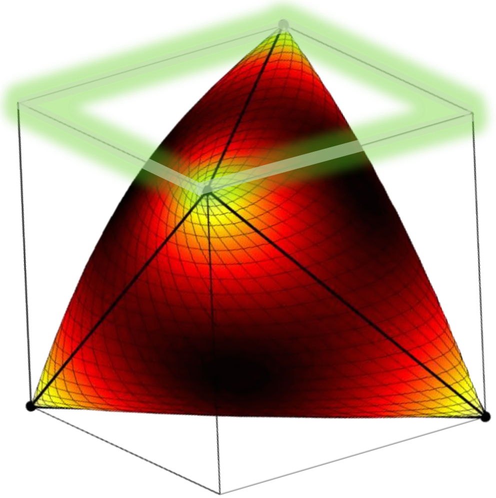
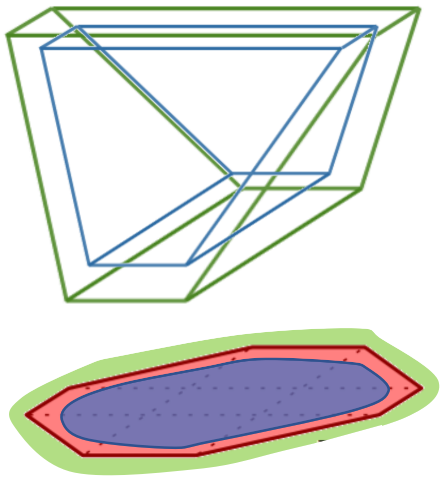
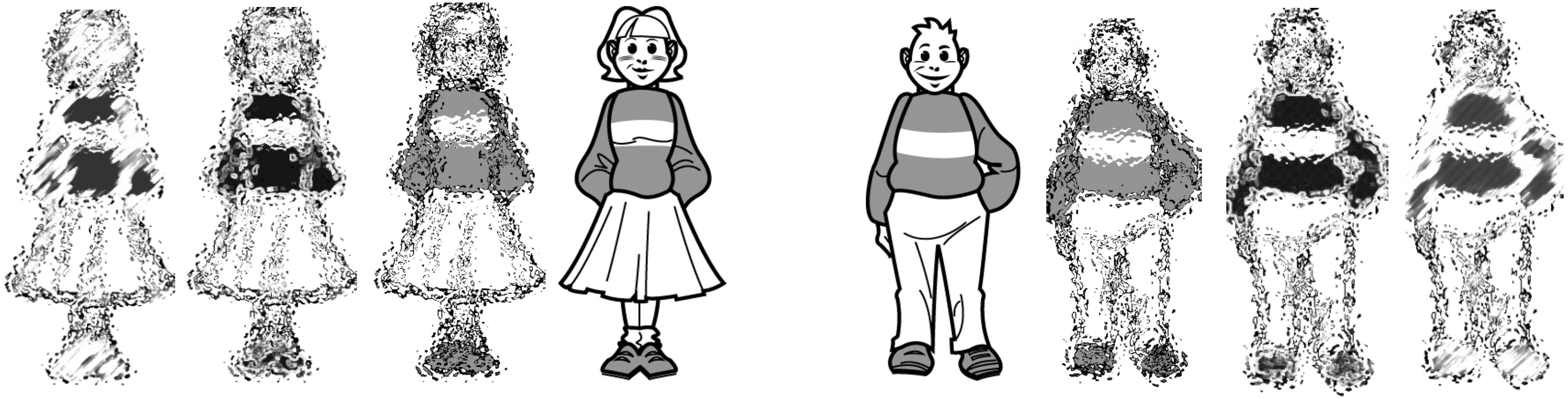
spectrahedral lifts (SDPs) and quantum learning

James R. Lee
University of Washington



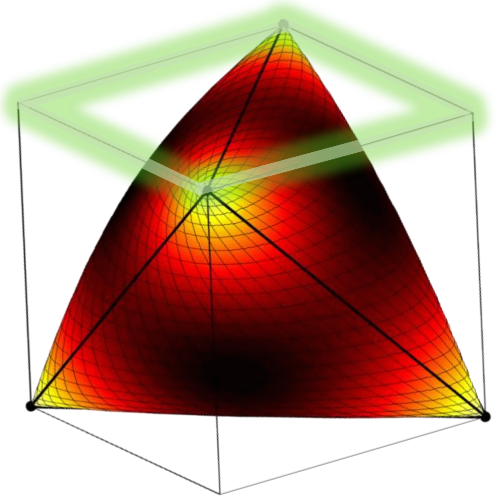


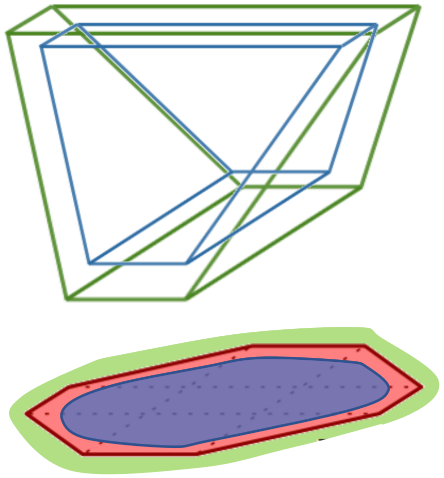


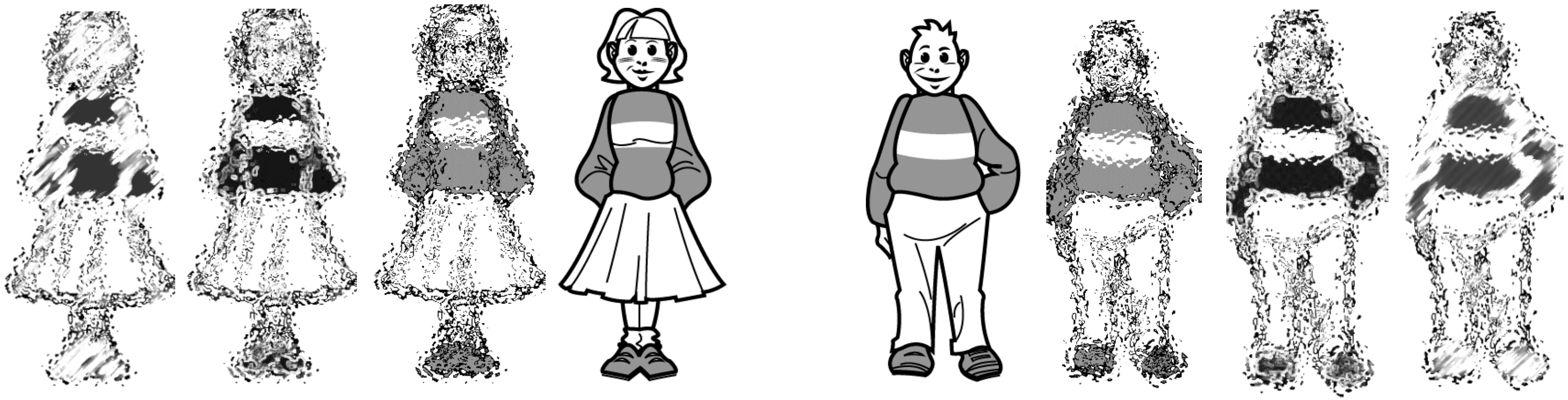
PSPACE 2017

spectrahedral lifts (SDPs) and quantum learning

James R. Lee
University of Washington





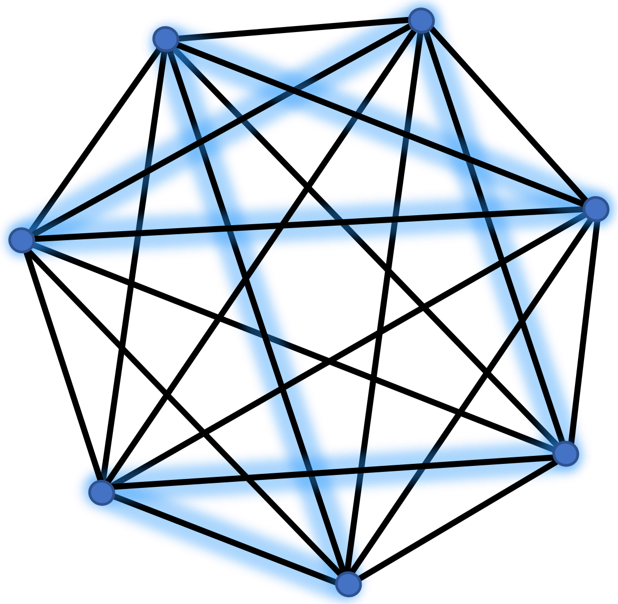


QIP 2017

Traveling Salesman Problem:

Given n cities $\{1,2,\ldots,n\}$ and costs $c_{ij} \geq 0$ for traveling between cities i and j, find the permutation π of $\{1,2,\ldots,n\}$ that minimizes

$$c_{\pi(1)\pi(2)} + c_{\pi(2)\pi(3)} + \cdots + c_{\pi(n)\pi(1)}$$



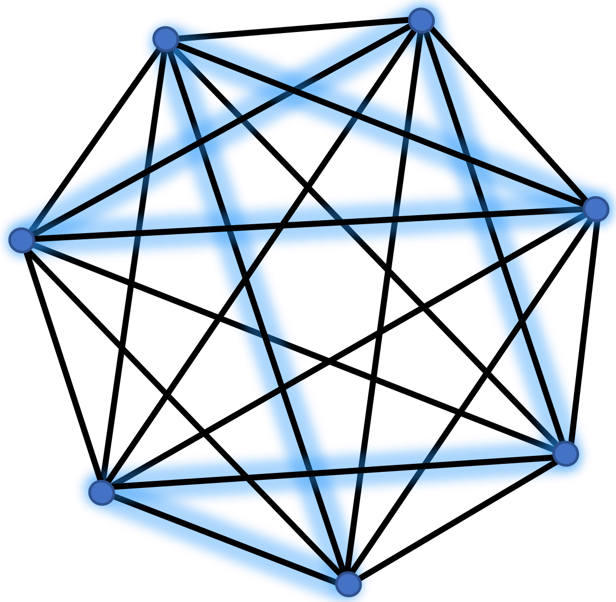
> Attempts to solve the traveling salesman problem and related problems of discrete minimization have led to a revival and a great development of the theory of polyhedra in spaces of *n* dimensions, which lay practically untouched—except for isolated results—since Archimedes. Recent work has created a field of unsuspected beauty and power, which is far from being exhausted.
>
> Gian Carlo Rota, 1969

