Introduction to LP and SDP Hierarchies

Madhur Tulsiani

Princeton University

Local Constraints in Approximation Algorithms

- Linear Programming (LP) or Semidefinite Programming (SDP) based approximation algorithms impose constraints on few variables at a time.

Local Constraints in Approximation Algorithms

- Linear Programming (LP) or Semidefinite Programming (SDP) based approximation algorithms impose constraints on few variables at a time.

- When can local constraints help in approximating a global property (eg. Vertex Cover, Chromatic Number)?

Local Constraints in Approximation Algorithms

- Linear Programming (LP) or Semidefinite Programming (SDP) based approximation algorithms impose constraints on few variables at a time.
- When can local constraints help in approximating a global property (eg. Vertex Cover, Chromatic Number)?
- How does one reason about increasingly larger local constraints?

Local Constraints in Approximation Algorithms

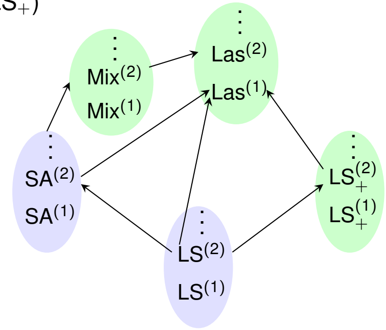
- Linear Programming (LP) or Semidefinite Programming (SDP) based approximation algorithms impose constraints on few variables at a time.
- When can local constraints help in approximating a global property (eg. Vertex Cover, Chromatic Number)?
- How does one reason about increasingly larger local constraints?
- Does approximation get better as constraints get larger?

LP/SDP Hierarchies

- Various hierarchies give increasingly powerful programs at different levels (rounds).
  - Lovász-Schrijver (LS, LS_+)
  - Sherali-Adams
  - Lasserre
  - “Mixed”

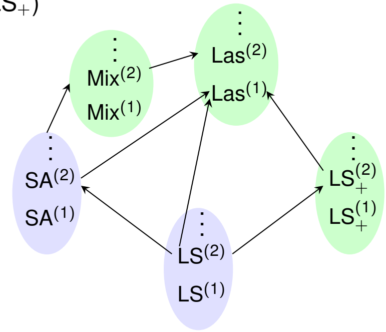
LP/SDP Hierarchies

- Various hierarchies give increasingly powerful programs at different levels (rounds).
  - Lovász-Schrijver (LS, LS_+)
  - Sherali-Adams
  - Lasserre
  - "Mixed"



LP/SDP Hierarchies

- Various hierarchies give increasingly powerful programs at different levels (rounds).
  - Lovász-Schrijver (LS, LS_+)
  - Sherali-Adams
  - Lasserre
  - "Mixed"



- Can optimize over r^{th} level in time $n^{O(r)}$. n^{th} level is tight.

LP/SDP Hierarchies

- Powerful computational model capturing most known LP/SDP algorithms within constant number of levels.

LP/SDP Hierarchies

- Powerful computational model capturing most known LP/SDP algorithms within constant number of levels.

- Lower bounds rule out large and natural class of algorithms.

LP/SDP Hierarchies

- Powerful computational model capturing most known LP/SDP algorithms within constant number of levels.
- Lower bounds rule out large and natural class of algorithms.
- Performance measured by considering integrality gap at various levels.

$$\text{Integrality Gap} = \frac{\text{Optimum of Relaxation}}{\text{Integer Optimum}} \qquad \text{(for maximization)}$$

Why bother?

- Conditional
- All polytime algorithms

NP-Hardness

UG-Hardness

- Unconditional
- Restricted class of algorithms

LP/SDP Hierarchies

What Hierarchies want

Example: Maximum Independent Set for graph $G = (V, E)$

$$\begin{aligned} \text{minimize} \quad & \sum_u x_u \\ \text{subject to} \quad & x_u + x_v \leq 1 \qquad \forall (u, v) \in E \\ & x_u \in [0, 1] \end{aligned}$$

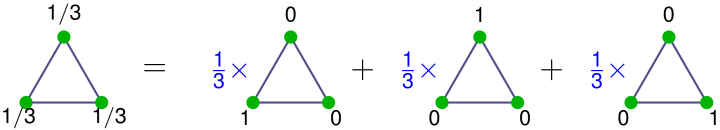
- Hope: $x_1, \ldots, x_n$ is convex combination of 0/1 solutions.

What Hierarchies want

Example: Maximum Independent Set for graph $G = (V, E)$

$$\begin{aligned} \text{minimize} \quad & \sum_u x_u \\ \text{subject to} \quad & x_u + x_v \leq 1 \qquad \forall\, (u, v) \in E \\ & x_u \in [0, 1] \end{aligned}$$

- Hope: $x_1, \ldots, x_n$ is convex combination of 0/1 solutions.

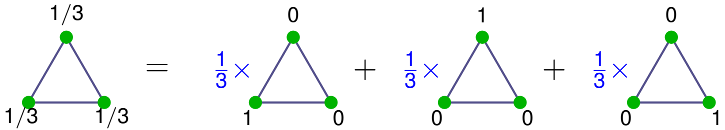


What Hierarchies want

Example: Maximum Independent Set for graph $G = (V, E)$

$$
\begin{aligned}
\text{minimize} \quad & \sum_u x_u & \\
\text{subject to} \quad & x_u + x_v \leq 1 & \forall\, (u, v) \in E \\
& x_u \in [0, 1] &
\end{aligned}
$$

- Hope: $x_1, \ldots, x_n$ is marginal of distribution over 0/1 solutions.

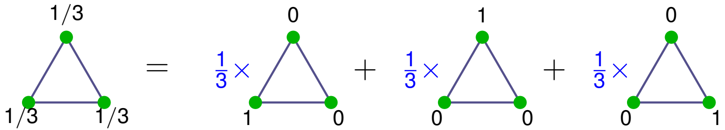


What Hierarchies want

Example: Maximum Independent Set for graph $G = (V, E)$

$$\begin{aligned} \text{minimize} \quad & \sum_u x_u \\ \text{subject to} \quad & x_u + x_v \leq 1 \qquad \forall\, (u, v) \in E \\ & x_u \in [0, 1] \end{aligned}$$

- Hope: $x_1, \ldots, x_n$ is marginal of distribution over 0/1 solutions.



- Hierarchies add variables for conditional/joint probabilities.

The Sherali-Adams Hierarchy

The Sherali-Adams Hierarchy

- Start with a 0/1 integer linear program.
- Hope: Fractional $(x_1, \ldots, x_n) = \mathbb{E}\left[(z_1, \ldots, z_n)\right]$ for integral $(z_1, \ldots, z_n)$

The Sherali-Adams Hierarchy

- Start with a 0/1 integer linear program.
- Hope: Fractional $(x_1, \ldots, x_n) = \mathbb{E}\left[(z_1, \ldots, z_n)\right]$ for integral $(z_1, \ldots, z_n)$
- Add "big variables" X_S for $|S| \leq r$
(think $X_S = \mathbb{E}\left[\prod_{i \in S} z_i\right] = \mathbb{P}\left[\text{All vars in } S \text{ are } 1\right]$)

The Sherali-Adams Hierarchy

- Start with a 0/1 integer linear program.
- Hope: Fractional $(x_1, \ldots, x_n) = \mathbb{E}[(z_1, \ldots, z_n)]$ for integral $(z_1, \ldots, z_n)$
- Add “big variables” X_S for $|S| \leq r$
(think $X_S = \mathbb{E}\left[\prod_{i \in S} z_i\right] = \mathbb{P}$ [All vars in S are 1])
- Constraints:

The Sherali-Adams Hierarchy

- Start with a 0/1 integer linear program.
- Hope: Fractional $(x_1, \ldots, x_n) = \mathbb{E}\left[(z_1, \ldots, z_n)\right]$ for integral $(z_1, \ldots, z_n)$
- Add "big variables" X_S for $|S| \leq r$
  (think $X_S = \mathbb{E}\left[\prod_{i \in S} z_i\right] = \mathbb{P}$[All vars in S are 1])
- Constraints:

$$\sum_i a_i z_i \leq b$$

$$\mathbb{E}\left[\left(\sum_i a_i z_i\right) \cdot z_5 z_7 (1 - z_9)\right] \leq \mathbb{E}\left[b \cdot z_5 z_7 (1 - z_9)\right]$$

The Sherali-Adams Hierarchy

- Start with a 0/1 integer linear program.
- Hope: Fractional $(x_1, \ldots, x_n) = \mathbb{E}[(z_1, \ldots, z_n)]$ for integral $(z_1, \ldots, z_n)$
- Add "big variables" X_S for $|S| \leq r$
  (think $X_S = \mathbb{E}\left[\prod_{i \in S} z_i\right] = \mathbb{P}[\text{All vars in } S \text{ are } 1]$)
- Constraints:

$$\begin{aligned}
\sum_i a_i z_i &\leq b \\
\mathbb{E}\left[\left(\sum_i a_i z_i\right) \cdot z_5 z_7 (1 - z_9)\right] &\leq \mathbb{E}[b \cdot z_5 z_7 (1 - z_9)] \\
\sum_i a_i \cdot (X_{\{i,5,7\}} - X_{\{i,5,7,9\}}) &\leq b \cdot (X_{\{5,7\}} - X_{\{5,7,9\}})
\end{aligned}$$

- LP on n^r variables.

