# Lower bounds for CSP refutation by SDP hierarchies

Ryuhei Mori[1]    David Witmer[2]

[1]Tokyo Institute of Technology

[2]Carnegie Mellon University

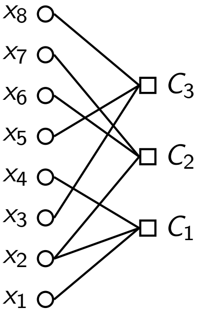
APPROX + RANDOM at Paris
Sep 9, 2016

# Constraints satisfaction problem (CSP)

*k*-CSP

- An alphabet set $[q] = \{1, 2, \dots, q\}$.
- Variables $(x_i \in [q])_{i=1,\dots,n}$.
- Constraints $(C_j : [q]^k \to \{0,1\})_{j=1,\dots m}$.

$$F = C_1(x_1, x_2, x_4) \wedge C_2(x_2, x_6, x_7) \wedge C_3(x_3, x_5, x_8)$$



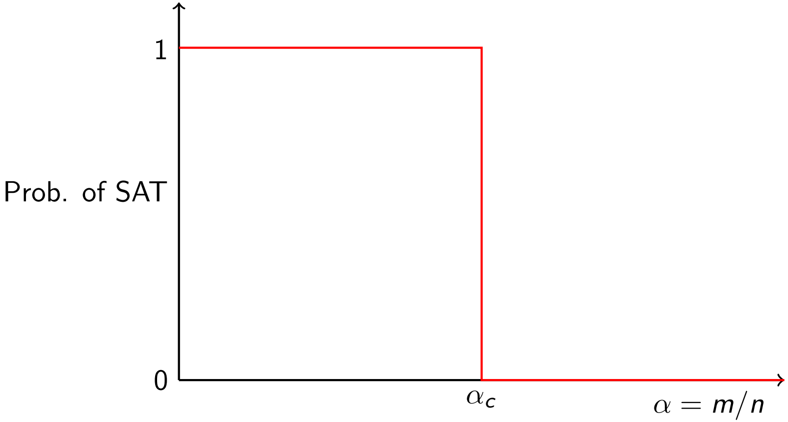
# Random $k$-CSP

Random CSP with parameters $n$, $m$ and $C\colon [q]^k \to \{0,1\}$.

Each constraint is generated independently in the following way.

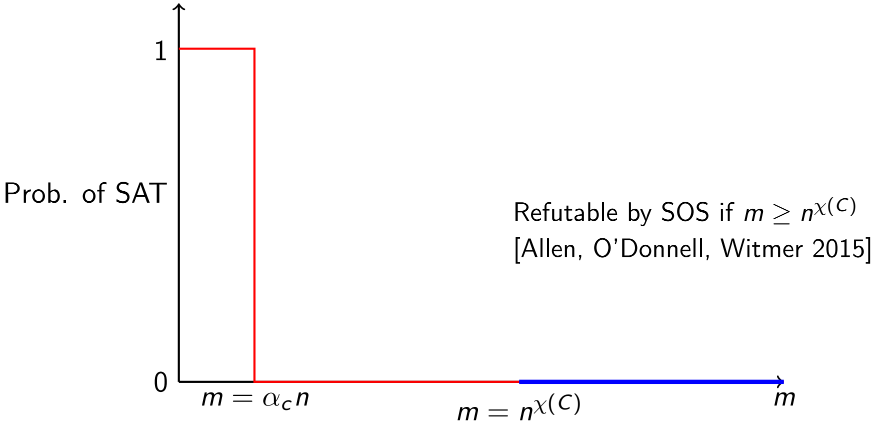
- Chose a set of $k$ variables from $\binom{n}{k}$ candidates uniformly.
- Apply random permutations on $[q]$ for all of the $k$ variables.

# Phase transition of the random $k$-CSP




Phase transition conjecture

# Refutation of random $k$-CSP




Main result:
Not refutable in poly-time by SA+ nor LS+ if $m < n^{\chi(C)-\delta}$

# Table of contents

# MAX-$k$-CSP

$$\max_{\mathbf{x} \in [q]^n} : \sum_{j=1}^{m} C_j(\mathbf{x}_j)$$

if $m >$ Optimum of relaxed MAX-$k$-CSP then

$m >$ Optimum of MAX-$k$-CSP $\implies$ The CSP is UNSAT

# LP/SDP relexation

Formulation of combinatorial optimization problem as linear program.

$$\max_{\mathbf{x}} : \sum_{j=1}^{m} C_j(\mathbf{x}_a)$$

$$\text{s.t.} : \mathbf{x} \in [q]^n$$

$$\Updownarrow$$

$$\max_{p} : \mathbb{E}_p \left[ \sum_{j=1}^{m} C_j(\mathbf{X}_a) \right]$$

$$\text{s.t.} : p \in \mathcal{P}([q]^n) = \text{ The set of distributions on } [q]^n$$

Then, relax the polytope $\mathcal{P}([q]^n)$.