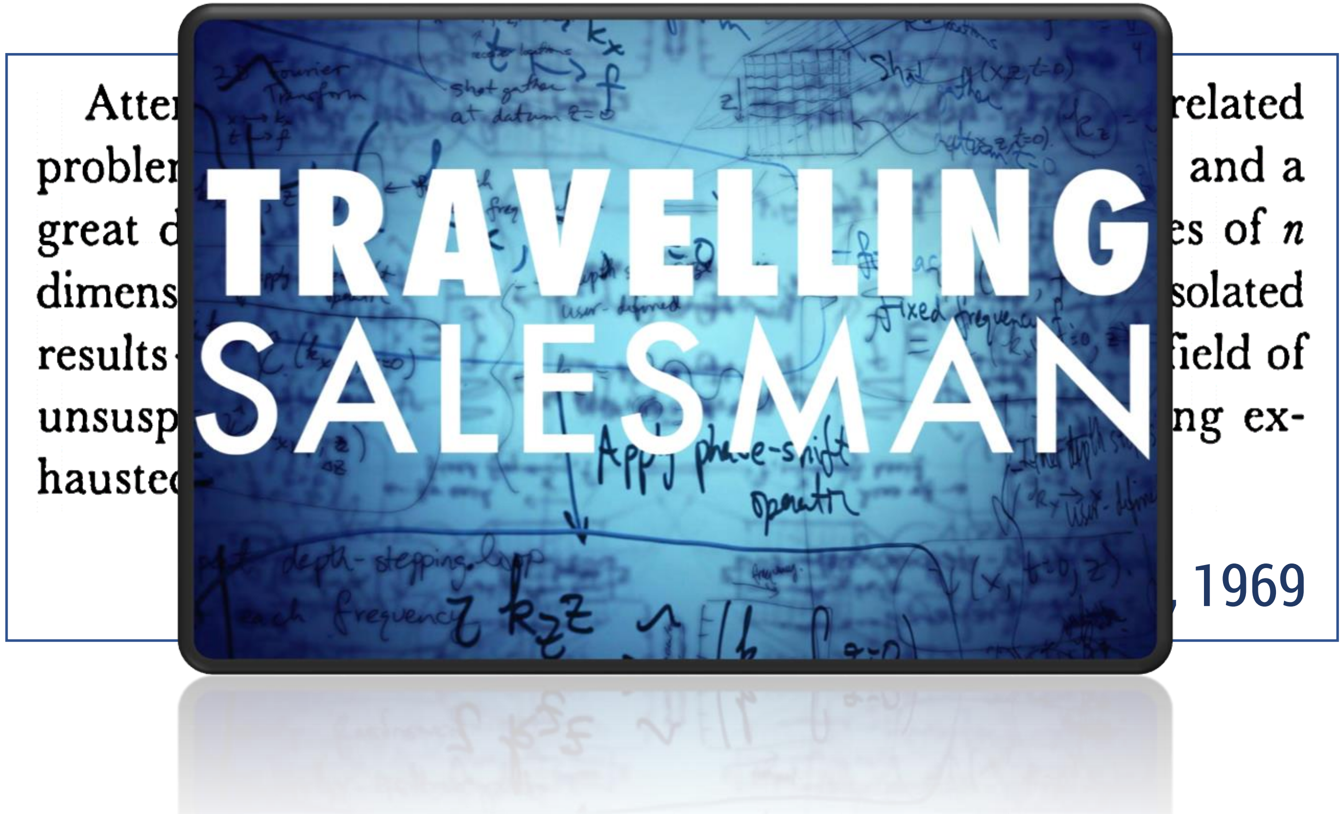
combinatorial optimization

Traveling Salesman Problem:

Given n cities $\{1,2,\ldots,n\}$ and costs $c_{ij} \geq 0$ for traveling between cities i and j, find the permutation π of $\{1,2,\ldots,n\}$ that minimizes

$$c_{\pi(1)\pi(2)} + c_{\pi(2)\pi(3)} + \cdots + c_{\pi(n)\pi(1)}$$

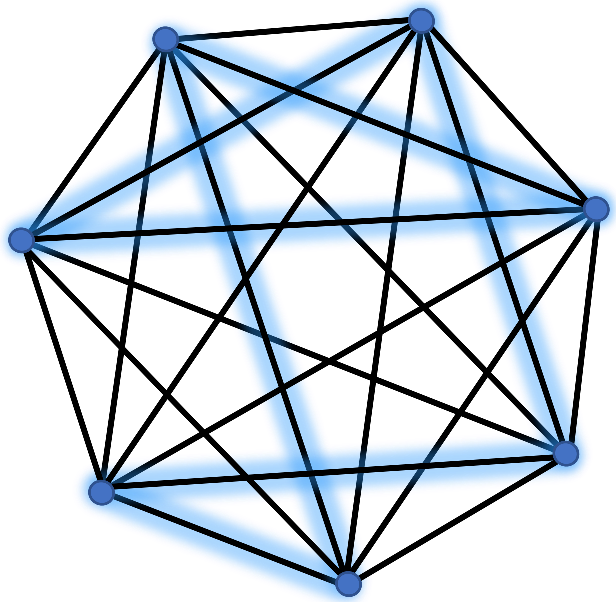




Traveling Salesman Problem:

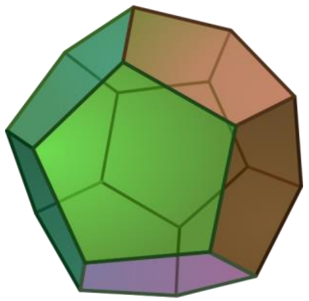
Given n cities $\{1,2,\dots,n\}$ and costs $c_{ij} \geq 0$ for traveling between cities i and j, find the permutation π of $\{1,2,\dots,n\}$ that minimizes

$$c_{\pi(1)\pi(2)} + c_{\pi(2)\pi(3)} + \cdots + c_{\pi(n)\pi(1)}$$



$$\mathrm{TSP}_n = \mathrm{conv}\left(\left\{1_\tau \in \{0,1\}^{\binom{n}{2}} : \tau \text{ is a tour}\right\}\right) \subseteq \mathbb{R}^{\binom{n}{2}}$$

Can find an optimal tour by minimizing a linear function over TSP_n: $\min \{ \langle c, x\rangle : x \in \mathrm{TSP}_n\}$

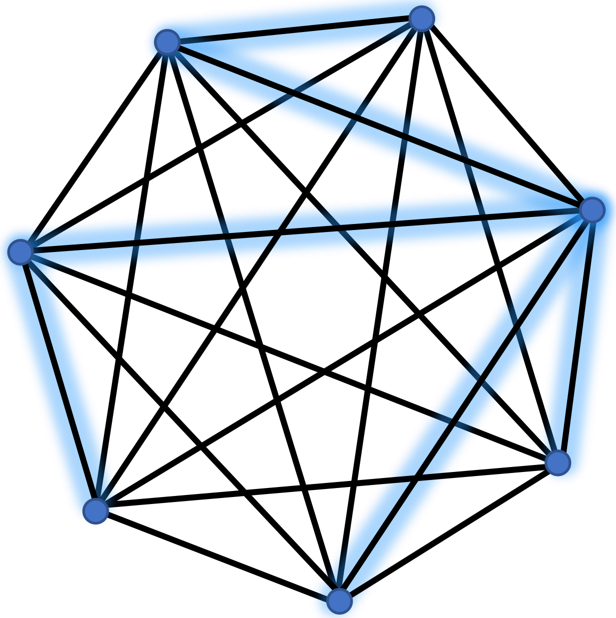


Problem: TSP_n has exponentially many facets!

One can tell the same (short) story for many polytopes associated to NP-complete problems.

Minimum Spanning Tree:

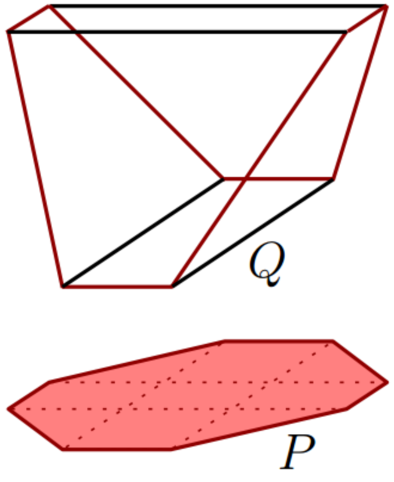
Given n cities $\{1,2,\ldots,n\}$ and costs $c_{ij} \geq 0$ between cities i and j, find a spanning tree of minimum cost.



$$\mathrm{ST}_n = \mathrm{conv}\left(\left\{1_\tau \in \{0,1\}^{\binom{n}{2}} : \tau \text{ is a spanning tree}\right\}\right)$$

Again, has exponentially many facets.

There is a **lift** of ST_n in n^3 dimensions with only $O(n^3)$ facets.
[Martin 1991]



a general model of (small) linear programs

Lifts of polytopes:

A **lift** Q of a polytope $P \subseteq \mathbb{R}^d$ is a polytope $Q \subseteq \mathbb{R}^N$ for $N \geq d$ such that Q linearly projects to P. If we can optimize linear functions over Q, then we can optimize over P.

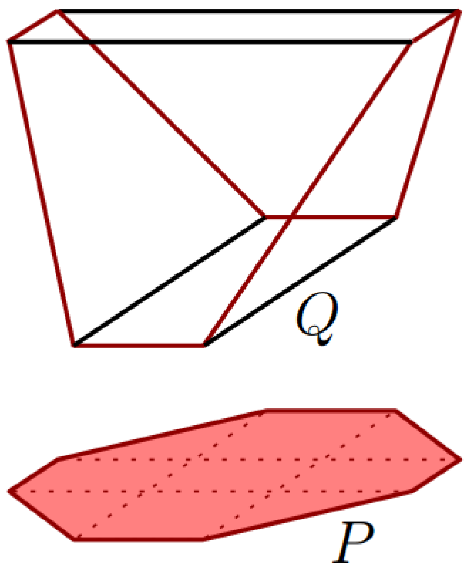
LP design point of view:

A lift corresponds to introducing (arbitrary) new variables and inequalities.

facets in the lift ⇔ # inequality constraints in the LP

Extension complexity:

The **extension complexity** of P is the minimal # of facets in a lift of P.