Sherali-Adams $\approx$ Locally Consistent Distributions

Sherali-Adams ≈ Locally Consistent Distributions

- Using $0 \leq z_1 \leq 1$, $0 \leq z_2 \leq 1$

$$
\begin{array}{rcccl}
0 & \leq & X_{\{1,2\}} & \leq & 1 \\
0 & \leq & X_{\{1\}} - X_{\{1,2\}} & \leq & 1 \\
0 & \leq & X_{\{2\}} - X_{\{1,2\}} & \leq & 1 \\
0 & \leq & 1 - X_{\{1\}} - X_{\{2\}} + X_{\{1,2\}} & \leq & 1
\end{array}
$$

Sherali-Adams ≈ Locally Consistent Distributions

- Using $0 \le z_1 \le 1$, $0 \le z_2 \le 1$

$$\begin{array}{rcccl} 0 & \le & X_{\{1,2\}} & \le & 1 \\ 0 & \le & X_{\{1\}} - X_{\{1,2\}} & \le & 1 \\ 0 & \le & X_{\{2\}} - X_{\{1,2\}} & \le & 1 \\ 0 & \le & 1 - X_{\{1\}} - X_{\{2\}} + X_{\{1,2\}} & \le & 1 \end{array}$$

- $X_{\{1\}}, X_{\{2\}}, X_{\{1,2\}}$ define a distribution $D(\{1,2\})$ over $\{0,1\}^2$.

Sherali-Adams ≈ Locally Consistent Distributions

- Using $0 \leq z_1 \leq 1$, $0 \leq z_2 \leq 1$

$$\begin{array}{ccccc} 0 & \leq & X_{\{1,2\}} & \leq & 1 \\ 0 & \leq & X_{\{1\}} - X_{\{1,2\}} & \leq & 1 \\ 0 & \leq & X_{\{2\}} - X_{\{1,2\}} & \leq & 1 \\ 0 & \leq & 1 - X_{\{1\}} - X_{\{2\}} + X_{\{1,2\}} & \leq & 1 \end{array}$$

- $X_{\{1\}}, X_{\{2\}}, X_{\{1,2\}}$ define a distribution $D(\{1,2\})$ over $\{0,1\}^2$.
- $D(\{1,2,3\})$ and $D(\{1,2,4\})$ must agree with $D(\{1,2\})$.

Sherali-Adams $\approx$ Locally Consistent Distributions

- Using $0 \leq z_1 \leq 1$, $0 \leq z_2 \leq 1$

$$
\begin{array}{rcccl}
0 & \leq & X_{\{1,2\}} & \leq & 1 \\
0 & \leq & X_{\{1\}} - X_{\{1,2\}} & \leq & 1 \\
0 & \leq & X_{\{2\}} - X_{\{1,2\}} & \leq & 1 \\
0 & \leq & 1 - X_{\{1\}} - X_{\{2\}} + X_{\{1,2\}} & \leq & 1
\end{array}
$$

- $X_{\{1\}}, X_{\{2\}}, X_{\{1,2\}}$ define a distribution $D(\{1,2\})$ over $\{0,1\}^2$.

- $D(\{1,2,3\})$ and $D(\{1,2,4\})$ must agree with $D(\{1,2\})$.

- $SA^{(r)} \implies LCD^{(r)}$. If each constraint has at most k vars, $LCD^{(r+k)} \implies SA^{(r)}$

The Lasserre Hierarchy

The Lasserre Hierarchy

- Start with a 0/1 integer quadratic program.

The Lasserre Hierarchy

- Start with a 0/1 integer quadratic program.
- Think "big" variables $Z_S = \prod_{i \in S} z_i$.

The Lasserre Hierarchy

- Start with a 0/1 integer quadratic program.
- Think “big" variables $Z_S = \prod_{i \in S} z_i$.
- Associated psd matrix Y (moment matrix)

$$Y_{S_1,S_2} = \mathbb{E}\left[Z_{S_1} \cdot Z_{S_2}\right] = \mathbb{E}\left[\prod_{i \in S_1 \cup S_2} z_i\right]$$

The Lasserre Hierarchy

- Start with a 0/1 integer quadratic program.
- Think “big" variables $Z_S = \prod_{i \in S} z_i$.
- Associated psd matrix Y (moment matrix)

$$Y_{S_1, S_2} = \mathbb{E}\left[Z_{S_1} \cdot Z_{S_2}\right] = \mathbb{E}\left[\prod_{i \in S_1 \cup S_2} z_i\right]$$

The Lasserre Hierarchy

- Start with a 0/1 integer quadratic program.
- Think “big" variables $Z_S = \prod_{i \in S} z_i$.
- Associated psd matrix Y (moment matrix)

$$Y_{S_1, S_2} = \mathbb{E}\left[Z_{S_1} \cdot Z_{S_2}\right] = \mathbb{E}\left[\prod_{i \in S_1 \cup S_2} z_i\right] = \mathbb{P}[\text{All vars in } S_1 \cup S_2 \text{ are } 1]$$

The Lasserre Hierarchy

- Start with a 0/1 integer quadratic program.
- Think “big" variables $Z_S = \prod_{i \in S} z_i$.
- Associated psd matrix Y (moment matrix)

$$Y_{S_1,S_2} = \mathbb{E}\left[Z_{S_1} \cdot Z_{S_2}\right] = \mathbb{E}\left[\prod_{i \in S_1 \cup S_2} z_i\right] = \mathbb{P}[\text{All vars in } S_1 \cup S_2 \text{ are } 1]$$

- $(Y \succeq 0)$ + original constraints + consistency constraints.

The Lasserre hierarchy (constraints)

- Y is psd. (i.e. find vectors $\mathbf{U}_S$ satisfying $Y_{S_1,S_2} = \langle \mathbf{U}_{S_1}, \mathbf{U}_{S_2} \rangle$)

The Lasserre hierarchy (constraints)

- Y is psd. (i.e. find vectors $\mathbf{U}_S$ satisfying $Y_{S_1,S_2} = \langle \mathbf{U}_{S_1}, \mathbf{U}_{S_2} \rangle$)
- Y_{S_1,S_2} only depends on $S_1 \cup S_2$. ($Y_{S_1,S_2} = \mathbb{P}[\text{All vars in } S_1 \cup S_2 \text{ are } 1]$)

The Lasserre hierarchy (constraints)

- Y is psd. (i.e. find vectors $\mathbf{U}_S$ satisfying $Y_{S_1,S_2} = \langle \mathbf{U}_{S_1}, \mathbf{U}_{S_2} \rangle$)
- Y_{S_1,S_2} only depends on $S_1 \cup S_2$. ($Y_{S_1,S_2} = \mathbb{P}[\text{All vars in } S_1 \cup S_2 \text{ are } 1]$)
- Original quadratic constraints as inner products.