# Sherali-Adams LP hierarchy

The tight polytope

$$\max_p : \mathbb{E}_p \left[ \sum_{j=1}^{m} C_j(\mathbf{X}_a) \right]$$

$$\text{s.t. } : p \in \mathcal{P}([q]^n) = \text{ The set of distributions on } [q]^n$$

The $r$-round Sherali-Adams relaxation

$$\max_{(p_S : S \subseteq [n], |S| \leq r)} : \sum_{j=1}^{m} \mathbb{E}_{p_{V(C_j)}} [C_j(\mathbf{x}_j)]$$

$$\text{s.t. } : p_S \in \mathcal{P}([q]^k), S \subseteq [n], |S| \leq r$$

All local distributions are locally-consistent

$$\text{Tight} = \text{SA}(n) \subseteq \text{SA}(n-1) \subseteq \cdots \subseteq \text{SA}(k)$$

# Sherali-Adams+ SDP hierarchy

The $r$-round SA+ = The $r$-round SA with PSDness condition for $\Sigma$

where $\Sigma$ is variance-covariance matrix, i.e.,

$$\Sigma_{(i,x),(j,y)} := p_{\{i,j\}}(x,y) - p_{\{i\}}(x)p_{\{j\}}(y).$$

$$\begin{aligned} \max_{(p_S : S \subseteq [n], |S| \le r)} : \quad & \sum_{j=1}^{m} \mathbb{E}_{p_{V(C_j)}}[C_j(\mathbf{x}_j)] \\ \text{s.t.} : \quad & (p_S)_{S \subseteq [n], |S| \le r} \in \mathrm{SA}(r) \\ & \Sigma \succeq 0 \end{aligned}$$

# Equivalence of PSDness

## Lemma (Schur complement)

$$\begin{bmatrix} 1 & p^T \\ p & B \end{bmatrix} \textit{ is PSD} \iff B - pp^T \textit{ is PSD.}$$

## Proof.

$$\begin{aligned}
&\begin{bmatrix} 1 & p^T \\ p & B \end{bmatrix} \text{ is PSD} \\
\iff & \left( [x_0\, x] \begin{bmatrix} 1 & p^T \\ p & B \end{bmatrix} [x_0\, x]^T \geq 0 \text{ for any } x_0 \in \mathbb{R},\ x \in \mathbb{R}^{nq} \right) \\
\iff & \left( x_0^2 + 2\langle p, x\rangle x_0 + \langle Bx, x\rangle \geq 0 \text{ for any } x_0 \in \mathbb{R},\ x \in \mathbb{R}^{nq} \right) \\
\iff & \left( (x_0 + \langle p, x\rangle)^2 - \langle p, x\rangle^2 + \langle Bx, x\rangle \geq 0 \text{ for any } x_0 \in \mathbb{R},\ x \in \mathbb{R}^{nq} \right) \\
\iff & \left( -\langle p, x\rangle^2 + \langle Bx, x\rangle \geq 0 \text{ for any } x \in \mathbb{R}^{nq} \right) \\
\iff & \left( x(B - pp^T)x^T \text{ for any } x \in \mathbb{R}^{nq} \right) \\
\iff & B - pp^T \text{ is PSD.}
\end{aligned}$$

# Traditional description of SA+

$$\max_{(p_S : S \subseteq n, |S| \leq r)} : \quad \sum_{j=1}^{m} \mathbb{E}_{p_j}\left[C_j(\mathbf{x}_j)\right]$$

$$\text{s.t.} : \quad (p_S)_{S \subseteq [n], |S| \leq r} \in \mathrm{SA}(r)$$

$$\langle \mathbf{v}_{i,a}, \mathbf{v}_{j,b} \rangle = p_{\{i,j\}}(a,b) \quad \forall i \neq j \in [n], a, b \in [q]$$

$$\langle \mathbf{v}_{i,a}, \mathbf{v}_{i,b} \rangle = 0 \quad \forall i \in [n], a \neq b \in [q]$$

$$\|\mathbf{v}_{i,a}\|^2 = \langle \mathbf{v}_{i,a}, \mathbf{v}_{\emptyset} \rangle = p_{\{i\}}(a) \quad \forall i \in [n], a \in [q]$$

$$\|\mathbf{v}_{\emptyset}\|^2 = 1$$

$$\exists V, \text{s.t. } V^T V = \begin{bmatrix} 1 & p^T \\ p & B \end{bmatrix} \iff \begin{bmatrix} 1 & p^T \\ p & B \end{bmatrix} \text{ is PSD}$$

# Other SDP hierarchies

- Lovász-Schrijver+ (LS+): SA with PSDness condition for conditional variance-covariance matrix [Tulsiani and Worah 2012]
- Lasserre/SOS: SA with PSDness condition for variance-covariance matrix for higher order statistics.