Examples with exponential savings:

Spanning trees, s-t flows, the permutahedron, …

a general model of (small) linear programs

Lifts of polytopes:

A **lift** Q of a polytope $P \subseteq \mathbb{R}^d$ is a polytope $Q \subseteq \mathbb{R}^N$ for $N \geq d$ such that Q linearly projects to P. If we can optimize linear functions over Q, then we can optimize over P.

Powerful model of computation.

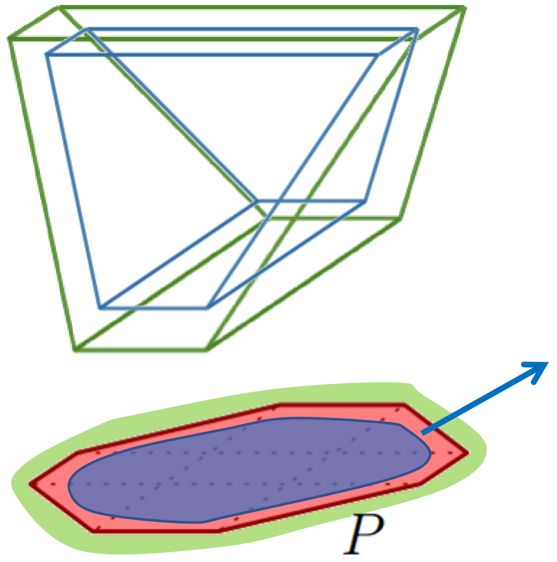
Even more powerful when we allow **approximation**.

Indication of power:

Dominant technique in the design of approximation algorithms.
Integrality gaps for LPs often lead to NP-hardness of approximation.

On the other hand:

Polynomial-size LPs for NP-hard problems would show that
$\mathrm{NP} \subseteq \mathrm{P/poly}$. [Rothvoss 2013]



a brief history of extended formulations

1980s: Fellow tries to prove that $P = NP$ by giving a linear program for TSP.

1989: Yannakakis (the referee) shows that every *symmetric* LP for TSP must have exponential size.

. . .

2011: Different set of fellows try to prove that $P = NP$ by giving an **asymmetric** LP for TSP

2012: Fiorini, Massar, Pokutta, Tiwary, de Wolf show that *every* LP for TSP must have exponential size.

2013: Chan, L, Raghavendra, Steurer show that no polynomial-size LP can approximate MAX-CUT within a factor better than 2.
[Goemans-Williamson 1998: SDPs can do factor ≈ 1.139.]

2014: Rothvoss shows that every LP for the matching polytope must have exponential size.

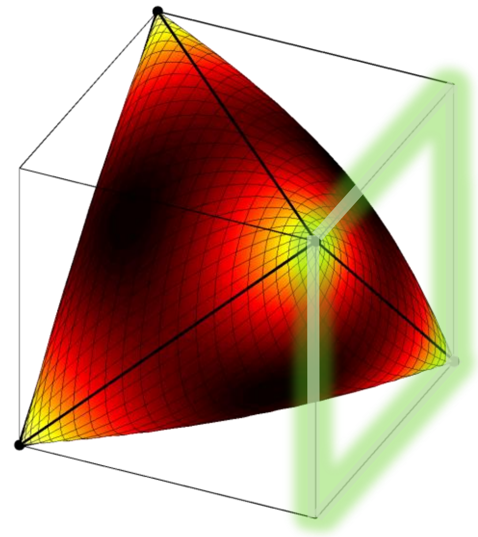
What about semidefinite programs?

semidefinite programs aka spectrahedral lifts

Let $\mathcal{S}_k^+$ denote the cone of $k \times k$ symmetric, positive semidefinite matrices.

A **spectrahedron** is the intersection $\mathcal{S}_k^+ \cap \mathcal{L}$ for some affine subspace $\mathcal{L}$.

This is precisely the feasible region of an SDP.



Definition: A polytope P admits a **PSD lift** of size k if P is a linear projection of a spectrahedron $\mathcal{S}_k^+ \cap \mathcal{L}$.

- Easy to see that minimal size of PSD lift is ≤ minimal size of polyhedral lift
- Assuming the Unique Games Conjecture, integrality gaps for SDPs translate mechanically into NP-hardness of approximation results.
  [Khot-Kindler-Mossel-O'Donnell 2004, Austrin 2007, Raghavendra 2008]
- Sometimes PSD lifts of polytopes are smaller than any polyhedral lift
  Gap of $O(d \log d)$ vs $\Omega(d^2)$ [Fawzi-Saunderson-Parrilo 2015]
  Exponential gaps known for approximation problems like MAX-CUT [Kothari-Meka-Raghavendra 2017]

SDP lifts cannot prove that $P = NP$

From [L-Raghavendra-Steurer 2015]:

Lower bounds on PSD lift size

The TSP_n, CUT_n, and $\mathrm{STAB}(G_n)$ polytopes do not admit PSD lifts of size $c^{n^{2/11}}$ (for some constant $c > 1$ and some family $\{G_n\}$ of n-vertex graphs)

Approximation hardness for constraint satisfaction problems

For max-constraint satisfaction problems, SDPs of polynomial size are equivalent in power to those arising from degree-$O(1)$ SoS relaxations.

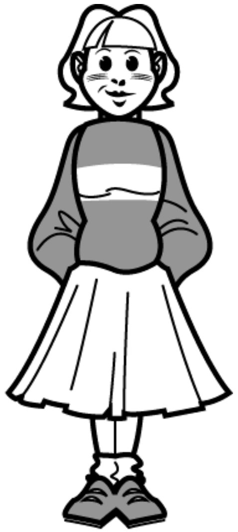
For instance, no family of polynomial-size SDP relaxations can achieve better than a $7/8$-approximation for MAX 3-SAT.

High level: Starting with a small SDP for some problem, we **quantum learn** a roughly equivalent sum-of-squares SDP on a subset of the variables.

communication (in expectation) model

A function $F: \mathcal{A} \times \mathcal{B} \to \mathbb{R}_+$

$a \in \mathcal{A}$



Alice

$\boldsymbol{M}_a$

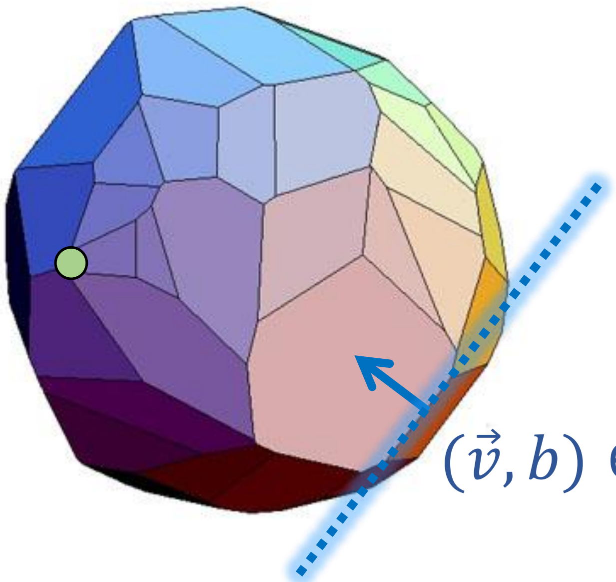
$b \in \mathcal{B}$

Bob

Successful protocol:
Bob outputs a random number $\boldsymbol{B}(b, \boldsymbol{M}_a) \geq 0$ such that for all $a \in \mathcal{A}, b \in \mathcal{B}$,

$$\mathbb{E}[\boldsymbol{B}(b, \boldsymbol{M}_a)] = F(a, b)$$

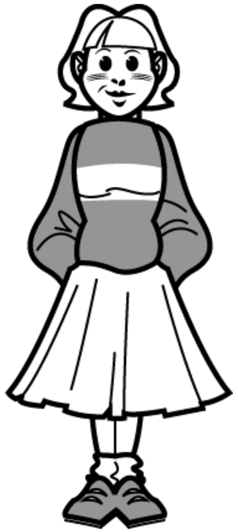
Polytope $P \subseteq \mathbb{R}^n$



$x \in V(P)$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$a \in \mathcal{A}$



Alice

$\boldsymbol{M}_a$

$b \in \mathcal{B}$

Bob

Successful protocol:

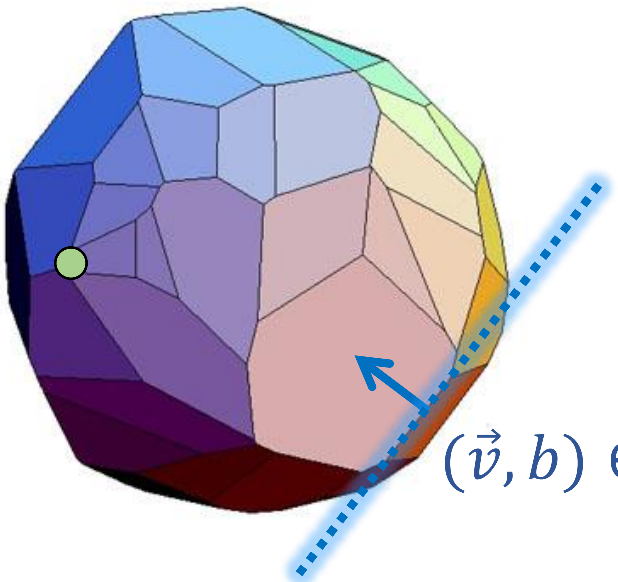
Bob outputs a random number $\boldsymbol{B}(b, \boldsymbol{M}_a) \geq 0$ such that for all $a \in \mathcal{A}, b \in \mathcal{B}$,

$$\mathbb{E}[\boldsymbol{B}(b, \boldsymbol{M}_a)] = F(a, b)$$

communication (in expectation) model

Polytope $P \subseteq \mathbb{R}^n$

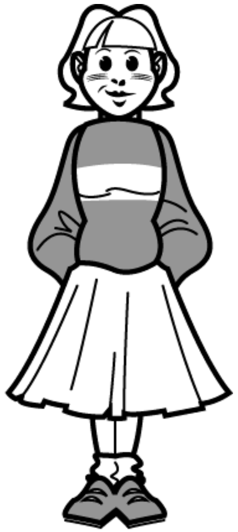
$x \in V(P)$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$x \in V(P)$