SDP for Independent Set

$$
\begin{array}{lll}
\text{maximize} & \sum_{i \in V} \left|\mathbf{U}_{\{i\}}\right|^2 & \\
\text{subject to} & \langle \mathbf{U}_{\{i\}}, \mathbf{U}_{\{j\}} \rangle = 0 & \forall\, (i,j) \in E \\
& \langle \mathbf{U}_{S_1}, \mathbf{U}_{S_2} \rangle = \langle \mathbf{U}_{S_3}, \mathbf{U}_{S_4} \rangle & \forall\, S_1 \cup S_2 = S_3 \cup S_4 \\
& \langle \mathbf{U}_{S_1}, \mathbf{U}_{S_2} \rangle \in [0, 1] & \forall S_1, S_2
\end{array}
$$

The “Mixed” hierarchy

- Motivated by [Raghavendra 08]. Used by [CS 08] for Hypergraph Independent Set.

- Captures what we actually know how to use about Lasserre solutions.

The "Mixed" hierarchy

- Motivated by [Raghavendra 08]. Used by [CS 08] for Hypergraph Independent Set.
- Captures what we actually know how to use about Lasserre solutions.
- Level r has
  - Variables X_S for $|S| \leq r$ and all Sherali-Adams constraints.
  - Vectors $\mathbf{U}_0, \mathbf{U}_1, \ldots, \mathbf{U}_n$ satisfying

$$\langle \mathbf{U}_i, \mathbf{U}_j \rangle = X_{\{i,j\}}, \langle \mathbf{U}_0, \mathbf{U}_i \rangle = X_{\{i\}} \text{ and } |\mathbf{U}_0| = 1.$$

Hands-on: Deriving some constraints

The triangle inequality

- $|\mathbf{U}_i - \mathbf{U}_j|^2 + |\mathbf{U}_j - \mathbf{U}_k|^2 \;\geq\; |\mathbf{U}_i - \mathbf{U}_k|^2$ is equivalent to

$$\langle \mathbf{U}_i - \mathbf{U}_j, \mathbf{U}_k - \mathbf{U}_j \rangle \;\geq\; 0$$

The triangle inequality

- $|\mathbf{U}_i - \mathbf{U}_j|^2 + |\mathbf{U}_j - \mathbf{U}_k|^2 \geq |\mathbf{U}_i - \mathbf{U}_k|^2$ is equivalent to
$$\langle \mathbf{U}_i - \mathbf{U}_j, \mathbf{U}_k - \mathbf{U}_j \rangle \geq 0$$
- $\mathrm{Mix}^{(3)} \implies \exists$ distribution on z_i, z_j, z_k such that $\mathbb{E}[z_i \cdot z_j] = \langle \mathbf{U}_i, \mathbf{U}_j \rangle$ (and so on).

The triangle inequality

- $|\mathbf{U}_i - \mathbf{U}_j|^2 + |\mathbf{U}_j - \mathbf{U}_k|^2 \ \geq \ |\mathbf{U}_i - \mathbf{U}_k|^2$ is equivalent to

$$\langle \mathbf{U}_i - \mathbf{U}_j, \mathbf{U}_k - \mathbf{U}_j \rangle \ \geq \ 0$$

- $\text{Mix}^{(3)} \implies \exists$ distribution on z_i, z_j, z_k such that $\mathbb{E}[z_i \cdot z_j] = \langle \mathbf{U}_i, \mathbf{U}_j \rangle$ (and so on).

- For all integer solutions $(z_i - z_j) \cdot (z_k - z_j) \ \geq \ 0$.

The triangle inequality

- $|\mathbf{U}_i - \mathbf{U}_j|^2 + |\mathbf{U}_j - \mathbf{U}_k|^2 \geq |\mathbf{U}_i - \mathbf{U}_k|^2$ is equivalent to
$$\langle \mathbf{U}_i - \mathbf{U}_j, \mathbf{U}_k - \mathbf{U}_j \rangle \geq 0$$
- $\text{Mix}^{(3)} \implies \exists$ distribution on z_i, z_j, z_k such that $\mathbb{E}[z_i \cdot z_j] = \langle \mathbf{U}_i, \mathbf{U}_j \rangle$ (and so on).
- For all integer solutions $(z_i - z_j) \cdot (z_k - z_j) \geq 0$.
$$\therefore \langle \mathbf{U}_i - \mathbf{U}_j, \mathbf{U}_k - \mathbf{U}_j \rangle = \mathbb{E}\left[(z_i - z_j) \cdot (z_k - z_j)\right] \geq 0$$

“Clique constraints” for Independent Set

- For every clique B in a graph, adding the constraint

$$\sum_{i \in B} x_i \leq 1$$

makes the independent set LP tight for perfect graphs.

“Clique constraints” for Independent Set

- For every clique B in a graph, adding the constraint

$$\sum_{i \in B} x_i \leq 1$$

makes the independent set LP tight for perfect graphs.

- Too many constraints, but all implied by one level of the mixed hierarchy.

“Clique constraints” for Independent Set

- For every clique B in a graph, adding the constraint

$$\sum_{i \in B} x_i \leq 1$$

  makes the independent set LP tight for perfect graphs.

- Too many constraints, but all implied by one level of the mixed hierarchy.

- For $i, j \in B$, $\langle \mathbf{U}_i, \mathbf{U}_j \rangle = 0$. By Pythagoras,

$$\sum_{i \in B} \left\langle \mathbf{U}_0, \frac{\mathbf{U}_i}{|\mathbf{U}_i|} \right\rangle^2 \leq |\mathbf{U}_0|^2 = 1 \implies \sum_{i \in B} \frac{x_i^2}{x_i} \leq 1.$$

- Derived by Lovász using the ϑ-function.

The Lovász-Schrijver Hierarchy

The Lovász-Schrijver Hierarchy

- Start with a 0/1 integer program and a relaxation P. Define tigher relaxation $LS(P)$.

The Lovász-Schrijver Hierarchy

- Start with a 0/1 integer program and a relaxation P. Define tigher relaxation $LS(P)$.
- Hope: Fractional $(x_1, \ldots, x_n) = \mathbb{E}\left[(z_1, \ldots, z_n)\right]$ for integral $(z_1, \ldots, z_n)$

The Lovász-Schrijver Hierarchy

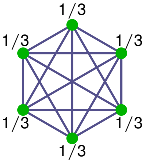
- Start with a 0/1 integer program and a relaxation P. Define tigher relaxation $LS(P)$.
- Hope: Fractional $(x_1, \ldots, x_n) = \mathbb{E}[(z_1, \ldots, z_n)]$ for integral $(z_1, \ldots, z_n)$
- Restriction: $\mathbf{x} = (x_1, \ldots, x_n) \in LS(P)$ if $\exists Y$ satisfying
  (think $Y_{ij} = \mathbb{E}[z_i z_j] = \mathbb{P}[z_i \wedge z_j]$)
  - $Y = Y^T$
  - $Y_{ii} = x_i \qquad \forall i$
  - $\frac{Y_i}{x_i} \in P, \ \frac{\mathbf{x} - Y_i}{1 - x_i} \in P \qquad \forall i$
  - $Y \succeq 0$

The Lovász-Schrijver Hierarchy

- Start with a 0/1 integer program and a relaxation P. Define tigher relaxation $LS(P)$.
- Hope: Fractional $(x_1, \ldots, x_n) = \mathbb{E}[(z_1, \ldots, z_n)]$ for integral $(z_1, \ldots, z_n)$
- Restriction: $\mathbf{x} = (x_1, \ldots, x_n) \in LS(P)$ if $\exists Y$ satisfying
  (think $Y_{ij} = \mathbb{E}[z_i z_j] = \mathbb{P}[z_i \wedge z_j]$)
  - $Y = Y^T$
  - $Y_{ii} = x_i \qquad \forall i$
  - $\frac{Y_i}{x_i} \in P, \; \frac{\mathbf{x} - Y_i}{1 - x_i} \in P \qquad \forall i$
  - $Y \succeq 0$
- Above is an LP (SDP) in $n^2 + n$ variables.

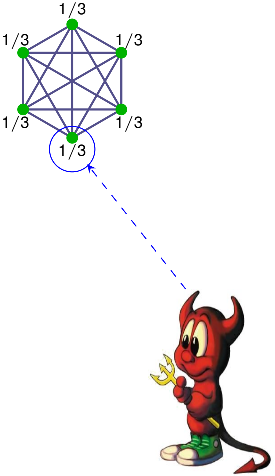
Lovász-Schrijver in action

- r^{th} level optimizes over distributions conditioned on r variables.



