Strong relaxations and cutting planes for the Stable Set Problem

Stefano Smriglio

University of L'Aquila
stefano.smriglio@univaq.it

joint work with
Monia Giandomenico
Adam N. Letchford
Fabrizio Rossi

Vienna, January 2012

Outline

- Part I: review of known results
  - Motivation
  - Linear programming relaxations
  - Semidefinite relaxations
- Part II: a non-standard lift-and-project relaxation
  - The Lovász-Schrijver $M(K, K)$ lifting operator
  - Projection by Benders decomposition
  - Cutting planes
- Part III: a new convex quadratic programming relaxation
  - A nice Ellipsoid
  - Separation
  - Cutting planes

The stable set problem (SSP)

$G = (V, E)$ simple undirected graph

A vertex set $S \subseteq V$ is called *stable* if the vertices in S are pairwise non-adjacent

stable set problem: determine a stable set of *maximum cardinality* (*weight*, if a weight vector $w \in \mathbb{Q}_+^{|V|}$ is given)

(equivalent to the *max-clique* problem)

A difficult case for IP

- "Unstructured stable set problems appear to be very difficult integer programming problems for LP-based branch-and-bound algorithms" (George Nemhauser '92)
- In fact, graphs with $\simeq 300$ vertices can be hard to solve to optimality (compare to the TSP!).
- Challenge/opportunity for integer programmers to develop new methods to overcome the difficulties experienced by the traditional ones.
- We'll start from using as much structure as we can (polyhedral combinatorics) and move towards methods which completely disdain structure

Edge formulation

$$
\begin{aligned}
\alpha_w(G) = \max \quad & \sum_{i \in V} w_i x_i \\
\text{s.t.} \quad & x_i + x_j \le 1 \quad \forall \{i,j\} \in E \\
& x_i \in \{0,1\} \quad \forall i \in V
\end{aligned}
\tag{1}
$$

Fractional Stable Set Polytope

$\mathrm{FRAC}(G) = \{x \in \mathbb{R}^{|V|} : x_i \ge 0, \forall i \in V$ and (1) hold $\}$

Stable Set Polytope

$\mathrm{STAB}(G)$ denotes the convex hull of the incidence vectors of stable sets in G.

The template paradigm for finding cuts

- a *template* is a set $\mathcal{C}$ of linear inequalities valid for $\mathrm{STAB}(G)$ with a combinatorial description
- a *separation algorithm* for $\mathcal{C}$ is an algorithm that, given any point x^*, returns either an inequality in $\mathcal{C}$ that is violated by x^* or a failure message
- the separation algorithm is *exact* if returns a failure message only if all inequalities in $\mathcal{C}$ are satisfied by x^*; otherwise it is *heuristic*
- the *template paradigm*:
  - (i) describe one or more templates
  - (ii) for each template design an "efficient" separation algorithm
- facet-inducing cuts typically play an important role

Clique inequalities (Padberg'73)

For any maximal clique $C \subseteq V$ the inequality

$$\sum_{i \in C} x_i \leq 1 \qquad (2)$$

induces a facet of $\mathrm{STAB}(G)$

The associated separation problem is NP-hard in the strong sense

Denote by Ω the set of all maximal cliques of G and let

$$\mathrm{Q}\,STAB(G) = \{x \in \mathbb{R}^{|V|} : (2) \text{ holds } \forall C \in \Omega, x_i \geq 0, \forall i \in V\}$$

Clearly, we have:

$$\boxed{\mathrm{STAB}(G) \subseteq \mathrm{QSTAB}(G) \subseteq \mathrm{FRAC}(G)}$$

Separation heuristic

- effective heuristics (Hoffman & Padberg '93, Borndorfer '99). Basic principle: pick a vertex v and greedily look for violated cliques in its neighbourhood
- Experiment:
  - start with a set of cliques covering all edges;
  - run a cutting plane algorithm embedding the heuristic and stop when it fails (violation tolerance $1E^{-9}$) $\Rightarrow UB_{CL}$
  - go on with an exact separation algorithm and stop when it fails (violation tolerance $1E^{-9}$) $\Rightarrow UB_{QSTAB}$

