Maximum Stable Set Formulations and Heuristics Based on Continuous Optimization

Sam Burer (Georgia Tech)

Renato Monteiro (Georgia Tech)

Yin Zhang (Rice University)

University of Waterloo

April 27, 2001

Outline

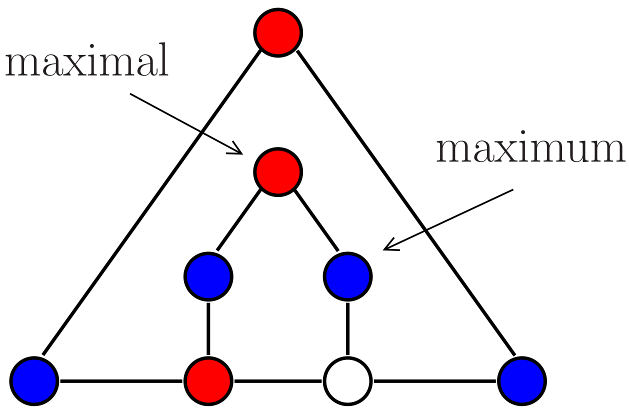
- The maximum stable set problem
- Formulation as rank-restricted SDP
- Local and global maxima, saddle points
- Finding high-quality stable sets
- Escaping to even better stable sets
- Computational results
- Conclusions and future Directions

The Maximum Stable Set Problem

Given: graph $G = (V, E)$

Defn: $S \subseteq V$ is a stable set if all its vertices are mutually nonadjacent

Goal: find a stable set S of maximum size α



- $V = \{1, \ldots, n\}$
- MSS is NP-Hard
- Equivalently known as
  - maximum independent set problem on G
  - minimum vertex cover problem on G
  - maximum clique problem on $\bar{G}$

Restricting the Lovász SDP

$$\begin{aligned} \alpha_r = \quad \max \quad & (ee^T) \bullet X \\ \text{s.t.} \quad & \text{trace}(X) = 1 \\ & X_{ij} = 0 \;\; \forall \;\; (i,j) \in E \\ & \text{rank}(X) \leq r \\ & X \succeq 0 \end{aligned}$$

$$\alpha \;\leq\; \alpha_1 \;\leq\; \alpha_2 \;\leq\; \cdots \;\leq\; \alpha_n \;=\; \vartheta$$

$$\alpha \;=\; \alpha_1 \;=\; \alpha_2$$

Rank-1 and Rank-2 Restrictions

$$X \text{ psd, rank-1} \quad \Longleftrightarrow \quad X = xx^T$$

$$\updownarrow$$

$$\begin{aligned} \alpha_1 = \quad \max \quad & (e^T x)^2 \\ \text{s.t.} \quad & \|x\|^2 = 1 \\ & x_i x_j = 0 \ \ \forall \ \ (i,j) \in E \end{aligned}$$

$$X \text{ psd, rank-2} \quad \Longleftrightarrow \quad X = xx^T + yy^T$$

$$\updownarrow$$

$$\begin{aligned} \alpha_2 = \quad \max \quad & (e^T x)^2 + (e^T y)^2 \\ \text{s.t.} \quad & \|x\|^2 + \|y\|^2 = 1 \\ & x_i x_j + y_i y_j = 0 \ \ \forall \ \ (i,j) \in E \end{aligned}$$

abbreviation: ($\mathcal{R}_1$) denotes the rank-1 problem

Lemma 1. Let x be feasible for ($\mathcal{R}_1$), and define

$$S = G \setminus \{i \in V : x_i = 0\}.$$

Then S is a stable set "induced by x" and

$$(e^T x)^2 \le |S|. \tag{1}$$

Theorem 1. The optimal value of ($\mathcal{R}_1$) equals α, i.e., $\alpha_1 = \alpha$. Moreover, for a feasible x that induces S:

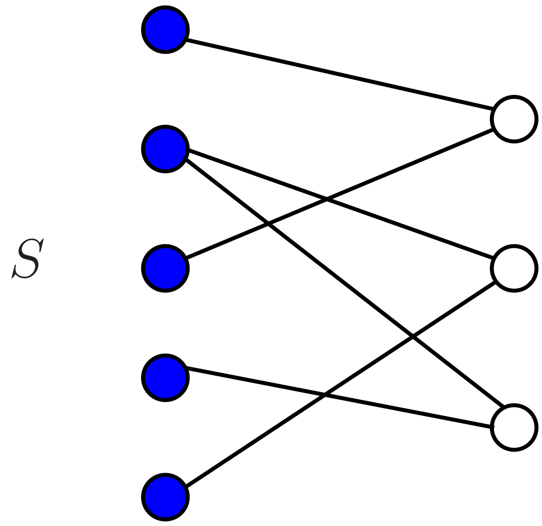
(i) x is a global max iff S is a maximum stable set and equality holds in (1);

(ii) x is a local max iff S is a maximal stable set and equality holds in (1).

Lemma 2a. Let (x, y) be feasible for $(\mathcal{R}_2)$, and define

$$\tilde{G} = G \setminus \{i : x_i = y_i = 0\}.$$

Then $\tilde{G}$ is bipartite, from which S is induced.



Lemma 2b. Let (x, y) be feasible for $(\mathcal{R}_2)$, and suppose (x, y) induces the stable set S. Then

$$(e^T x)^2 + (e^T y)^2 \leq |S|. \tag{2}$$

Theorem 2. The optimal value of $(\mathcal{R}_2)$ equals α, i.e., $\alpha_2 = \alpha_1 = \alpha$. Moreover, for a feasible (x, y) that induces S:

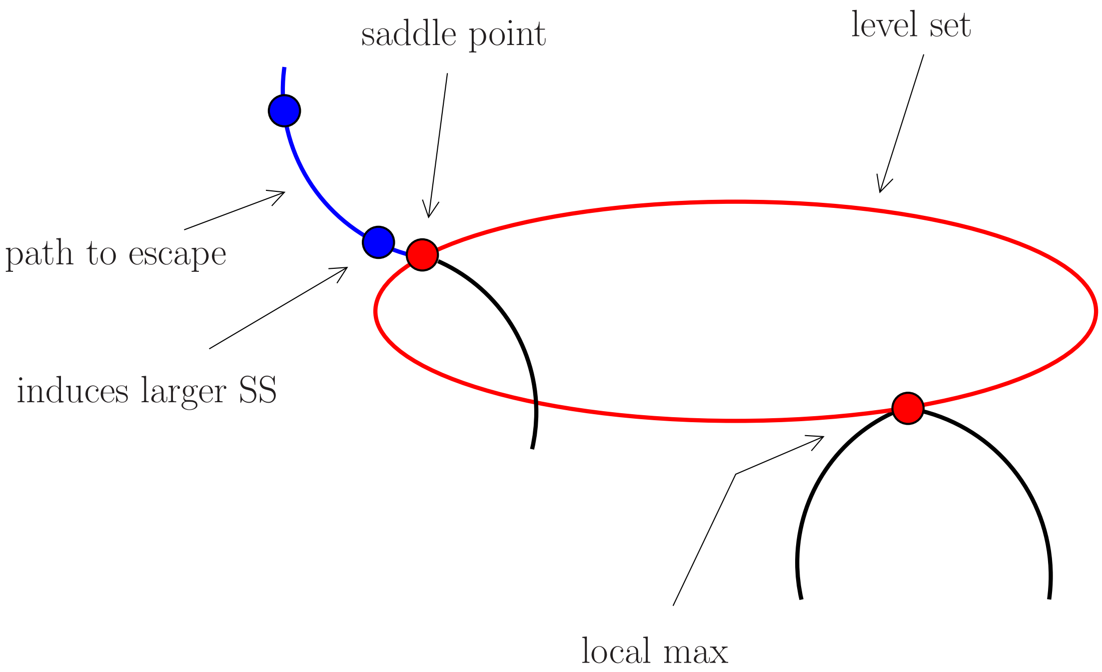
(i) (x, y) is a global max iff S is a maximum stable set and equality holds in (2);