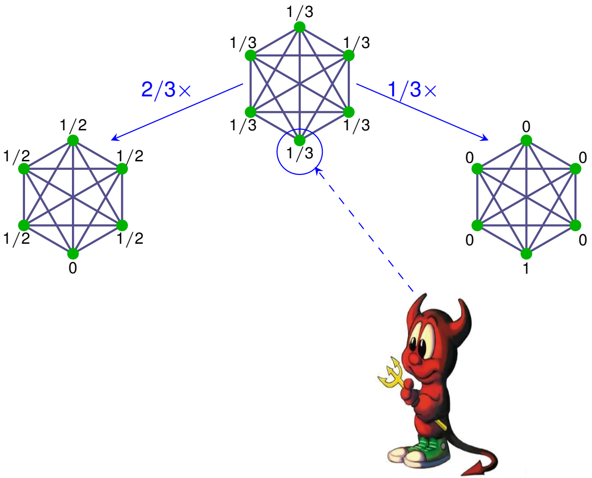
Lovász-Schrijver in action

- r^{th} level optimizes over distributions conditioned on r variables.



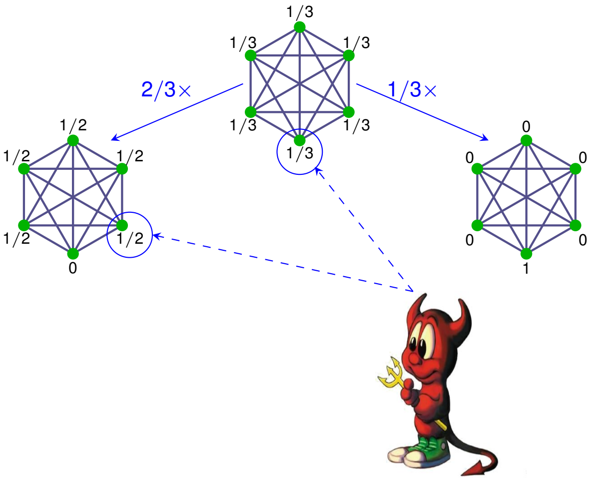
Lovász-Schrijver in action

- r^{th} level optimizes over distributions conditioned on r variables.



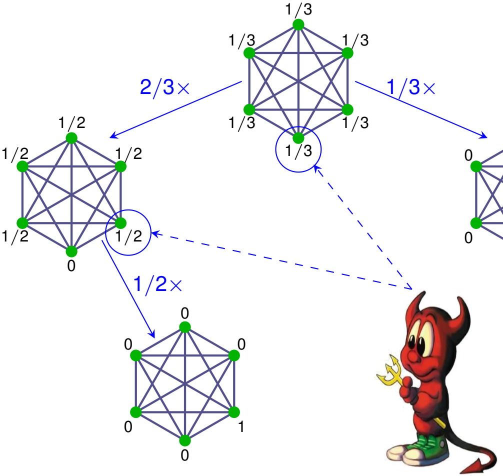
Lovász-Schrijver in action

- r^{th} level optimizes over distributions conditioned on r variables.



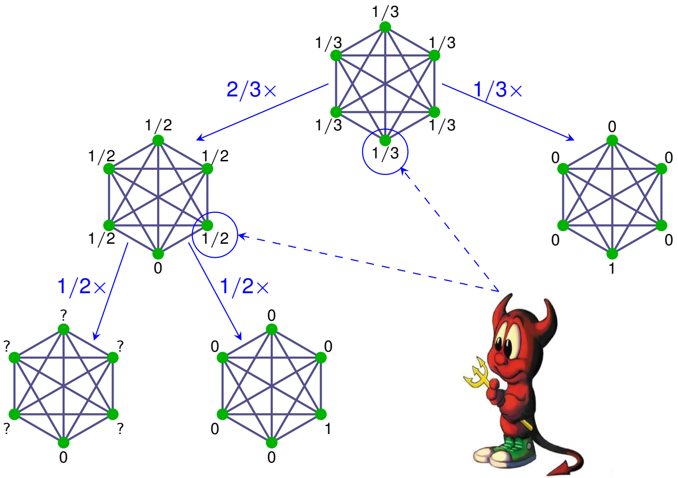
Lovász-Schrijver in action

- r^{th} level optimizes over distributions conditioned on r variables.



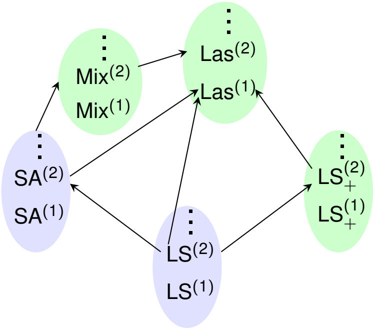
Lovász-Schrijver in action

- r^{th} level optimizes over distributions conditioned on r variables.

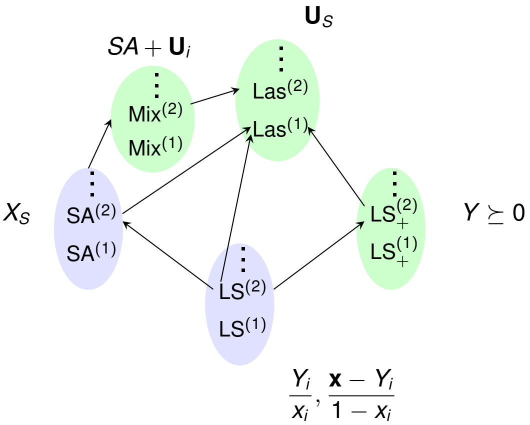


And if you just woke up . . .

And if you just woke up . . .



And if you just woke up . . .



Integrality Gaps for Expanding CSPs

CSP Expansion

- MAX k-CSP: m constraints on k-tuples of (n) boolean variables. Satisfy maximum. e.g. MAX 3-XOR (linear equations mod 2)

$$z_1 + z_2 + z_3 = 0 \qquad z_3 + z_4 + z_5 = 1 \qquad \cdots$$

CSP Expansion

- MAX k-CSP: m constraints on k-tuples of (n) boolean variables. Satisfy maximum. e.g. MAX 3-XOR (linear equations mod 2)

$$z_1 + z_2 + z_3 = 0 \qquad z_3 + z_4 + z_5 = 1 \qquad \cdots$$

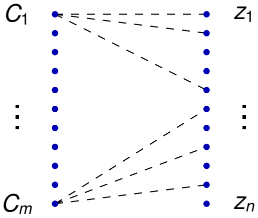
- Expansion: Every set S of constraints involves at least $\beta|S|$ variables (for $|S| < \alpha m$).

CSP Expansion

- MAX k-CSP: m constraints on k-tuples of (n) boolean variables. Satisfy maximum. e.g. MAX 3-XOR (linear equations mod 2)

$$z_1 + z_2 + z_3 = 0 \qquad z_3 + z_4 + z_5 = 1 \qquad \cdots$$

- Expansion: Every set S of constraints involves at least $\beta|S|$ variables (for $|S| < \alpha m$).

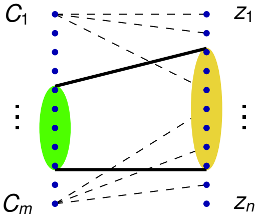


CSP Expansion

- MAX k-CSP: m constraints on k-tuples of (n) boolean variables. Satisfy maximum. e.g. MAX 3-XOR (linear equations mod 2)

$$z_1 + z_2 + z_3 = 0 \qquad z_3 + z_4 + z_5 = 1 \qquad \cdots$$

- Expansion: Every set S of constraints involves at least $\beta|S|$ variables (for $|S| < \alpha m$).

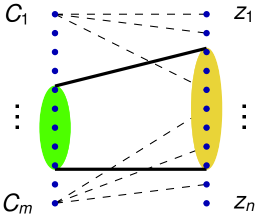


CSP Expansion

- MAX k-CSP: m constraints on k-tuples of (n) boolean variables. Satisfy maximum. e.g. MAX 3-XOR (linear equations mod 2)

$$z_1 + z_2 + z_3 = 0 \qquad z_3 + z_4 + z_5 = 1 \qquad \cdots$$

- Expansion: Every set S of constraints involves at least $\beta|S|$ variables (for $|S| < \alpha m$).