# LP/SDP in Computational Complexity

- If poly-size SA cannot refute then any poly-size LP relaxation cannot refute [Chan, Lee, Raghavendra, and Steurer 2013]
- If poly-size SOS cannot refute then any poly-size SDP relaxation cannot refute [Lee, Raghavendra, and Steurer 2015]

# Known facts and Main results

- Worst-case lower bound of SA+ for pairwise uniform MAX-$k$-CSP with linearly many constraints [Benabbas, Georgiou, Magen, and Tulsiani 2012]
- Average-case lower bound of LS+ for pairwise uniform random MAX-$k$-CSP with linearly many constraints [Tulsiani and Worah 2013]
- Average-case lower bound of SA+ for $(t-1)$-wise uniform random MAX-$k$-CSP with $n^{t/2-\delta}$ constraints with small modification [O'Donnell and Witmer 2014]

**Theorem ([This Work])**

*Poly-size SA+/LS+ cannot refute random $(t-1)$-wise uniform CSP with $n^{t/2-\delta}$ constraints with high probability.*

# Summary for LP/SDP hierarchy

All LP/SDP hierarchies can be understood by the idea of local distributions.

All SDP hierarchies have PSDness condition of variance-covariance matrix.

# Table of contents

# The local distribution

## Definition ($(t-1)$-wise uniform constraint)

A constraint $C\colon [q]^k \to \{0,1\}$ is said to be $(t-1)$-wise uniform if there exists a distribution $\mu$ on the support of $C$ such that the marginal distribution $\mu_T$ for $T \subseteq [k]$ is uniform distribution on $[q]^{|T|}$ if $|T| \leq t-1$.

Specific choice of local distribution [BGMT 2012].

$$D'_S(x_S) := \frac{1}{Z_S} \prod_{C \in \mathcal{C}(S)} \mu_C(x_{\Gamma C})$$

$$D_S(x_S) := \sum_{x_{\bar{S} \setminus S}} D'_{\bar{S}}(x_{\bar{S}})$$

$\bar{S}$: The closure of $S$.

# Local consistency

Theorem ([BGMT 2012] [O'Donnell and Witmer 2014])
*For* $m = O(n^{t/2-\delta})$ *and* $r = O(n^{\frac{\delta}{t-2}})$, $(D_S)_{|S|\leq r}$ *is locally consistent with high probability.*

# Summary and our goal

Sherali-Adams+ Relaxation.

$$\Sigma_{(i,x),(j,y)} := p_{\{i,j\}}(x,y) - p_{\{i\}}(x)p_{\{j\}}(y).$$

$$\begin{aligned} \max_{(p_S : S\subseteq[n], |S|\le r)} \quad : \quad & \sum_{j=1}^{m} \mathbb{E}_{p_{V(C_j)}}[C_j(\mathbf{x}_j)] \\ \text{s.t.} \quad : \quad & (p_S)_{S\subseteq[n],|S|\le r} \in \mathrm{SA}(r) \\ & \Sigma \succeq 0 \end{aligned}$$

- From [BGMT 2012] and [O'Donnell and Witmer 2014], $(p_S = D_S)_{S\subseteq[n],|S|\le r} \in \mathrm{SA}(r)$.
- Our goal is to show that $\Sigma$ for $(D_S)_{S\subseteq[n],|S|\le r}$ is PSD.

# Table of contents

- LP/SDP hierarchies
- locally consistent distribution [BGMT 2012]
- Sketch of the proof

# Proof strategy

Goal is to show

$\Sigma$ for $(D_S)_{S \subseteq [n], |S| \leq r}$ is PSD.

- In fact, $\Sigma$ is not diagonal.
- We will not find the vectors explicitly.
- We will show the PSDness in more implicit way.

# Correlation graph

- Vertex set: $V = [n]$.
- Edge set: $E = \{(i, j) \mid \exists (x, y) \in [q]^2 \text{ s.t. } \Sigma_{(i,x),(j,y)} \neq 0\}$.

**Theorem ([This Work])**

*Every connected components of the correlation graph for $D$ has size $O(1)$.*

By rearranging the indices of $\Sigma$, we obtain

$$\begin{bmatrix} \Sigma_1 & 0 & 0 & 0 \\ 0 & \Sigma_2 & 0 & 0 \\ 0 & 0 & \Sigma_3 & 0 \\ 0 & 0 & 0 & I \end{bmatrix}$$

Since each block is PSD, the whole matrix $\Sigma$ is PSD.

# Summary

Main result:

Poly-size SA+/LS+ cannot refute random $(t-1)$-wise uniform CSP with $n^{t/2-\delta}$ constraints with high probability.

Proof technique:

- We regard the PSDness condition as $\Sigma \succeq 0$.
- We show that the connected components in the correlation graph are small.
- This immediately implies the PSDness of $\Sigma$.