$\boldsymbol{M}_a$

Successful protocol:
Bob outputs a random number $\boldsymbol{B}(b, \boldsymbol{M}_a) \geq 0$
such that for all $a \in \mathcal{A}, b \in \mathcal{B}$,

$$\mathbb{E}[\boldsymbol{B}(b, \boldsymbol{M}_a)] = F(a, b)$$

Alice

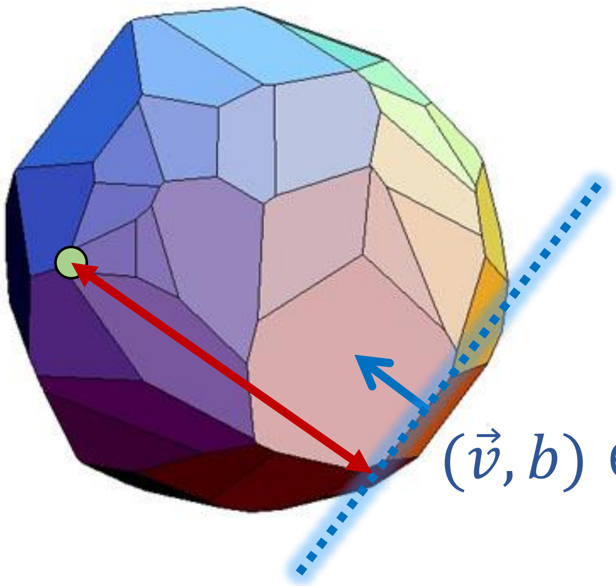
Bob

communication (in expectation) model

Polytope $P \subseteq \mathbb{R}^n$

$F\big((\vec{v}, b), x\big) = b - \langle \vec{v}, x \rangle$

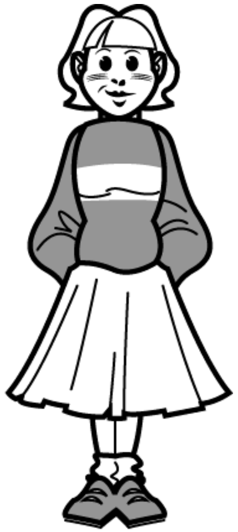
$x \in \{0,1\}^n$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$x \in \{0,1\}^n$



$\boldsymbol{M}_a$

Successful protocol:
Bob outputs a random number $\boldsymbol{B}(b, \boldsymbol{M}_a) \geq 0$
such that for all $a \in \mathcal{A}, b \in \mathcal{B}$,

$$\mathbb{E}[\boldsymbol{B}(b, \boldsymbol{M}_a)] = F(a, b)$$

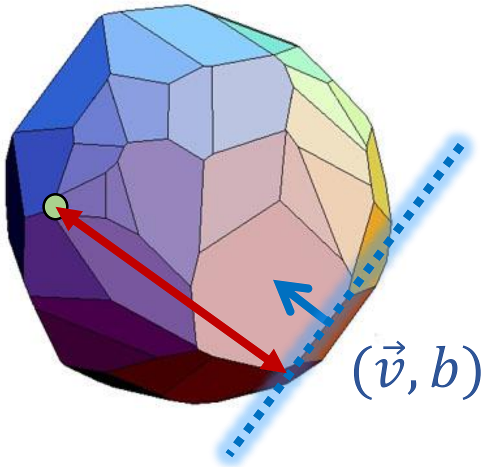
Alice

Bob

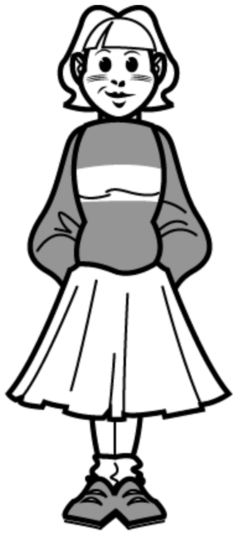
communication (in expectation) model

$P \subseteq \mathbb{R}^n$

$x \in \{0,1\}^n$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$



$\boldsymbol{M}_{(\vec{v},b)}$

$F((\vec{v}, b), x) = b - \langle \vec{v}, x \rangle$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

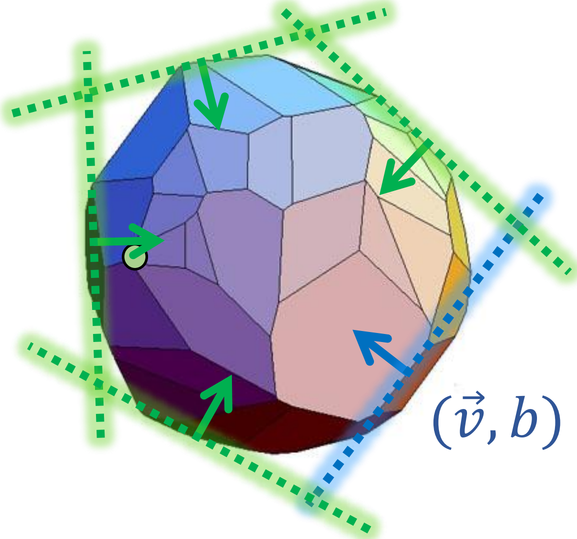
$x \in \{0,1\}^n$

$\boldsymbol{M}$ m-bit classical message $\Leftrightarrow$
P has a polyhedral lift of size 2^m

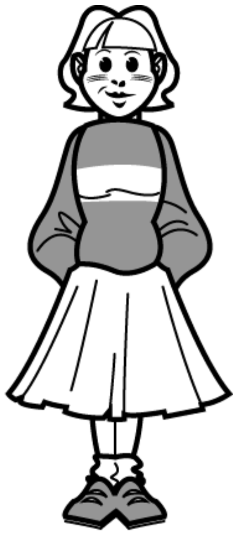
$\boldsymbol{M}$ m-qubit quantum message $\Leftrightarrow$
P has an SDP lift of size 2^m

Yannakakis factorization theorem
[+ Fiorini-Massar-Pokutta-Tiwary-de Wolf]

communication (in expectation) model



$(\vec{v}, b) \in \mathbb{R}^{n+1}$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$\boldsymbol{M}_{(\vec{v},b)}$

$F\big((\vec{v}, b), x\big) = b - \langle \vec{v}, x \rangle$

$x \in \{0,1\}^n$

$\boldsymbol{M}$ m-bit classical message $\Leftrightarrow$
P has a polyhedral lift of size 2^m

$\boldsymbol{M}$ m-qubit quantum message $\Leftrightarrow$
P has an SDP lift of size 2^m

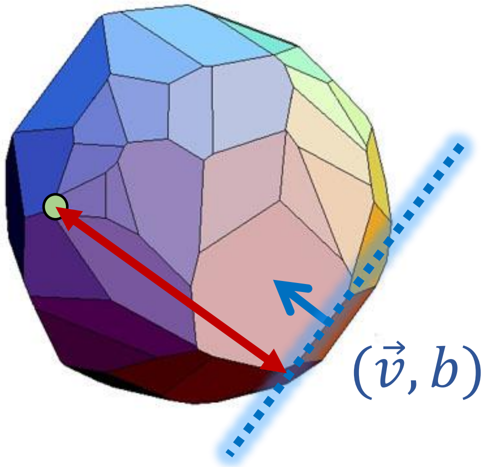
Yannakakis factorization theorem

[+ Fiorini-Massar-Pokutta-Tiwary-de Wolf]

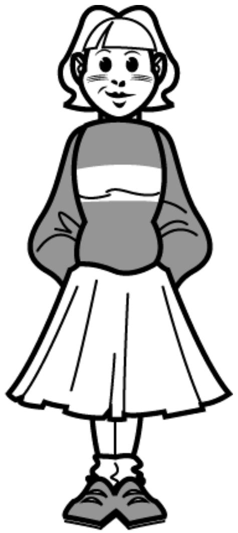
communication (in expectation) model

$P \subseteq \mathbb{R}^n$

$x \in \{0,1\}^n$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$



$\boldsymbol{M}_{(\vec{v},b)}$

$F\big((\vec{v}, b), x\big) = b - \langle \vec{v}, x \rangle$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$x \in \{0,1\}^n$

$\boldsymbol{M}$ m-bit classical message $\Leftrightarrow$
P has a polyhedral lift of size 2^m

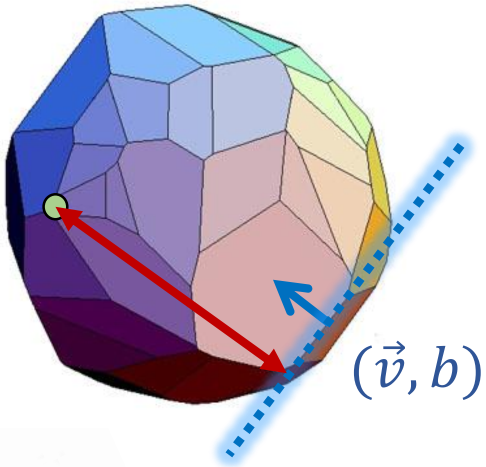
$\boldsymbol{M}$ m-qubit quantum message $\Leftrightarrow$
P has an SDP lift of size 2^m

Yannakakis factorization theorem
[+ Fiorini-Massar-Pokutta-Tiwary-de Wolf]

In this model, one-way communication and arbitrary communication are equivalent.
[Kaniewski-T. Lee-de Wolf 2014]

query protocols (k-hapless Bob)

$P \subseteq \mathbb{R}^n$



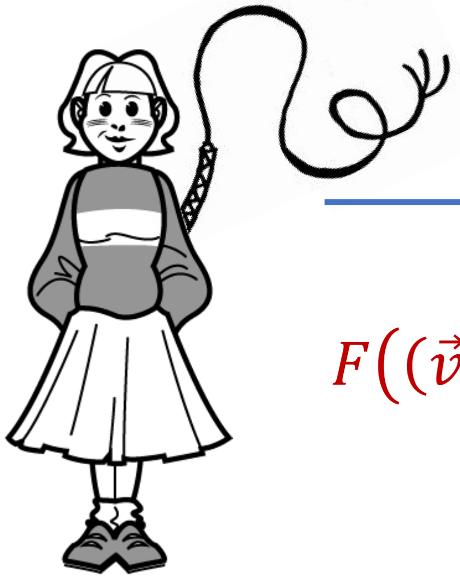
$x \in \{0,1\}^n$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

Alice specifies k bits
$\Leftrightarrow$ P has a "Sherali Adams" lift of size $\binom{n}{k} 2^k$

Alice specifies k bits in **superposition** + Bob measures
$\Leftrightarrow$ P has a "sum of squares" lift of degree $2k$