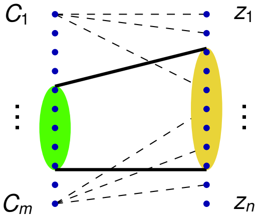
In fact, $\gamma|S|$ variables appearing in only one constraint in S.

CSP Expansion

- MAX k-CSP: m constraints on k-tuples of (n) boolean variables. Satisfy maximum. e.g. MAX 3-XOR (linear equations mod 2)

$$z_1 + z_2 + z_3 = 0 \qquad z_3 + z_4 + z_5 = 1 \qquad \cdots$$

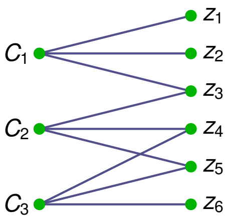
- Expansion: Every set S of constraints involves at least $\beta|S|$ variables (for $|S| < \alpha m$).



In fact, $\gamma|S|$ variables appearing in only one constraint in S.

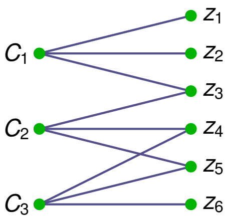
- Used extensively in proof complexity e.g. [BW01], [BGHMP03]. For LS_+ by [AAT04].

Local Satisfiability



- Take $\gamma = 0.9$
- Can show any three 3-XOR constraints are simultaneously satisfiable.

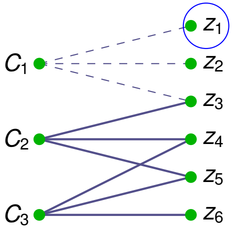
Local Satisfiability



- Take $\gamma = 0.9$
- Can show any three 3-XOR constraints are simultaneously satisfiable.

$$\mathbb{E}_{z_1 \ldots z_6}\left[C_1(z_1, z_2, z_3) \cdot C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6)\right]$$

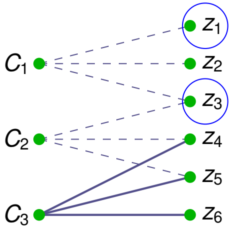
Local Satisfiability



- Take $\gamma = 0.9$
- Can show any three 3-XOR constraints are simultaneously satisfiable.

$$\begin{aligned} &\mathbb{E}_{z_1 \ldots z_6}\left[C_1(z_1, z_2, z_3) \cdot C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6)\right] \\ = \; &\mathbb{E}_{z_2 \ldots z_6}\left[C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6) \cdot \mathbb{E}_{z_1}\left[C_1(z_1, z_2, z_3)\right]\right] \end{aligned}$$

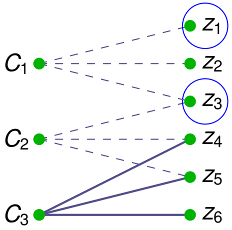
Local Satisfiability



- Take $\gamma = 0.9$
- Can show any three 3-XOR constraints are simultaneously satisfiable.

$$
\begin{aligned}
& \mathbb{E}_{z_1 \ldots z_6}\left[C_1(z_1, z_2, z_3) \cdot C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6)\right] \\
= \; & \mathbb{E}_{z_2 \ldots z_6}\left[C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6) \cdot \mathbb{E}_{z_1}\left[C_1(z_1, z_2, z_3)\right]\right] \\
= \; & \mathbb{E}_{z_4, z_5, z_6}\left[C_3(z_4, z_5, z_6) \cdot \mathbb{E}_{z_3}\left[C_2(z_3, z_4, z_5)\right] \cdot (1/2)\right]
\end{aligned}
$$

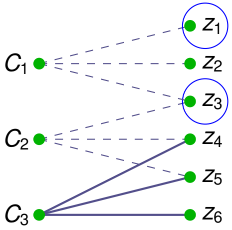
Local Satisfiability



- Take $\gamma = 0.9$
- Can show any three 3-XOR constraints are simultaneously satisfiable.

$$
\begin{aligned}
& \mathbb{E}_{z_1 \ldots z_6}\left[C_1(z_1, z_2, z_3) \cdot C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6)\right] \\
= \; & \mathbb{E}_{z_2 \ldots z_6}\left[C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6) \cdot \mathbb{E}_{z_1}\left[C_1(z_1, z_2, z_3)\right]\right] \\
= \; & \mathbb{E}_{z_4, z_5, z_6}\left[C_3(z_4, z_5, z_6) \cdot \mathbb{E}_{z_3}\left[C_2(z_3, z_4, z_5)\right] \cdot (1/2)\right] \\
= \; & 1/8
\end{aligned}
$$

Local Satisfiability



- Take $\gamma = 0.9$
- Can show any three 3-XOR constraints are simultaneously satisfiable.
- Can take $\gamma \approx (k-2)$ and any αn constraints.
- Just require $\mathbb{E}[C(z_1, \ldots, z_k)]$ over any $k-2$ vars to be constant.

$$\begin{aligned} & \mathbb{E}_{z_1 \ldots z_6}\left[C_1(z_1, z_2, z_3) \cdot C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6)\right] \\ = \; & \mathbb{E}_{z_2 \ldots z_6}\left[C_2(z_3, z_4, z_5) \cdot C_3(z_4, z_5, z_6) \cdot \mathbb{E}_{z_1}\left[C_1(z_1, z_2, z_3)\right]\right] \\ = \; & \mathbb{E}_{z_4, z_5, z_6}\left[C_3(z_4, z_5, z_6) \cdot \mathbb{E}_{z_3}\left[C_2(z_3, z_4, z_5)\right] \cdot (1/2)\right] \\ = \; & 1/8 \end{aligned}$$

Sherali-Adams LP for CSPs

Variables: $X_{(S,\alpha)}$ for $|S| \leq t$, partial assignments $\alpha \in \{0,1\}^S$

$$\begin{array}{llll}
\text{maximize} & \displaystyle\sum_{i=1}^{m} \sum_{\alpha \in \{0,1\}^{T_i}} C_i(\alpha) \cdot X_{(T_i,\alpha)} & & \\
\text{subject to} & X_{(S\cup\{i\},\alpha\circ 0)} + X_{(S\cup\{i\},\alpha\circ 1)} = X_{(S,\alpha)} & & \forall i \notin S \\
 & X_{(S,\alpha)} \geq 0 & & \\
 & X_{(\emptyset,\emptyset)} = 1 & &
\end{array}$$

Sherali-Adams LP for CSPs

Variables: $X_{(S,\alpha)}$ for $|S| \le t$, partial assignments $\alpha \in \{0,1\}^S$

$$\begin{array}{lrcll} \text{maximize} & \displaystyle\sum_{i=1}^{m} \sum_{\alpha \in \{0,1\}^{T_i}} C_i(\alpha) \cdot X_{(T_i,\alpha)} & & & \\ \text{subject to} & X_{(S\cup\{i\},\alpha\circ 0)} + X_{(S\cup\{i\},\alpha\circ 1)} & = & X_{(S,\alpha)} & \forall i \notin S \\ & X_{(S,\alpha)} & \ge & 0 & \\ & X_{(\emptyset,\emptyset)} & = & 1 & \end{array}$$

- $X_{(S,\alpha)} \sim \mathbb{P}[\text{Vars in } S \text{ assigned according to } \alpha]$

Sherali-Adams LP for CSPs

Variables: $X_{(S,\alpha)}$ for $|S| \leq t$, partial assignments $\alpha \in \{0,1\}^S$

$$\begin{array}{lrcll} \text{maximize} & \multicolumn{3}{c}{\displaystyle\sum_{i=1}^{m} \sum_{\alpha \in \{0,1\}^{T_i}} C_i(\alpha) \cdot X_{(T_i,\alpha)}} & \\ \text{subject to} & X_{(S \cup \{i\}, \alpha \circ 0)} + X_{(S \cup \{i\}, \alpha \circ 1)} & = & X_{(S,\alpha)} & \forall i \notin S \\ & X_{(S,\alpha)} & \geq & 0 & \\ & X_{(\emptyset,\emptyset)} & = & 1 & \end{array}$$

- $X_{(S,\alpha)} \sim \mathbb{P}[\text{Vars in } S \text{ assigned according to } \alpha]$
- Need distributions $D(S)$ such that $D(S_1), D(S_2)$ agree on $S_1 \cap S_2$.