Clique inequalities: heuristic vs. exact separation

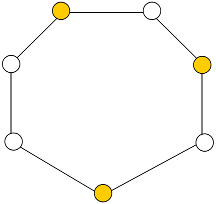
Graph	$\|V\|$	Density	$\alpha(G)$	UB_{CL}	% Gap	# cuts	UB QSTAB	% Gap closed	# add. cuts
brock200_1	200	0.25	21	38.03	81.1	2,001	38.02	0.1	8
brock200_2	200	0.5	12	21.27	77.3	3,975	21.13	1.6	111
brock200_3	200	0.39	15	27.35	82.3	2,861	27.23	1.0	64
brock200_4	200	0.34	17	30.70	80.6	2,529	30.63	0.5	35
C.125.9	125	0.1	34	43.07	26.7	486	43.07	0.0	0
C.250.9	250	0.1	44	71.38	62.23	1,722	-	-	-
c-fat200-5	200	0.57	58	66.67	14.9	7,561	66.67	0.0	0
DSJC125.1	125	0.09	34	43.14	26.9	460	43.14	0.0	0
DSJC125.5	125	0.5	10	15.46	54.6	1,522	15.37	1.6	42
DSJC125.9	125	0.9	4	4.66	16.6	2,904	4.58	12.2	236
mann_a27	378	0.01	126	135.00	7.1	468	135.00	0.0	0
mann_a45	1035	0	345	360.00	4.3	1,320	360.00	0.0	0
hamming6-4	64	0.65	4	5.33	33.4	170	5.33	0.0	0
keller4	171	0.35	11	14.83	34.8	942	14.83	0.0	11
p_hat300_1	300	0.76	8	15.30	91.25	7,197	-	-	-
p_hat300_2	300	0.51	25	33.59	34.36	3,280	-	-	-
p_hat300_3	300	0.26	36	54.37	51.0	3,968	54.31	0.3	20
san200_0.7-2	200	0.3	18	19.17	6.5	1,489	18.65	44.4	26
sanr200_0.7	200	0.3	18	33.40	85.5	2,350	33.34	0.4	16
sanr200_0.9	200	0.1	42	59.82	42.4	1,170	59.82	0.0	0

Clique inequalities: branch-and-cut results

	CPLEX 11.2 default		Clique-B&C	
Graph	Time	Subprob.	Time	Subprob.
brock200_1	1,691.88	303,352	1,153.54	119,613
brock200_2	118.70	9,808	91.42	2,498
brock200_3	202.18	19,461	233.10	8,239
brock200_4	484.30	71,580	435.08	22,488
C.125.9	6.27	3,458	3.51	3,291
C.250.9	+++	+++	+++	+++
c-fat200-5	12.11	47	7.04	47
DSJC125.1	7.07	4,887	3.16	3,805
DSJC125.5	11.00	1,626	7.93	412
mann_a9	**0.02**	3	0.02	7
mann_a27	1.87	1,888	1.57	6,961
mann_a45	109.45	67,300	61.09	55,044
hamming6-4	0.11	6	0.09	5
keller4	23.40	5,814	27.86	2,713
p_hat300-1	136.23	4549	245.10	2736
p_hat300-2	210.02	5,773	94.31	1,345
p_hat300-3	+++	+++	4,040.84	79,090
san200_0.7-2	6.34	517	4.07	170
sanr200_0.7	1,008.83	161,673	689.58	34,250
sanr200_0.9	2,795.50	853,581	1,357.64	395,538

Odd hole and antihole inequalities (Padberg '73)

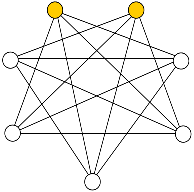
$H \subseteq V$ induces a simple (chordless) cycle of odd cardinality



$\sum_{i \in H} x_i \leq \left\lfloor \frac{|H|}{2} \right\rfloor$

Separation: polynomial (Gerards and Schrijver '86)

it's complement



$\sum_{i \in A} x_i \leq 2$

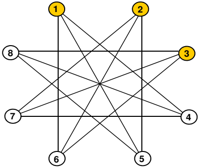
open

Web and antiweb Inequalities (Trotter '75)

Let $p, q \in \mathbb{Z}, p \geq 2q+1, q \geq 1$, $V = \{1, \dots p\}$.

- *Web*: $E = \{\{i,j\}, i \in V, j \in \{i+q, \dots, i-q\}\}$
- *Antiweb*: $E = \{\{i,j\}, i \in V, j \in \{i-q+1, \dots, i+q-1\}\}$

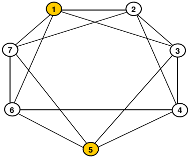
$W(8,3)$



$\sum_{i=1}^{p} x_i \leq q$

Separation: open

$AW(7,3)$



$\sum_{i=1}^{p} x_i \leq \lfloor p/q \rfloor$

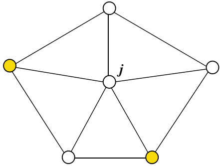
polynomial for fixed q
(Cheng and De Vries '02)

Lifting

- Odd hole, odd antihole, web and antiweb inequalities are not facet-inducing in general and need to be strengthened by lifting
- lifting requires the solution of a sequence of (small) weighted stable set problems (Nemhauser and Trotter '74, Padberg '75)
- In some cases the lifted inequalities can still be separated in polynomial time, e.g., the odd wheel inequalities (Grötschel, Lovász and Schrijver)

$$\sum_{i \in H} x_i + \left\lfloor \frac{|H|}{2} \right\rfloor x_j \leq \left\lfloor \frac{|H|}{2} \right\rfloor$$

$H \subseteq V$ induces an odd hole
$j \notin H$ is adjacent to all vertices in H.