[Kaniewski-T. Lee-de Wolf 2014]



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$\boldsymbol{M}_{(\vec{v},b)}$ →

$F((\vec{v}, b), x) = b - \langle \vec{v}, x \rangle$

$x \in \{0,1\}^n$

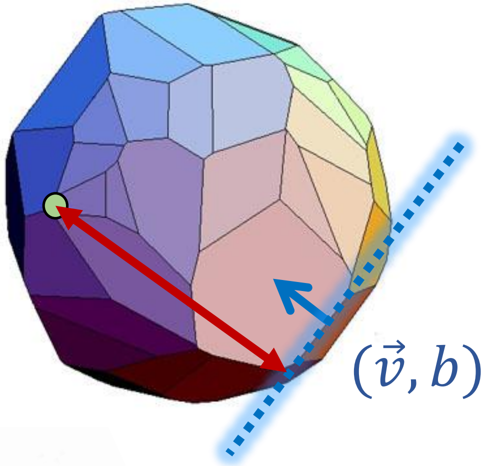
Suppose Alice's message says:
Look at k bits of your input in positions:
$i_1, i_2, \ldots, i_k$ and output 1 if you see values
$y_1, y_2, \ldots, y_k$

	1		0	0				1			0
1	1	0	0	0	1	0	1	0	1	1	0

query protocols (k-hapless Bob)

$P \subseteq \mathbb{R}^n$

$x \in \{0,1\}^n$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

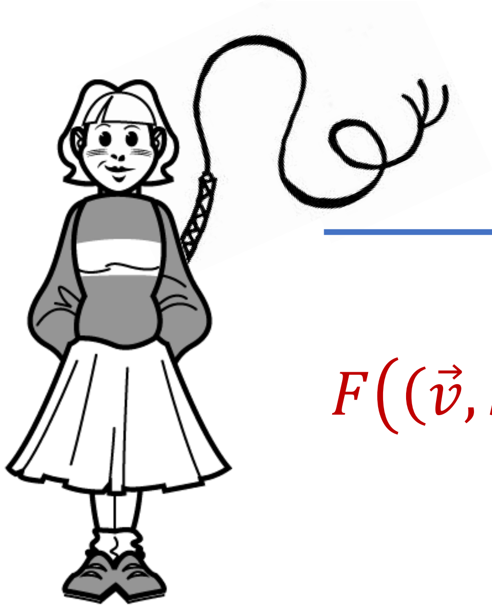
Alice specifies k bits

$\Leftrightarrow$ P has a "Sherali Adams" lift of size $\binom{n}{k}2^k$

Alice specifies k bits in **superposition** + Bob measures

$\Leftrightarrow$ P has a "sum of squares" lift of degree $2k$

[Kaniewski-T. Lee-de Wolf 2014]



$\boldsymbol{M}_{(\vec{v},b)}$

$F((\vec{v}, b), x) = b - \langle \vec{v}, x \rangle$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$x \in \{0,1\}^n$

Alice sends

$$\sum_{|S|=k} \sum_{y \in \{0,1\}^k} a_{S,y} \, |S, y\rangle$$

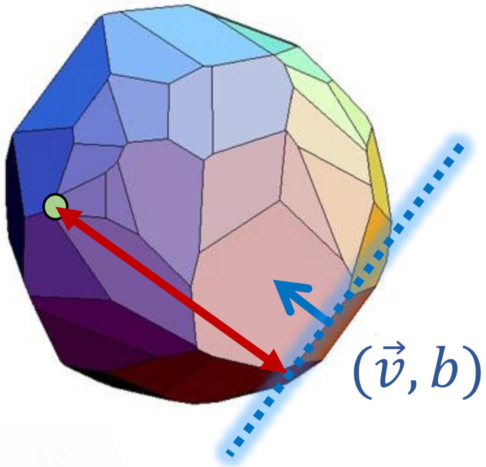
Bob computes

$$\sum_{|S|=k} \sum_{y \in \{0,1\}^k} a_{S,y} \, |S, y\rangle |\mathbf{1}_{x|_S = y}\rangle$$

and then does a computation + measurement (independent of x).

query protocols (k-hapless Bob)

$P \subseteq \mathbb{R}^n$



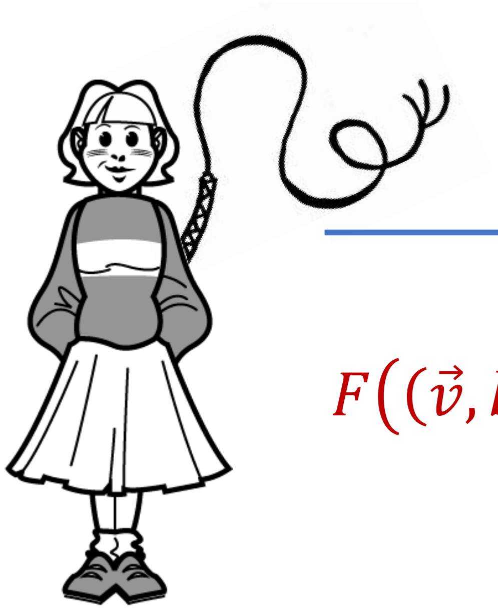
$x \in \{0,1\}^n$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

Alice specifies k bits
$\Leftrightarrow$ P has a "Sherali Adams" lift of size $\binom{n}{k} 2^k$

Alice specifies k bits in **superposition** + Bob measures
$\Leftrightarrow$ P has a "sum of squares" lift of degree $2k$

[Kaniewski-T. Lee-de Wolf 2014]



$\boldsymbol{M}_{(\vec{v},b)}$

$F((\vec{v}, b), x) = b - \langle \vec{v}, x \rangle$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$x \in \{0,1\}^n$

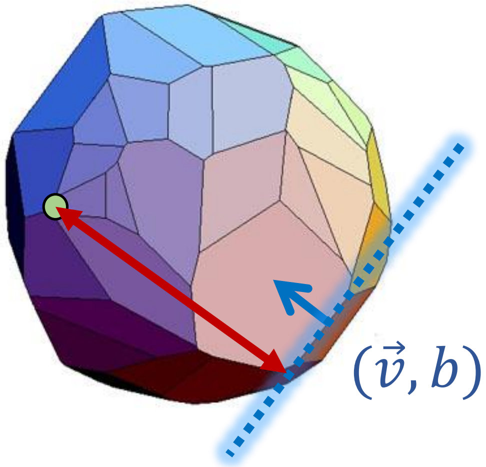
Low-degree sum of squares lifts are well-studied objects in optimization and proof complexity.

Goal:
Relate complexity of arbitrary quantum protocol to complexity of query protocols

simulation by a query protocol

$P \subseteq \mathbb{R}^n$

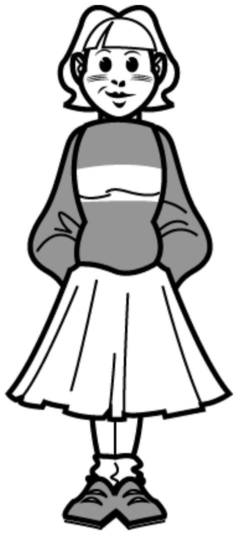
$x \in \{0,1\}^n$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

Translator converts Alice's message into a "query protocol" message

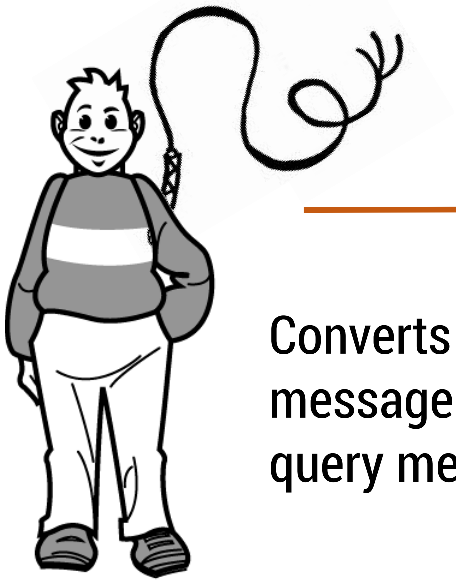
If a good enough translator exists, then
general protocols ≈ query protocols
⇒ general PSD lifts ≈ SoS lifts



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$\boldsymbol{M}_{(\vec{v},b)}$

$F\big((\vec{v}, b), x\big) = b - \langle \vec{v}, x \rangle$



Translator Bob

$\widetilde{\boldsymbol{M}}$

Converts Alice's m-bit message to an $\approx m$-bit query message

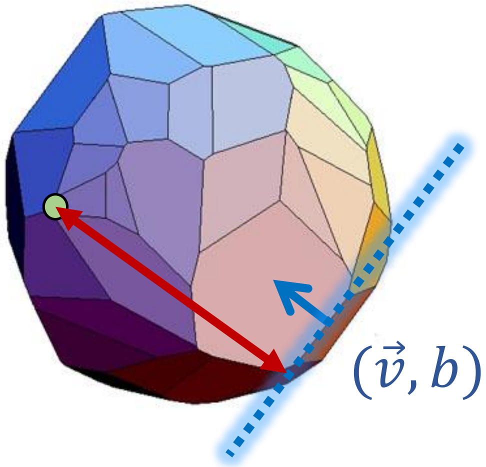
1	1	0	0	0	1	0	1	0	1	1	0

$x \in \{0,1\}^n$

simulation by a query protocol

$P \subseteq \mathbb{R}^n$

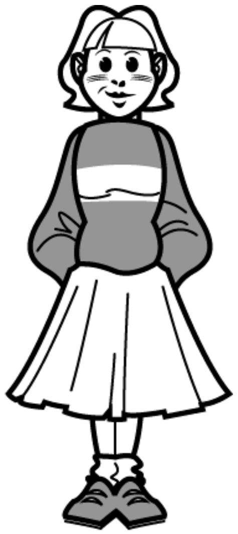
$x \in \{0,1\}^n$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

Strategy 1: Be smart

Look at Alice's message, make a judgement about what "bits" are most influential (now in a quantum sense), query them, condition(?), measure(?), recurse on the conditional(?) quantum state.



