geq \text{poly}(\varepsilon/d)$

<u>Theorem</u>

$\lambda_{min}(H) \approx_\varepsilon h_{Sep(d,n)}(H)$

and can compute this to error ε with $r \cdot \text{poly}(d/\varepsilon)$ rounds of SOS, i.e. time $n^{r \cdot \text{poly}(d/\varepsilon)}$.

[Brandão-H, '13] based on [Barak, Raghavendra, Steurer '11]

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

$h_{Sep}(M) = \max_{\alpha,\beta} tr[M(\alpha \otimes \beta)] = \max_{p\in S} ||p||_B$
$S = \{p : \exists\, \alpha \text{ s.t. } p_i = tr[A_i \alpha]\} \subseteq \Delta_m$
$||x||_B = ||\sum_i x_i B_i||_{op}$

Lemma: $\forall p \in \Delta_m$ $\exists q$ k-sparse (each q_i = integer / k)
$||p-q||_B \leq c(\log(n)/k)^{1/2}$. 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

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

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

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

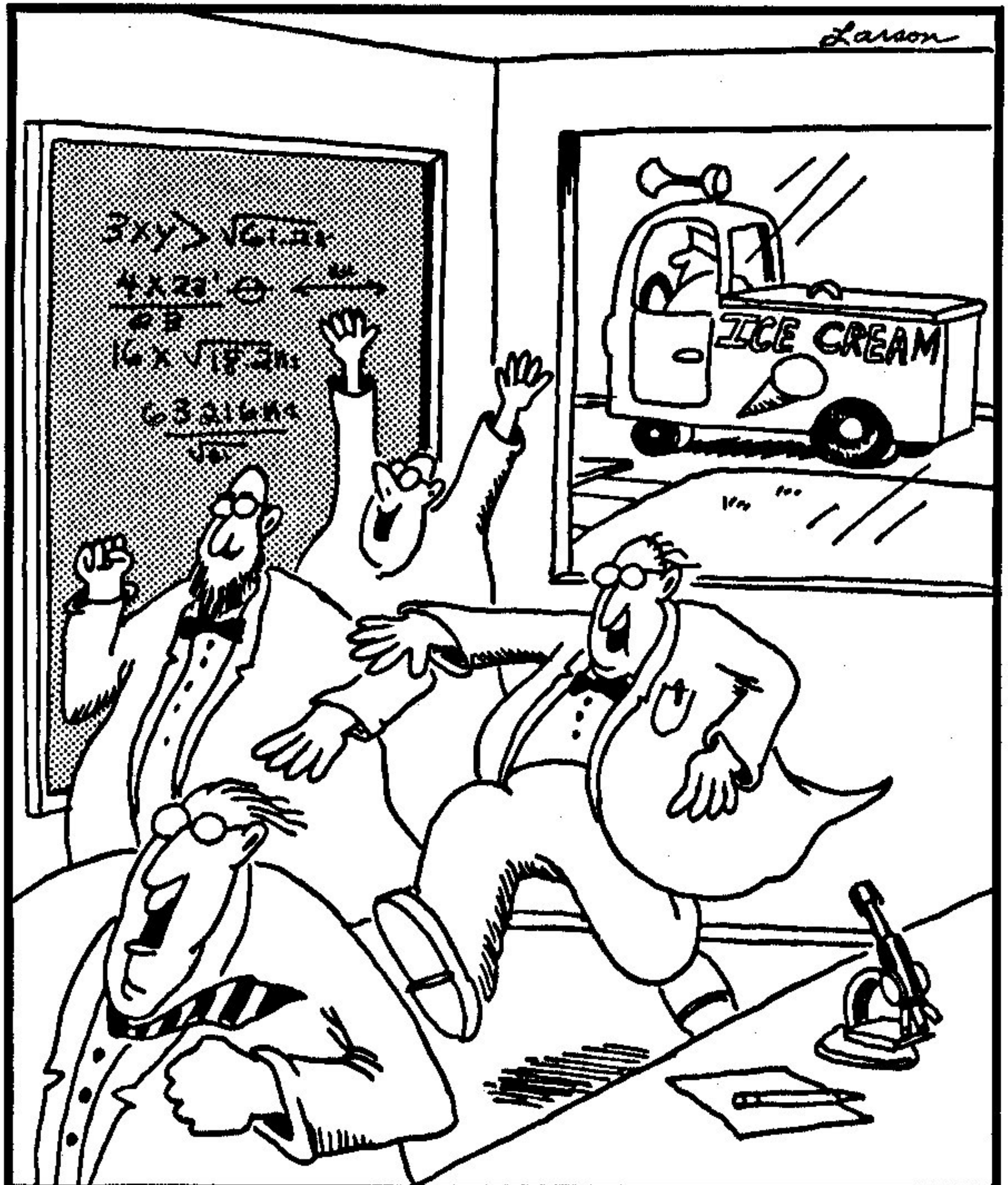
ε -nets vs. SOS

Problem	ε -nets	SOS/info theory
$\max_{p \in \Delta} p^T A p$	KLP '06	DF '80 KLP '06
approx Nash	LMM '03	H. '14
free games	AIM '14	Brandão-H '13
h_{Sep}	Shi-Wu '11 Brandão-H '14	BCY '10 Brandão-H '12 BKS '13

questions / references

- Application to 2->4 norm and small-set expansion.
- Matching quasipolynomial algorithms and hardness.
- simulating noisy/low-entanglement dynamics
- conditions under which Hamiltonians are easy to simulate
- Relation between hierarchies and nets
- Meaning of low quantum conditional mutual information

Hardness/connections	1001.0017
Relation to 2->4 norm, SSE	1205.4484
SOS for h_{Sep}	1210.6367
SOS for λ_{min}	1310.0017


Larson
ICE CREAM