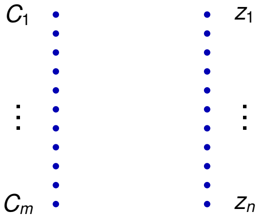
Sherali-Adams LP for CSPs

Variables: $X_{(S,\alpha)}$ for $|S| \leq t$, partial assignments $\alpha \in \{0,1\}^S$

$$\begin{array}{lrcll} \text{maximize} & \multicolumn{3}{c}{\sum_{i=1}^{m} \sum_{\alpha \in \{0,1\}^{T_i}} C_i(\alpha) \cdot X_{(T_i,\alpha)}} & \\ \text{subject to} & X_{(S \cup \{i\}, \alpha \circ 0)} + X_{(S \cup \{i\}, \alpha \circ 1)} & = & X_{(S,\alpha)} & \forall i \notin S \\ & X_{(S,\alpha)} & \geq & 0 & \\ & X_{(\emptyset,\emptyset)} & = & 1 & \end{array}$$

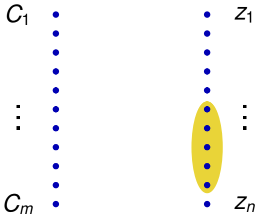
- $X_{(S,\alpha)} \sim \mathbb{P}[\text{Vars in } S \text{ assigned according to } \alpha]$
- Need distributions $D(S)$ such that $D(S_1)$, $D(S_2)$ agree on $S_1 \cap S_2$.
- Distributions should “locally look like" supported on satisfying assignments.

Obtaining integrality gaps for CSPs [BGMT 09]



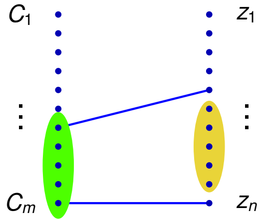
- Want to define distribution $D(S)$ for set S of variables.

Obtaining integrality gaps for CSPs [BGMT 09]



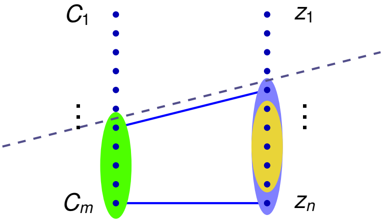
- Want to define distribution $D(S)$ for set S of variables.

Obtaining integrality gaps for CSPs [BGMT 09]



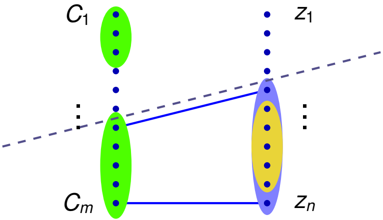
- Want to define distribution $D(S)$ for set S of variables.

Obtaining integrality gaps for CSPs [BGMT 09]



- Want to define distribution $D(S)$ for set S of variables.
- Find set of constraints $\mathcal{C}$ such that $G - \mathcal{C} - S$ remains expanding.
  $D(S)$ = uniform over assignments satisfying $\mathcal{C}$

Obtaining integrality gaps for CSPs [BGMT 09]



- Want to define distribution $D(S)$ for set S of variables.
- Find set of constraints $\mathcal{C}$ such that $G - \mathcal{C} - S$ remains expanding.
  $D(S)$ = uniform over assignments satisfying $\mathcal{C}$
- Remaining constraints “independent" of this assignment.
- Gives optimal integrality gaps for $\Omega(n)$ levels in the mixed hierarchy.

Vectors for Linear CSPs

A “new look” Lasserre

- Start with a $\{-1, 1\}$ quadratic integer program.
  $(z_1, \ldots, z_n) \to ((-1)^{z_1}, \ldots, (-1)^{z_n})$

A “new look” Lasserre

- Start with a $\{-1, 1\}$ quadratic integer program.
  $(z_1, \ldots, z_n) \to ((-1)^{z_1}, \ldots, (-1)^{z_n})$

- Define big variables $\tilde{Z}_S = \prod_{i \in S} (-1)^{z_i}$.

A "new look" Lasserre

- Start with a $\{-1,1\}$ quadratic integer program.
  $(z_1,\ldots,z_n) \to ((-1)^{z_1},\ldots,(-1)^{z_n})$
- Define big variables $\tilde{Z}_S = \prod_{i\in S}(-1)^{z_i}$.
- Consider the psd matrix $\tilde{Y}$

$$\tilde{Y}_{S_1,S_2} = \mathbb{E}\left[\tilde{Z}_{S_1}\cdot\tilde{Z}_{S_2}\right] = \mathbb{E}\left[\prod_{i\in S_1\Delta S_2}(-1)^{z_i}\right]$$

A "new look" Lasserre

- Start with a $\{-1, 1\}$ quadratic integer program.
  $(z_1, \ldots, z_n) \to ((-1)^{z_1}, \ldots, (-1)^{z_n})$

- Define big variables $\tilde{Z}_S = \prod_{i \in S} (-1)^{z_i}$.

- Consider the psd matrix $\tilde{Y}$

$$\tilde{Y}_{S_1, S_2} = \mathbb{E}\left[\tilde{Z}_{S_1} \cdot \tilde{Z}_{S_2}\right] = \mathbb{E}\left[\prod_{i \in S_1 \Delta S_2} (-1)^{z_i}\right]$$

- Write program for inner products of vectors $\mathbf{W}_S$ s.t.
  $\tilde{Y}_{S_1, S_2} = \langle \mathbf{W}_{S_1}, \mathbf{W}_{S_2} \rangle$

Gaps for 3-XOR

SDP for MAX 3-XOR

$$\text{maximize} \quad \sum_{C_i \equiv (z_{i_1}+z_{i_2}+z_{i_3}=b_i)} \frac{1+(-1)^{b_i}\left\langle \mathbf{W}_{\{i_1,i_2,i_3\}}, \mathbf{W}_{\emptyset}\right\rangle}{2}$$

$$\text{subject to} \quad \left\langle \mathbf{W}_{S_1}, \mathbf{W}_{S_2}\right\rangle = \left\langle \mathbf{W}_{S_3}, \mathbf{W}_{S_4}\right\rangle \qquad \forall\ S_1 \Delta S_2 = S_3 \Delta S_4$$

$$|\mathbf{W}_S| = 1 \qquad \forall S,\ |S| \le r$$

Gaps for 3-XOR

SDP for MAX 3-XOR

$$\begin{array}{lll}
\text{maximize} & \displaystyle\sum_{C_i \equiv (z_{i_1} + z_{i_2} + z_{i_3} = b_i)} \frac{1 + (-1)^{b_i} \langle \mathbf{W}_{\{i_1, i_2, i_3\}}, \mathbf{W}_{\emptyset} \rangle}{2} & \\
\text{subject to} & \langle \mathbf{W}_{S_1}, \mathbf{W}_{S_2} \rangle = \langle \mathbf{W}_{S_3}, \mathbf{W}_{S_4} \rangle & \forall\ S_1 \Delta S_2 = S_3 \Delta S_4 \\
& |\mathbf{W}_S| = 1 & \forall S,\ |S| \leq r
\end{array}$$

- [Schoenebeck'08]: If *width* $2r$ resolution does not derive contradiction, then SDP value =1 after r levels of Lasserre.

Gaps for 3-XOR

SDP for MAX 3-XOR

$$\text{maximize} \quad \sum_{C_i \equiv (z_{i_1} + z_{i_2} + z_{i_3} = b_i)} \frac{1 + (-1)^{b_i} \left\langle \mathbf{W}_{\{i_1, i_2, i_3\}}, \mathbf{W}_{\emptyset} \right\rangle}{2}$$

$$\text{subject to} \quad \left\langle \mathbf{W}_{S_1}, \mathbf{W}_{S_2} \right\rangle = \left\langle \mathbf{W}_{S_3}, \mathbf{W}_{S_4} \right\rangle \qquad \forall\ S_1 \Delta S_2 = S_3 \Delta S_4$$

$$|\mathbf{W}_S| = 1 \qquad \forall S,\ |S| \leq r$$

- [Schoenebeck'08]: If *width* $2r$ resolution does not derive contradiction, then SDP value =1 after r levels of Lasserre.
- Expansion guarantees there are no width $2r$ contradictions.