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

$\boldsymbol{M}_{(\vec{v},b)}$

$F\big((\vec{v}, b), x\big) = b - \langle \vec{v}, x \rangle$

Translator Bob

$\widetilde{\boldsymbol{M}}$

Converts Alice's m-bit message to an $\approx m$-bit query message

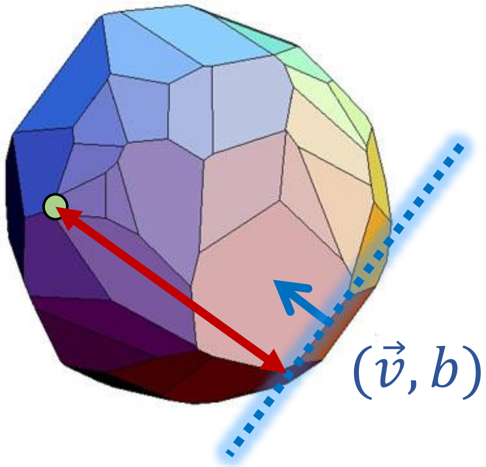
1	1	0	0	0	1	0	1	0	1	1	0

$x \in \{0,1\}^n$

simulation by a query protocol

$P \subseteq \mathbb{R}^n$

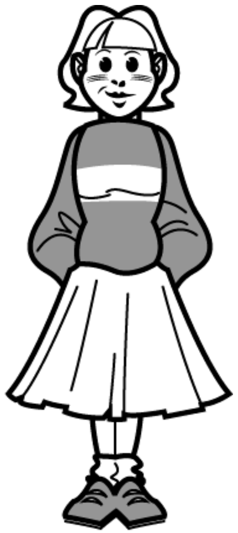
$x \in \{0,1\}^n$



$(\vec{v}, b) \in \mathbb{R}^{n+1}$

Strategy 2: Uhh... machine learning?

AlphaGo



$\boldsymbol{M}_{(\vec{v},b)}$

$$F\big((\vec{v}, b), x\big) = b - \langle \vec{v}, x \rangle$$

$(\vec{v}, b) \in \mathbb{R}^{n+1}$

Translator Bob

Converts Alice's m-bit message to an $\approx m$-bit query message

$\widetilde{\boldsymbol{M}}$

1	1	0	0	0	1	0	1	0	1	1	0

$x \in \{0,1\}^n$

$\approx$

Represent Bob as a QC state:

$$\Phi_{\text{Bob}} = \sum_{x \in \{0,1\}^n} |x\rangle\langle x| \otimes \rho_B^x$$

Key property:

$$\mathcal{D}(\Phi_{\text{Bob}} \,||\, \mathcal{U}) \leq O(m)$$

$\mathcal{D}$ is the relative von Neumann entropy,
$\mathcal{U}$ is the maximally mixed state

Recall: m is the # of qubits in Alice's message.

This holds when F (the function they are computing) is (mildly) reasonable.

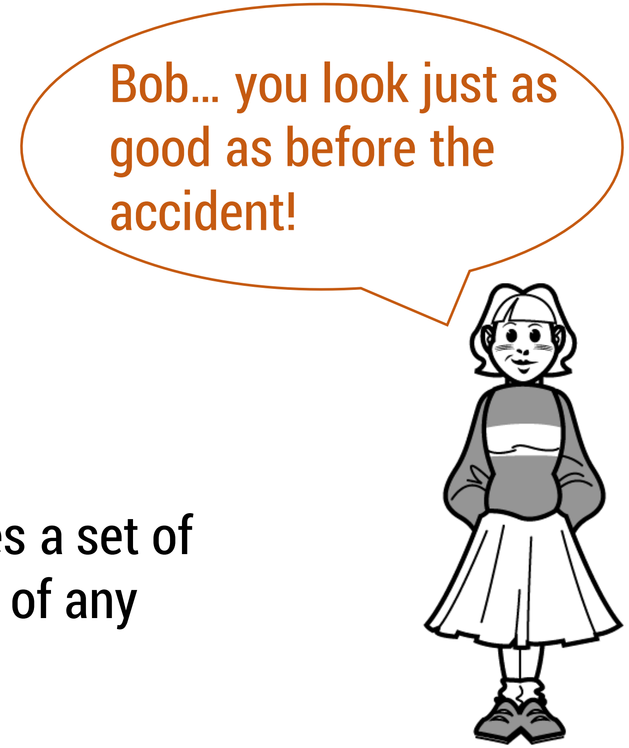
the approximation of Blob

$$\Phi_{\text{Bob}} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \otimes \rho_B^x$$

$$\mathcal{D}(\Phi_{\text{Bob}} \,\|\, \mathcal{U}) \leq O(m)$$

$\mathcal{D}$ is the relative von Neumann entropy,
$\mathcal{U}$ is the maximally mixed state

$\approx$



Notion of approximation:

It's enough that Alice cannot distinguish Bob from $\widetilde{\text{Bob}}$.

By the min-max theorem, Alice encodes a set of QC measurements that ensure validity of any potential Bob Φ:

$$\mathrm{Tr}(A_1\Phi) \approx \epsilon_1$$
$$\mathrm{Tr}(A_2\Phi) \approx \epsilon_2$$
$$\mathrm{Tr}(A_3\Phi) \approx \epsilon_3$$
$$\dots$$

$$A = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \sum_{(\vec{v},b)} \alpha_{(\vec{v},b)}(x)\ \rho_A^{(\vec{v},b)}$$

Jaynes' principle of maximum entropy

$$\Phi_{\text{Bob}} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \otimes \rho_B^x$$

$$\mathcal{D}(\Phi_{\text{Bob}} \,\|\, \mathcal{U}) \le O(m)$$

$\mathcal{D}$ is the relative von Neumann entropy,
$\mathcal{U}$ is the maximally mixed state

$\approx$

$$A = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \sum_{(\vec{v},b)} \alpha_{(\vec{v},b)}(x)\, \rho_A^{(\vec{v},b)}$$

[validity tests for potential Bob Φ]

Find the "simplest" Bob that passes all the tests:

Minimize $\mathcal{D}(\Phi \,\|\, \mathcal{U})$

subject to:

$$\mathrm{Tr}(A_1\Phi) \approx \epsilon_1$$
$$\mathrm{Tr}(A_2\Phi) \approx \epsilon_2$$
$$\mathrm{Tr}(A_3\Phi) \approx \epsilon_3$$
$$\ldots$$

$$\mathrm{Tr}(\Phi) = 1$$
$$\Phi \succcurlyeq 0$$

Jaynes' principle of maximum entropy

$$\Phi_{\text{Bob}} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \otimes \rho_B^x$$

$$\mathcal{D}(\Phi_{\text{Bob}} \,\|\, \mathcal{U}) \leq O(m)$$

$\mathcal{D}$ is the relative von Neumann entropy,
$\mathcal{U}$ is the maximally mixed state

$\approx$

$$A = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \sum_{(\vec{v},b)} \alpha_{(\vec{v},b)}(x)\, \rho_A^{(\vec{v},b)}$$

[validity tests for potential Bob Φ]

Find the "simplest" $\widetilde{\text{Bob}}$ that passes all the tests:

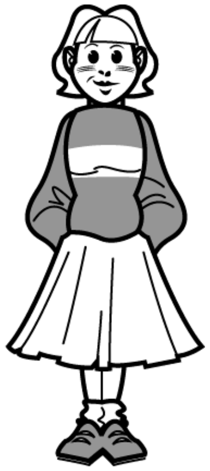
Minimize $\mathcal{D}(\Phi \,\|\, \mathcal{U})$

subject to:

$$\mathrm{Tr}(A_1\Phi) \approx \mathrm{Tr}(A_1\Phi_{\text{Bob}})$$
$$\mathrm{Tr}(A_2\Phi) \approx \mathrm{Tr}(A_2\Phi_{\text{Bob}})$$
$$\mathrm{Tr}(A_3\Phi) \approx \mathrm{Tr}(A_3\Phi_{\text{Bob}})$$
$$\dots$$

$$\mathrm{Tr}(\Phi) = 1$$
$$\Phi \succcurlyeq 0$$

… and just hope?



Jaynes' principle of maximum entropy

$$\Phi_{\text{Bob}} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \otimes \rho_B^x$$

$$\mathcal{D}(\Phi_{\text{Bob}} \,\|\, \mathcal{U}) \le O(m)$$

$\mathcal{D}$ is the relative von Neumann entropy,
$\mathcal{U}$ is the maximally mixed state

$\approx$

$$\tilde{A} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \sum_{(\vec{v},b)} \tilde{\alpha}_{(\vec{v},b)}(x)\, \rho_A^{(\vec{v},b)}$$

[validity tests for potential Bob Φ]

Find the "simplest" $\widetilde{\text{Bob}}$ that passes all the tests:

Minimize $\mathcal{D}(\Phi \,\|\, \mathcal{U})$

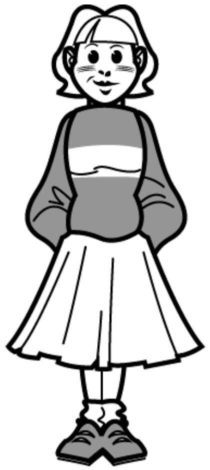
subject to:

$$\text{Tr}(\widetilde{A_1}\Phi) \approx \text{Tr}(\widetilde{A_1}\Phi_{\text{Bob}})$$
$$\text{Tr}(\widetilde{A_2}\Phi) \approx \text{Tr}(\widetilde{A_2}\Phi_{\text{Bob}})$$
$$\text{Tr}(\widetilde{A_3}\Phi) \approx \text{Tr}(\widetilde{A_3}\Phi_{\text{Bob}})$$
$$\cdots$$

$$\text{Tr}(\Phi) = 1$$
$$\Phi \succcurlyeq 0$$

… and just hope?

If we hope to learn a k-hapless Bob, then any such Bob is **orthogonal** to the Fourier expansion of $\alpha_{(\vec{v},b)}$ above degree k.