Rank inequalities

$$\sum_{i \in W} x_i \leq \alpha(G[W]), \qquad W \subseteq V$$

- Separation problem NP-hard in the strong sense
- Not facet-inducing in general
- A characterization of rank facets for $\mathrm{STAB}(G)$ from facets of $\mathrm{STAB}(G')$, G' subgraph of G, is given by Balas and Zemel '77
- Effective project-and-lift separation heuristic (Rossi, S. '01)
  - violated rank inequalities are detected by separation of clique inequalities on a reduced graph
  - non-template: the structure of the inequality is not known a priori

Computational experiences

1992	2001	2009
Nemhauser, Sigismondi	Rossi, S.	Rebennack, Oswald, Theis, Seitz, Reinelt, Pardalos
clique, (lifted) odd-hole	general rank inequalities	clique, odd-hole, rank, mod k, local cuts
random $\|V\| \leq 120$	DIMACS $\|V\| \leq 700$ random $\|V\| \leq 200$	DIMACS $\|V\| \leq 500$ random $\|V\| \leq 150$

- rank inequalities (and lifted odd-holes) help especially on random graphs, but their contribution is often disappointing
- large enumeration trees are unavoidable to certify optimality (sparse graphs)

Quadratic formulation

non-convex quadratically-constrained formulation:

$$\begin{aligned} \max \quad & \textstyle\sum_{i \in V} x_i \\ \text{s.t.} \quad & x_i^2 - x_i = 0 \quad (i \in V) \\ & x_i x_j = 0 \quad (\{i,j\} \in E). \end{aligned}$$

To linearize, introduce the *matrix variable* $X = xx^T$, along with the augmented matrix

$$Y := \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$$

Y is *positive semidefinite*

The Lovász theta relaxation

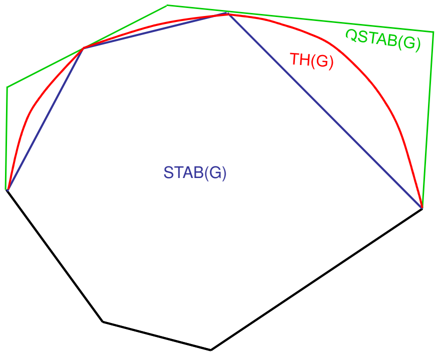
This leads to the SDP relaxation:

$$\begin{aligned} \theta(G) = \max \quad & tr(X) \\ \text{s.t.} \quad & Y_{0i} = Y_{ii} \quad (i \in V) \\ & Y_{ij} = 0 \quad (\{i,j\} \in E) \\ & Y \in \mathcal{S}^+_{|V|+1} \end{aligned}$$

where $\mathcal{S}^+_{|V|+1}$ is the cone of symmetric psd matrices of order $|V|+1$.

Theta Body $TH(G)$
projection of the feasible region onto the x-subspace

$TH(G)$ vs. $\mathrm{QSTAB}(G)$

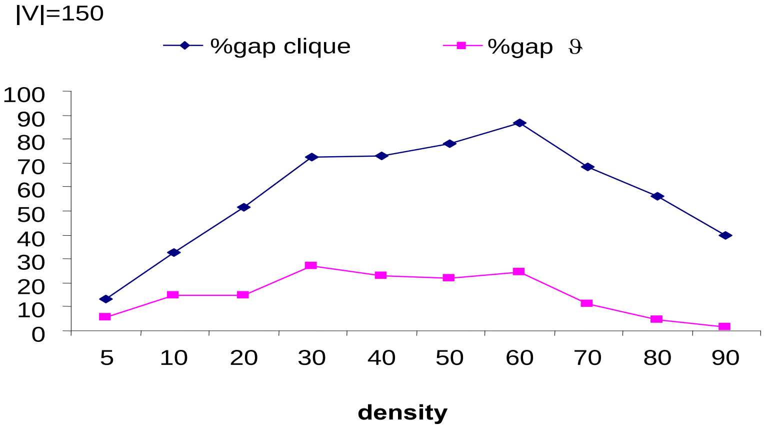


- $TH(G)$ convex but not polyhedral in general
- $\mathrm{STAB}(G) \subseteq TH(G) \subseteq \mathrm{QSTAB}(G)$ [equality iff G is perfect]