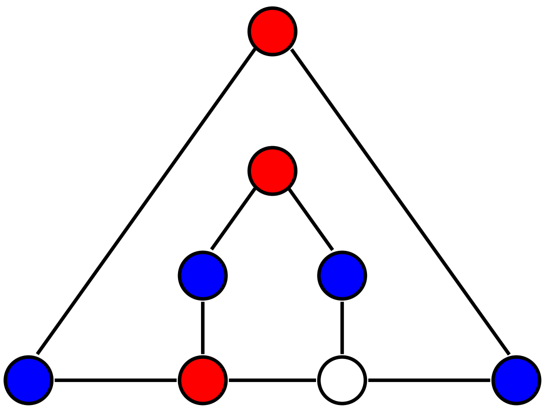
(ii) if (x, y) is a local max, then equality holds in (2).

Remark. In (ii), S need not be maximal.

Remark. The converse of (ii) does not hold.

The fact that the converse of (ii) does not hold exhibits an interesting characteristic of $(\mathcal{R}_2)$...





- Know the saddle point but not the escape path
- Restart near a saddle point
- In practice, escaping works often but not always

Computational Results

Tested three specific heuristics:

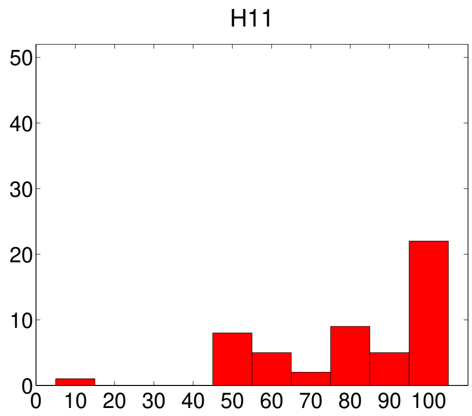
- H11 — one local max/stable set for $(\mathcal{R}_1)$
- H21 — one local max/stable set for $(\mathcal{R}_2)$
- H25 — five local max/stable sets for $(\mathcal{R}_2)$; employs the idea of escaping

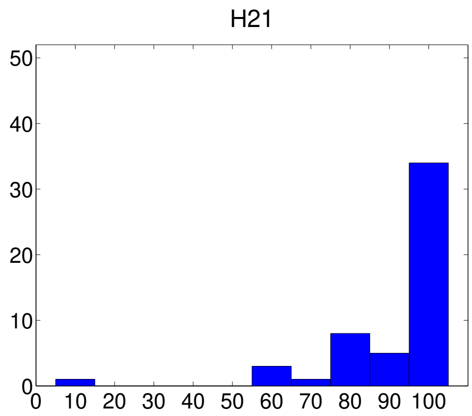
- First-order aug Lag algo to obtain local max
- Able to exploit graph sparsity
- Tested on SGI Origin2000 at 300MHz
- Test graphs are the complements of 64 graphs used in the Second DIMACS Implementation Challenge on the Maximum Clique Problem
- Each heuristic is randomized, and so each heuristic was run ten times on each graph