Jaynes' principle of maximum entropy

$$\Phi_{\text{Bob}} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \otimes \rho_B^x$$

$$\mathcal{D}(\Phi_{\text{Bob}} \,\|\, \mathcal{U}) \leq O(m)$$

$\mathcal{D}$ is the relative von Neumann entropy,
$\mathcal{U}$ is the maximally mixed state

$\approx$

$$\tilde{A} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \sum_{(\vec{v},b)} \tilde{\alpha}_{(\vec{v},b)}(x)\, \rho_A^{(\vec{v},b)}$$

[validity tests for potential Bob Φ]

Find the "simplest" $\widetilde{\text{Bob}}$ that passes all the tests:

Minimize $\mathcal{D}(\Phi \,\|\, \mathcal{U})$

subject to:

$$\mathrm{Tr}(\widetilde{A_1}\Phi) \approx \mathrm{Tr}(\widetilde{A_1}\Phi_{\text{Bob}})$$
$$\mathrm{Tr}(\widetilde{A_2}\Phi) \approx \mathrm{Tr}(\widetilde{A_2}\Phi_{\text{Bob}})$$
$$\mathrm{Tr}(\widetilde{A_3}\Phi) \approx \mathrm{Tr}(\widetilde{A_3}\Phi_{\text{Bob}})$$
$$\ldots$$

$$\mathrm{Tr}(\Phi) = 1$$
$$\Phi \succcurlyeq 0$$

Optimal solution:

$$\Phi^* \propto \exp\left(\sum_{j\geq 1} \lambda_j \widetilde{A_j}\right)$$

$$= \exp\left(\sum_{j\geq 1} \lambda_j \frac{\widetilde{A_j}}{2}\right)^2$$

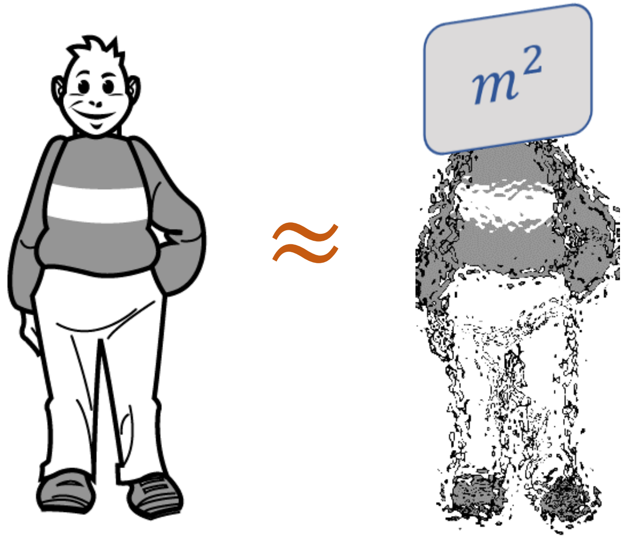
$$\approx \mathrm{poly}\left(\sum_{j\geq 1} \lambda_j \widetilde{A_j}\right)^2$$

Jaynes' principle of maximum entropy

$$\Phi_{\text{Bob}} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \otimes \rho_B^x$$

$$\mathcal{D}(\Phi_{\text{Bob}} \,\|\, \mathcal{U}) \le O(m)$$

$\mathcal{D}$ is the relative von Neumann entropy,
$\mathcal{U}$ is the maximally mixed state



$$\tilde{A} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \sum_{(\vec{v},b)} \tilde{\alpha}_{(\vec{v},b)}(x)\, \rho_A^{(\vec{v},b)}$$

[validity tests for potential Bob Φ]

Optimal solution:

$$\Phi^* \propto \exp\left(\sum_{j\ge 1} \lambda_j \widetilde{A_j}\right)$$

$$= \exp\left(\sum_{j\ge 1} \lambda_j \frac{\widetilde{A_j}}{2}\right)^2$$

$$\approx \text{poly}\left(\sum_{j\ge 1} \lambda_j \widetilde{A_j}\right)^2$$

Find the "simplest" $\widetilde{\text{Bob}}$ that passes all the tests:

Minimize $\mathcal{D}(\Phi \,\|\, \mathcal{U})$

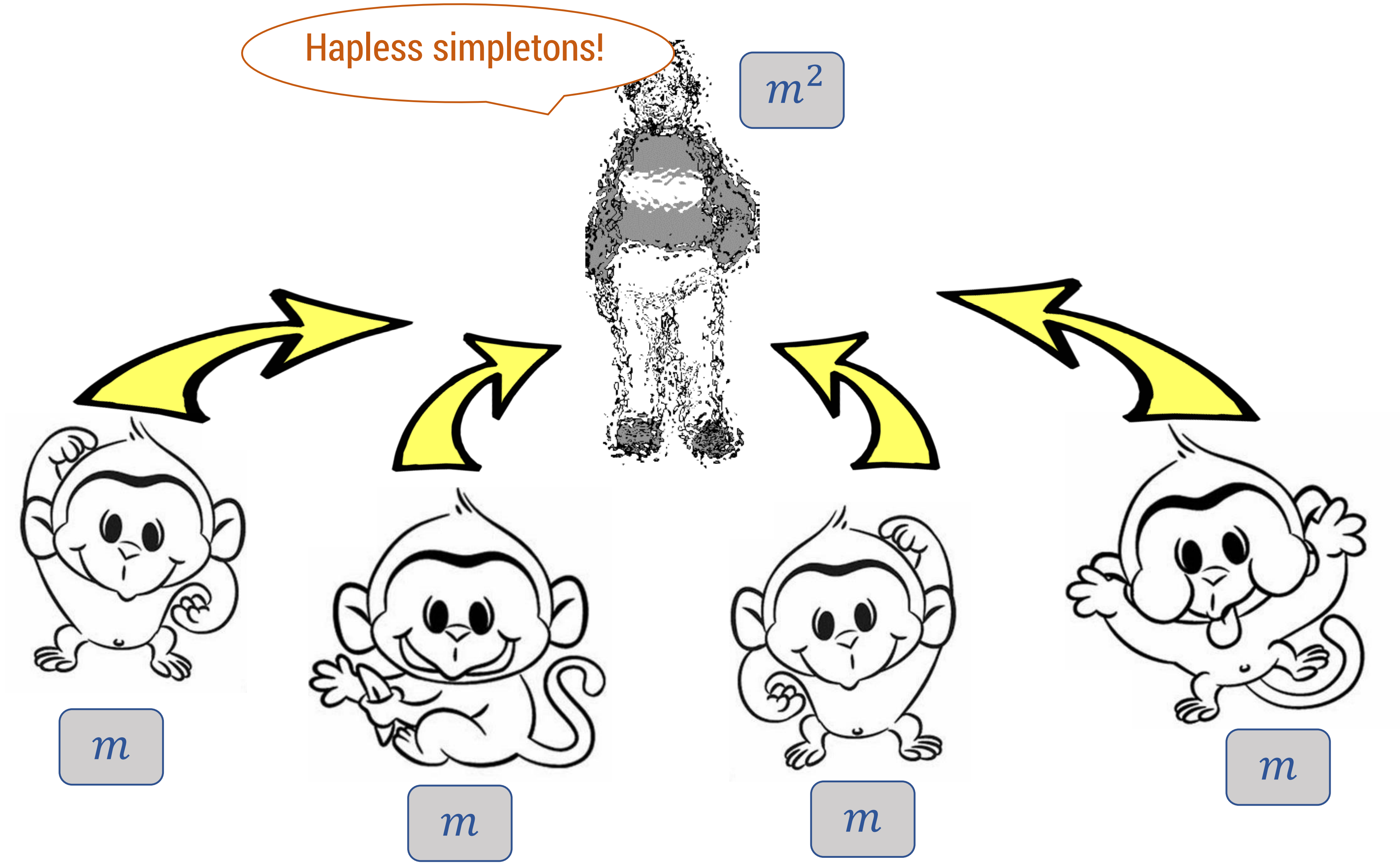
Approximation is an $O(m^2)$-hapless Bob.

subject to:

$$\mathrm{Tr}(\widetilde{A_1}\Phi) \approx \mathrm{Tr}(\widetilde{A_1}\Phi_{\text{Bob}})$$
$$\mathrm{Tr}(\widetilde{A_2}\Phi) \approx \mathrm{Tr}(\widetilde{A_2}\Phi_{\text{Bob}})$$
$$\mathrm{Tr}(\widetilde{A_3}\Phi) \approx \mathrm{Tr}(\widetilde{A_3}\Phi_{\text{Bob}})$$
$$\dots$$

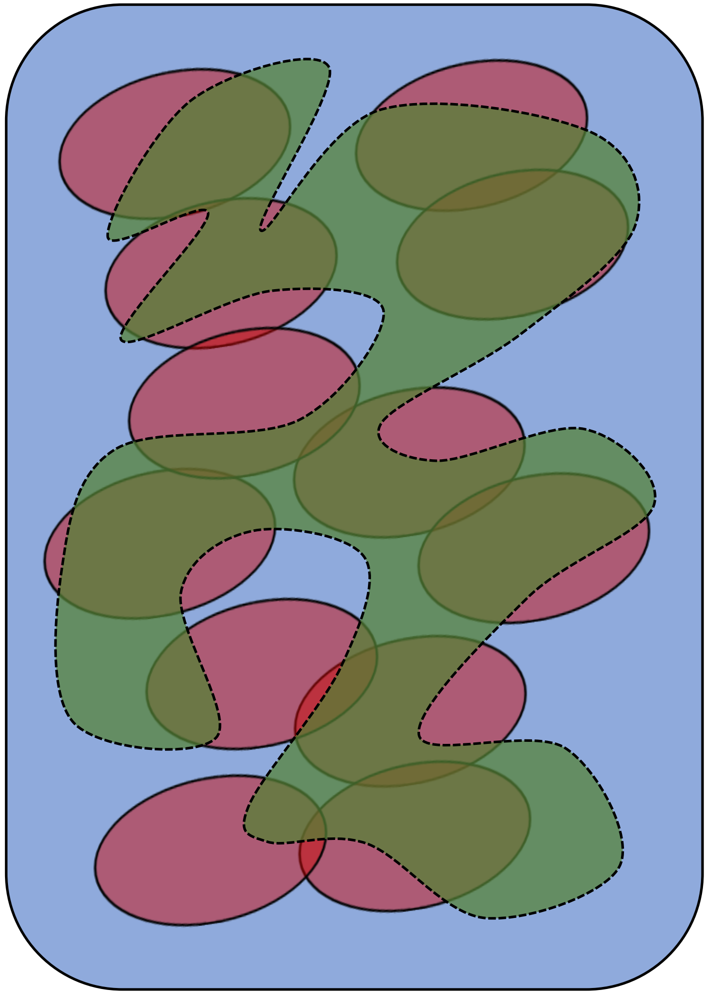
$$\mathrm{Tr}(\Phi) = 1$$
$$\Phi \succcurlyeq 0$$