Gaps for 3-XOR

SDP for MAX 3-XOR

$$\begin{array}{lll}
\text{maximize} & \displaystyle\sum_{C_i \equiv (z_{i_1}+z_{i_2}+z_{i_3}=b_i)} \frac{1+(-1)^{b_i}\left\langle \mathbf{W}_{\{i_1,i_2,i_3\}}, \mathbf{W}_{\emptyset}\right\rangle}{2} & \\
\text{subject to} & \left\langle \mathbf{W}_{S_1}, \mathbf{W}_{S_2}\right\rangle = \left\langle \mathbf{W}_{S_3}, \mathbf{W}_{S_4}\right\rangle & \forall\ S_1 \Delta S_2 = S_3 \Delta S_4 \\
& |\mathbf{W}_S| = 1 & \forall S,\ |S| \le r
\end{array}$$

- [Schoenebeck'08]: If *width* $2r$ resolution does not derive contradiction, then SDP value =1 after r levels of Lasserre.
- Expansion guarantees there are no width $2r$ contradictions.
- Used by [FO 06], [STT 07] for LS_+ hierarchy.

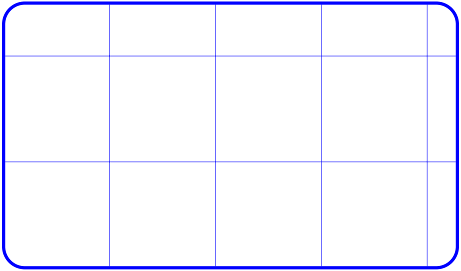
Schonebeck's construction

Schonebeck's construction

- $z_1 + z_2 + z_3 = 1 \mod 2 \implies (-1)^{z_1+z_2} = -(-1)^{z_3}$
  $\implies \mathbf{W}_{\{1,2\}} = -\mathbf{W}_{\{3\}}$

Schonebeck's construction

- $z_1 + z_2 + z_3 = 1 \mod 2 \implies (-1)^{z_1+z_2} = -(-1)^{z_3}$
  $\implies \mathbf{W}_{\{1,2\}} = -\mathbf{W}_{\{3\}}$
- Equations of width $2r$ divide $|S| \leq r$ into equivalence classes. Choose orthogonal $\mathbf{e}_{\mathcal{C}}$ for each class $\mathcal{C}$.

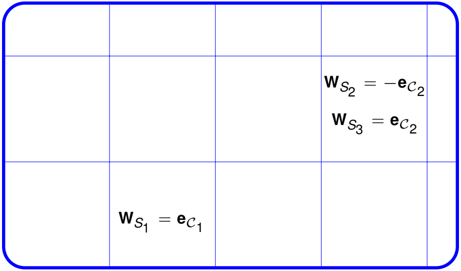


Schonebeck's construction

- $z_1 + z_2 + z_3 = 1 \mod 2 \implies (-1)^{z_1+z_2} = -(-1)^{z_3}$
  $\implies \mathbf{W}_{\{1,2\}} = -\mathbf{W}_{\{3\}}$

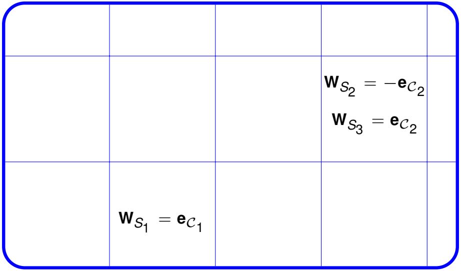
- Equations of width $2r$ divide $|S| \leq r$ into equivalence classes. Choose orthogonal $\mathbf{e}_{\mathcal{C}}$ for each class $\mathcal{C}$.

- No contradictions ensure each $S \in \mathcal{C}$ can be uniquely assigned $\pm\mathbf{e}_{\mathcal{C}}$.



Schonebeck's construction

- $z_1 + z_2 + z_3 = 1 \mod 2 \implies (-1)^{z_1+z_2} = -(-1)^{z_3}$
  $\implies \mathbf{W}_{\{1,2\}} = -\mathbf{W}_{\{3\}}$
- Equations of width $2r$ divide $|S| \leq r$ into equivalence classes. Choose orthogonal $\mathbf{e}_{\mathcal{C}}$ for each class $\mathcal{C}$.
- No contradictions ensure each $S \in \mathcal{C}$ can be uniquely assigned $\pm\mathbf{e}_{\mathcal{C}}$.
- Relies heavily on constraints being linear equations.



Reductions

Spreading the hardness around (Reductions) [T]

- If problem A reduces to B, can we say
  Integrality Gap for A $\implies$ Integrality Gap for B?

Spreading the hardness around (Reductions) [T]

- If problem A reduces to B, can we say
  Integrality Gap for A $\implies$ Integrality Gap for B?

- Reductions are (often) local algorithms.

Spreading the hardness around (Reductions) [T]

- If problem A reduces to B, can we say
  Integrality Gap for A $\implies$ Integrality Gap for B?

- Reductions are (often) local algorithms.

- Reduction from integer program A to integer program B. Each variable z_i' of B is a boolean function of few (say 5) variables $z_{i_1}, \ldots, z_{i_5}$ of A.

Spreading the hardness around (Reductions) [T]

- If problem A reduces to B, can we say
  Integrality Gap for A $\implies$ Integrality Gap for B?
- Reductions are (often) local algorithms.
- Reduction from integer program A to integer program B. Each variable z_i' of B is a boolean function of few (say 5) variables $z_{i_1}, \ldots, z_{i_5}$ of A.
- To show: If A has good vector solution, so does B.

Spreading the hardness around (Reductions) [T]

- If problem A reduces to B, can we say
  Integrality Gap for A $\implies$ Integrality Gap for B?

- Reductions are (often) local algorithms.

- Reduction from integer program A to integer program B. Each variable z'_i of B is a boolean function of few (say 5) variables $z_{i_1}, \ldots, z_{i_5}$ of A.

- To show: If A has good vector solution, so does B.

- Question posed in [AAT 04]. First done by [KV 05] from Unique Games to Sparsest Cut.

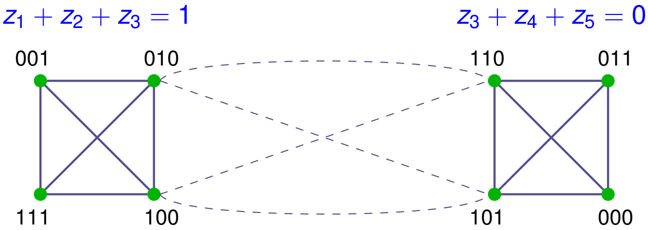
What can be proved

	NP-hard	UG-hard	Gap	Levels
MAX k-CSP	$\frac{2^k}{2^{\sqrt{2k}}}$	$\frac{2^k}{k+o(k)}$	$\frac{2^k}{2k}$	$\Omega(n)$
Independent Set	$\frac{n}{2^{(\log n)^{3/4+\epsilon}}}$		$\frac{n}{2^{c_1\sqrt{\log n \log\log n}}}$	$2^{c_2\sqrt{\log n \log\log n}}$
Approximate Graph Coloring	l vs. $2^{\frac{1}{25}\log^2 l}$		l vs. $\frac{2^{l/2}}{4l^2}$	$\Omega(n)$
Chromatic Number	$\frac{n}{2^{(\log n)^{3/4+\epsilon}}}$		$\frac{n}{2^{c_1\sqrt{\log n \log\log n}}}$	$2^{c_2\sqrt{\log n \log\log n}}$
Vertex Cover	1.36	2 - ϵ	1.36	$\Omega(n^\delta)$

- All the above results are for the Lasserre hierarchy.

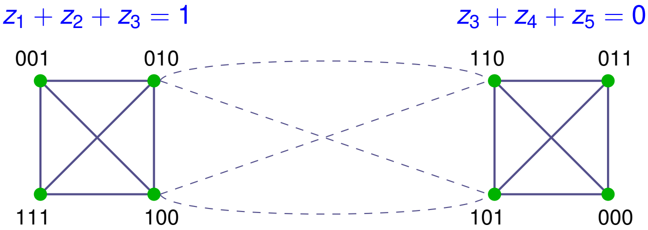
The FGLSS Construction

- Reduces MAX k-CSP to Independent Set in graph G_Φ.