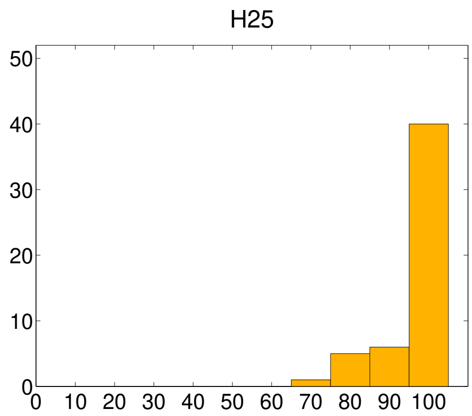
graph	$\lvert V\rvert$	$\lvert E\rvert$	dens %	α
MANN-a9.co	45	72	7.3	16
brock200-2.co	200	10024	50.4	12
brock400-3.co	400	20119	25.2	31
brock800-4.co	800	111957	35.0	26
c-fat200-1.co	200	18366	92.3	12
c-fat500-5.co	500	101559	81.4	64
hamming6-4.co	64	1312	65.1	4
hamming8-2.co	256	1024	3.1	128
hamming10-4.co	1024	89600	17.1	-1
johnson8-4-4.co	70	560	23.2	14
johnson16-2-4.co	120	1680	23.5	8
johnson32-2-4.co	496	14880	12.1	16
keller5.co	776	74710	24.8	27
p-hat300-1.co	300	33917	75.6	8
p-hat500-2.co	500	61804	49.5	36
p-hat700-2.co	700	122922	50.2	-1
p-hat1000-1.co	1000	377247	75.5	-1
p-hat1500-1.co	1500	839327	74.7	-1
san200-0.7-1.co	200	5970	30.0	30
san200-0.9-3.co	200	1990	10.0	44
san400-0.7-1.co	400	23940	30.0	40
san400-0.9-1.co	400	7980	10.0	100
san1000.co	1000	249000	49.8	15
sanr200-0.7.co	200	6032	30.3	18
sanr400-0.5.co	400	39816	49.9	13

Distribution of 52 graphs into "scores."

For a given graph, the score k is achieved if a stable set of size $k\%$ of α is found *over all ten runs*.

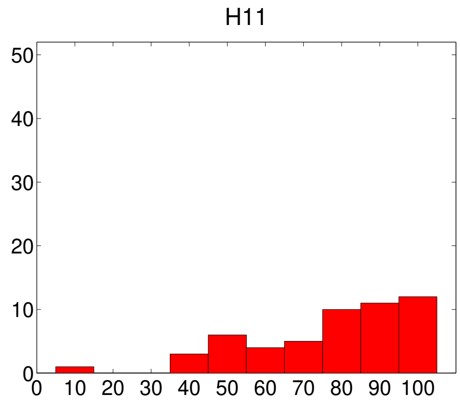


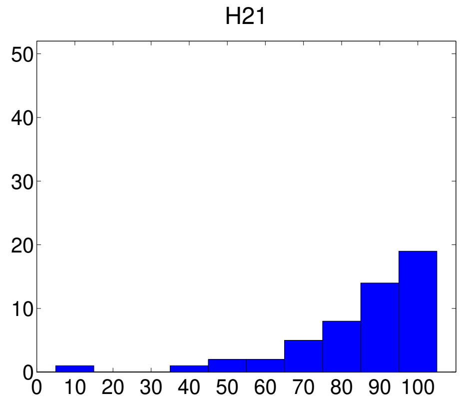


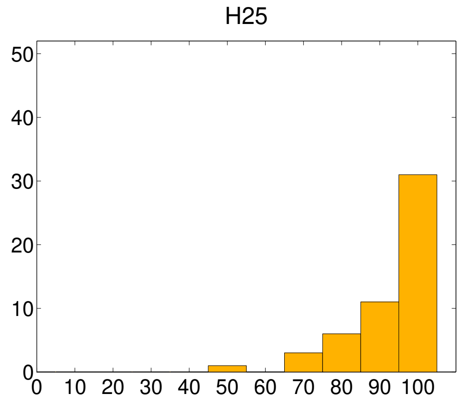


Distribution of 52 graphs into "scores."

For a given graph, the score k is achieved if a stable set of size $k\%$ of α is found *on average*.







Average total time for all 64 graphs:

H11	1960.9 sec
H21	12597.8 sec
H25	35697.1 sec

For each graph...

- Stable set found over all ten runs:

$$\text{H11} \leq \text{H21} \leq \text{H25}$$

- Stable set found on average:

$$\text{H11} \leq \text{H21} \leq \text{H25}$$

- Average time (except for two graphs):

$$\text{H11} \leq \text{H21} \leq \text{H25}$$

Conclusions and Future Directions

- New formulations of MSS
- Interesting link between MSS and Lovász SDP
- Is there a better way to escape saddle points? (positive curvature)
- Can a local max be found more quickly?
- How about the max weighted SS problem?

- Ideas applicable to more NP-Hard problems?
- Advantages over discrete heuristics?