model is more complex than the training data



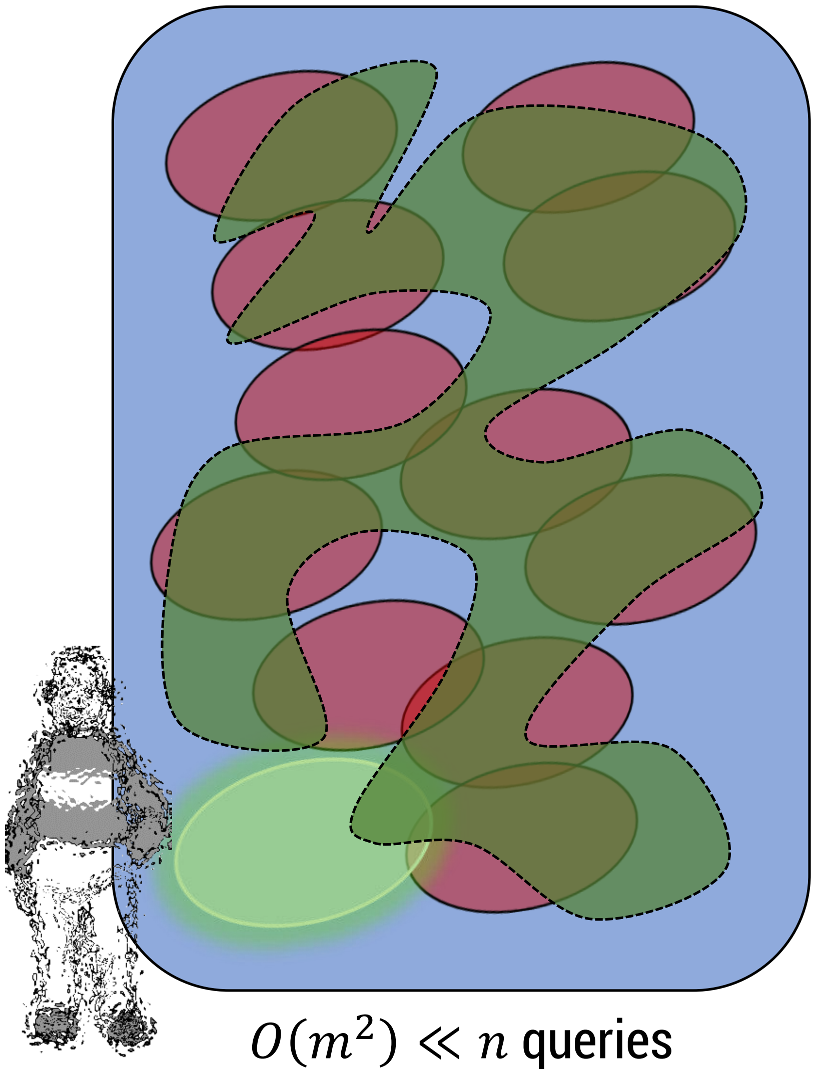
the use of symmetry

n variables



$O(m^2) \ll n$ queries

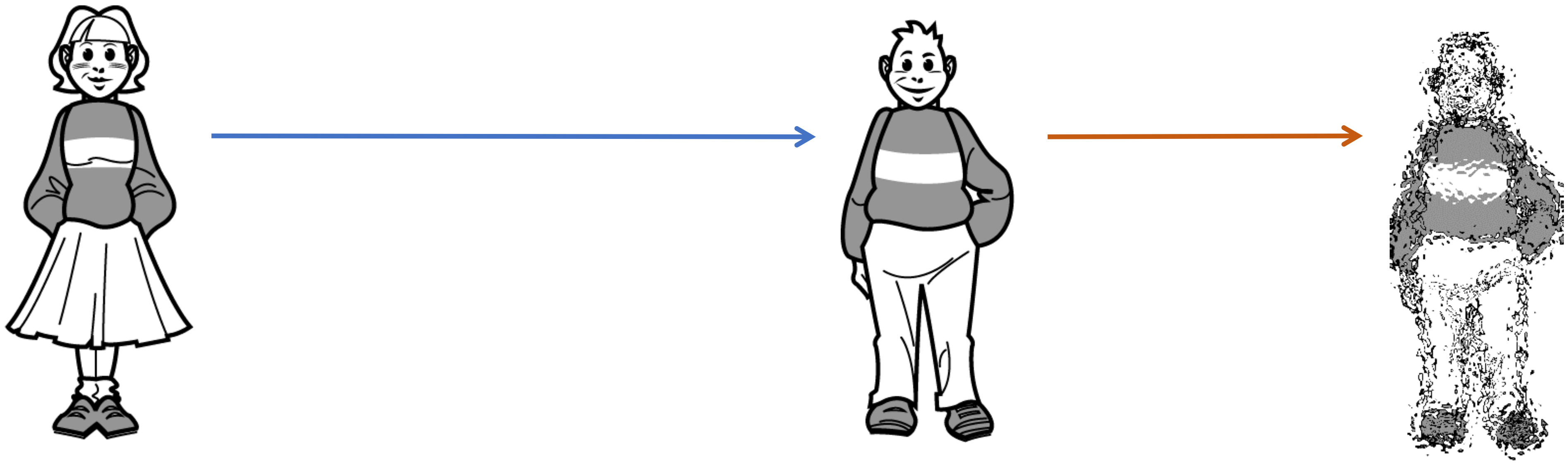
n variables



$O(m^2) \ll n$ queries

Application from yesterday:
"Small SDPs are bad at recognizing separable states"
[Harrow, Natarajan, Wu 2016]

Open problem: Are there small SDP lifts of the perfect matching polytope?
Exponential lower bounds for approximating CSPs?



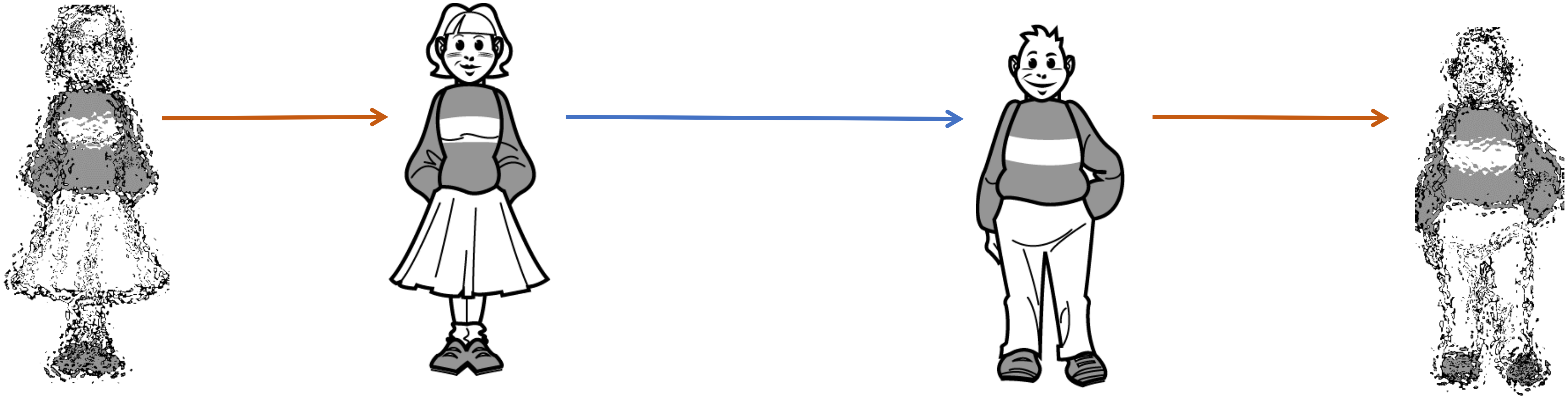
into the future

Application from yesterday:

"Small SDPs are bad at recognizing separable states"

[Harrow, Natarajan, Wu 2016]

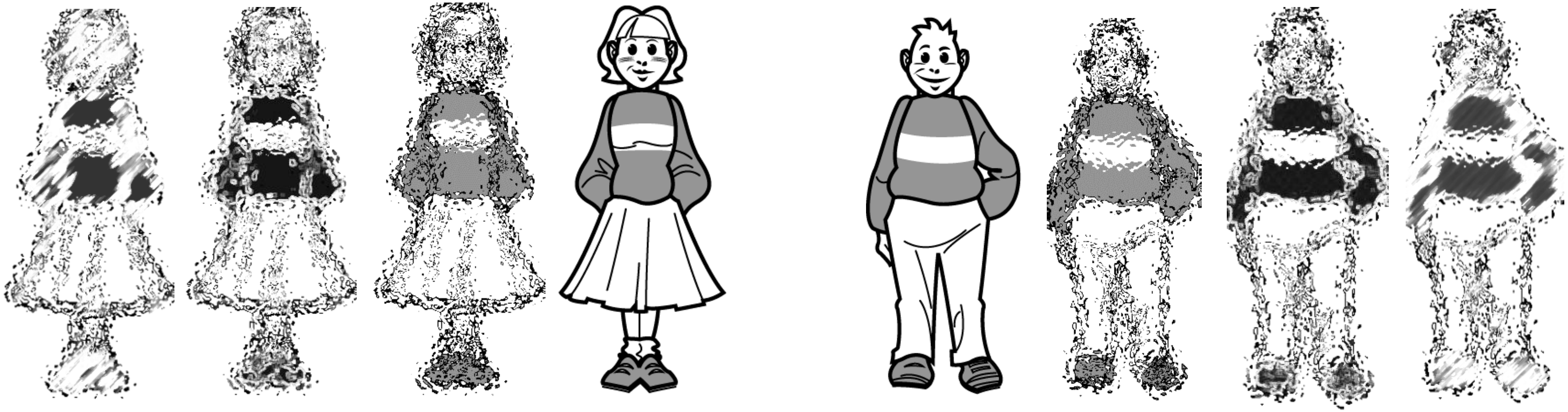
Open problem: Are there small SDP lifts of the perfect matching polytope?
Exponential lower bounds for approximating CSPs?



[Kothari-Meka-Raghavendra 2017] do this in the LP setting.

Application from yesterday:
"Small SDPs are bad at recognizing separable states"
[Harrow, Natarajan, Wu 2016]

Open problem: Are there small SDP lifts of the perfect matching polytope?
Exponential lower bounds for approximating CSPs?



Open problem: Are there small SDP lifts of the perfect matching polytope?
Exponential lower bounds for approximating CSPs?

Log rank conjecture:

[FMPTW'12] observed that for **Boolean** communication problems, this is equivalent to classical-quantum simulation with polynomial overhead.