The FGLSS Construction

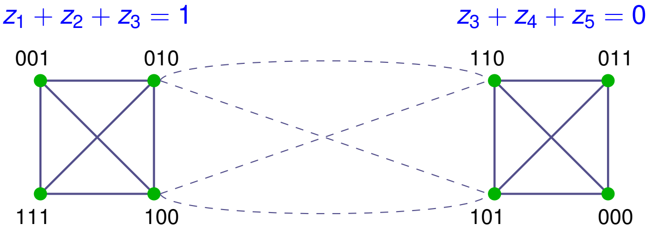
- Reduces MAX k-CSP to Independent Set in graph G_Φ.



- Need vectors for subsets of vertices in the G_Φ.

The FGLSS Construction

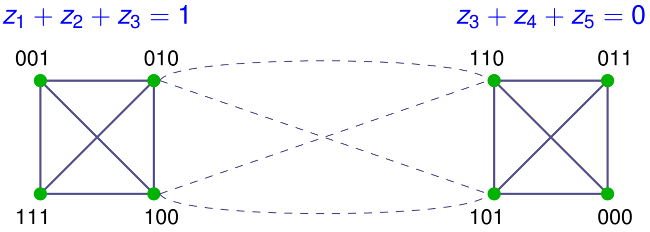
- Reduces MAX k-CSP to Independent Set in graph G_Φ.



- Need vectors for subsets of vertices in the G_Φ.
- Every vertex (or set of vertices) in G_Φ is an indicator function!

The FGLSS Construction

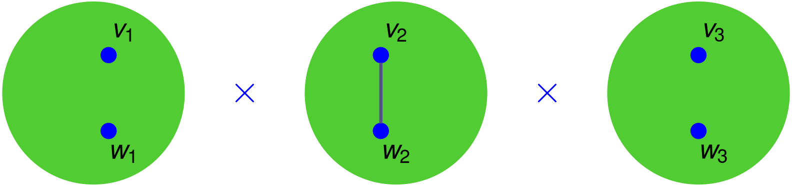
- Reduces MAX k-CSP to Independent Set in graph G_Φ.



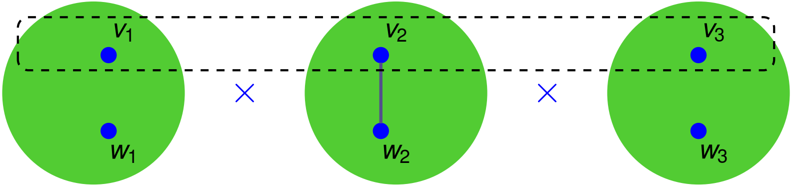
- Need vectors for subsets of vertices in the G_Φ.
- Every vertex (or set of vertices) in G_Φ is an indicator function!

$\mathbf{U}_{\{(z_1,z_2,z_3)=(0,0,1)\}} = \frac{1}{8}(\mathbf{W}_\emptyset + \mathbf{W}_{\{1\}} + \mathbf{W}_{\{2\}} - \mathbf{W}_{\{3\}} + \mathbf{W}_{\{1,2\}} - \mathbf{W}_{\{2,3\}} - \mathbf{W}_{\{1,3\}} - \mathbf{W}_{\{1,2,3\}})$

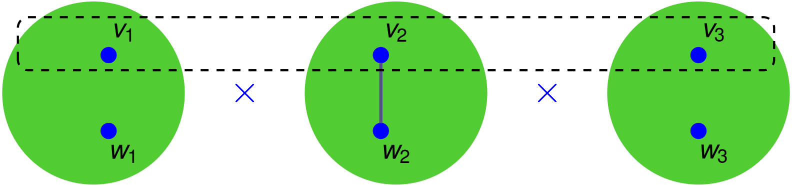
Graph Products



Graph Products

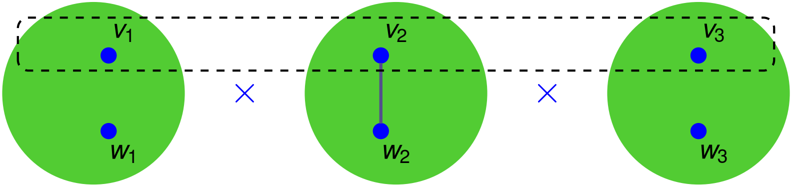


Graph Products



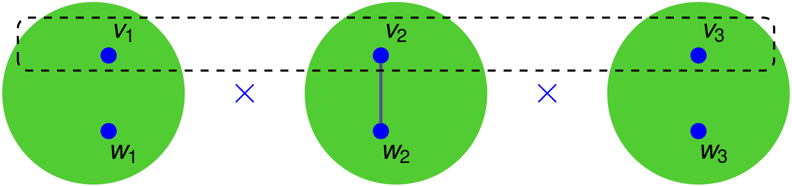
- $\overline{\mathbf{U}}_{\{(v_1,v_2,v_3)\}} = ?$

Graph Products



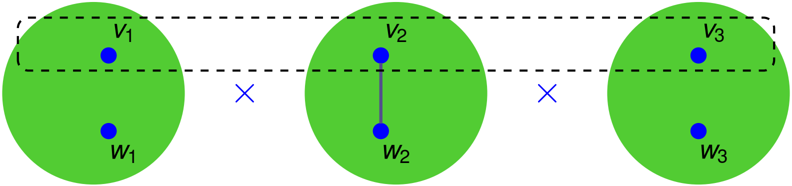
- $\overline{\mathbf{U}}_{\{(v_1,v_2,v_3)\}} = \mathbf{U}_{\{v_1\}} \otimes \mathbf{U}_{\{v_2\}} \otimes \mathbf{U}_{\{v_3\}}$

Graph Products



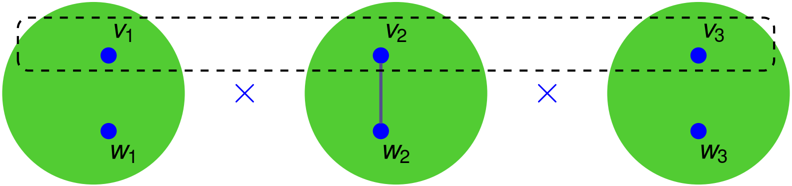
- $\overline{\mathbf{U}}_{\{(v_1,v_2,v_3)\}} = \mathbf{U}_{\{v_1\}} \otimes \mathbf{U}_{\{v_2\}} \otimes \mathbf{U}_{\{v_3\}}$
- Similar transformation for sets (project to each copy of G).

Graph Products



- $\overline{\mathbf{U}}_{\{(v_1,v_2,v_3)\}} = \mathbf{U}_{\{v_1\}} \otimes \mathbf{U}_{\{v_2\}} \otimes \mathbf{U}_{\{v_3\}}$
- Similar transformation for sets (project to each copy of G).
- Intuition: Independent set in product graph is product of independent sets in G.

Graph Products



- $\overline{\mathbf{U}}_{\{(v_1,v_2,v_3)\}} = \mathbf{U}_{\{v_1\}} \otimes \mathbf{U}_{\{v_2\}} \otimes \mathbf{U}_{\{v_3\}}$
- Similar transformation for sets (project to each copy of G).
- Intuition: Independent set in product graph is product of independent sets in G.
- Together give a gap of $\frac{n}{2^{O(\sqrt{\log n \log\log n})}}$.

A few problems

Problem 1: Lasserre Gaps

- Show an integrality gap of $2 - \epsilon$ for Vertex Cover, even for $O(1)$ levels of the Lasserre hierarchy.
- Obtain integrality gaps Unique Games (and Small-Set Expansion)
  - Gaps for $O((\log \log n)^{1/4})$ levels of mixed hierarchy were obtained by [RS 09] and [KS 09].
  - Extension to Lasserre?

Problem 2: Generalize Schoenebeck's technique

Problem 2: Generalize Schoenebeck's technique

- Technique seems specialized for linear equations.
- Breaks down even if there are few local contradictions (which doesn't rule out a gap).

Problem 2: Generalize Schoenebeck's technique

- Technique seems specialized for linear equations.
- Breaks down even if there are few local contradictions (which doesn't rule out a gap).
- We have distributions, but not vectors for other type of CSPs.

Problem 2: Generalize Schoenebeck's technique

- Technique seems specialized for linear equations.
- Breaks down even if there are few local contradictions (which doesn't rule out a gap).
- We have distributions, but not vectors for other type of CSPs.
- What extra constraints do vectors capture?

Thank You

Questions?