$\theta(G)$ vs. UB_{QSTAB}

Graph	$\|V\|$	$\|E\|$	$\alpha(G)$	UB_{QSTAB}	$\theta(G)$	$\frac{UB_{QSTAB}-\theta}{UB_{\mathrm{QSTAB}}-\alpha}\%$
brock200_1	200	0.25	21	38.02	27.5	61.81
brock200_2	200	0.50	12	21.13	14.22	75.68
brock200_3	200	0.39	15	27.23	18.82	68.77
brock200_4	200	0.34	17	30.63	21.29	68.53
C.125.9	125	0.10	34	43.07	37.89	57.11
C.250.9	250	0.10	44	71.38	56.24	55.30
c-fat200-5	200	0.57	58	66.67	60.34	73.01
DSJC125.1	125	0.09	34	43.14	38.39	51.97
DSJC125.5	125	0.50	10	15.37	11.47	72.63
DSJC125.9	125	0.90	4	4.58	4	100.00
mann_a9	45	0.01	16	135	17.47	98.76
mann_a27	378	0.00	126	360	132.76	97.11
hamming6-4	64	0.65	4	5.33	5.33	0.00
keller4	171	0.35	11	14.83	14.01	21.41
p_hat300_1	300	0.76	8	15.30	10.10	71.23
p_hat300_2	300	0.51	25	33.59	27.00	76.72
p_hat300_3	300	0.26	36	54.31	41.16	71.82
san200_0.7-2	200	0.30	18	18.65	18	100.00
sanr200_0.7	200	0.30	18	33.34	23.8	62.19
sanr200_0.9	200	0.10	42	59.82	49.3	59.03

$\theta(G)$ vs. UB_{QSTAB}



In practice

- $\theta(G)$ can be computed in polynomial time to arbitrary precision (Grötschel, Lovász and Schrijver '88)

- recent advances in practical implementations (Povh, Rendl and Wiegele '06; Malik Povh, Rendl and Wiegele '07)
- using $\theta(G)$ in branch-and-bound looks now viable (Wilson '09)
- numerical instability of SDP solvers and slow reoptimization still create problems
- when density $\geq 60\%$ clique inequalities are quite effective;
- the promising density range for θ-based methods is $[10, 40]$

Even stronger: the Lovász & Schrijver relaxation

$\mathrm{M}_+(\mathrm{FRAC(G)})$ equivalent to adding to the Lovász (SDP) relaxation some linear inequalities:

$$
\begin{aligned}
x_{ij} &\geq 0 && \{i,j\} \notin E \\
x_{ik} + x_{jk} &\leq x_k && \{i,j\} \in E, k \neq i,j \\
x_i + x_j + x_k &\leq 1 + x_{ik} + x_{jk} && \{i,j\} \in E, k \neq i,j
\end{aligned}
$$

Theorem (Lovász and Schrijver '91)

The projection $\mathrm{N}_+(\mathrm{FRAC}(G))$ *onto the non-quadratic space satisfies all clique, odd-hole, odd-antihole and odd-wheel inequalities.*

Theorem (Giandomenico and Letchford '06)

$\mathrm{N}_+(\mathrm{FRAC}(G))$ *satisfies also all web inequalities.*

Computational experiences

Burer & Vandenbussche '06:	$M_+(FRAC(G)$ Lagrangian method
Gruber & Rendl '03:	$TH(G)+$ odd-cycles + all triangle ineq. (stronger than $M_+(FRAC(G))$) Interior point cutting plane algorithm
Dukanovich & Rendl '07:	$TH(G)+$ some triangle ineq incomparable with $M_+(FRAC(G))$ Interior point exploiting some structure

optimizing over these relaxations is much harder than computing θ and requires highly specialised algorithms and often huge CPU times

SDP vs. LP lift-and-project relaxations

- Burer & Vandenbussche (2006). $M_+(FRAC(G))$ is substantially tighter than the Sherali-Adams (LP) relaxation $\mathrm{M}(FRAC(G))$:

Graph	$\alpha(G)$	UB_{clique}	$\mathrm{M}(FRAC(G))$	$\mathrm{M}_+(FRAC(G))$
brock200_2	12	21.53	66.66	17.08
brock200_4	17	30.84	66.66	22.84
c-fat200-5	58	66.67	66.66	58.17
mann_a9	16	18.50	18.00	17.17
hamming6-4	4	5.33	21.33	4.54
keller4	11	14.82	57.00	15.41
p_hat300_2	25	34.01	100.00	30.10
p_hat300_3	36	54.74	100.00	43.32
san200_0.7-2	18	21.14	66.66	20.01
sanr200_07	18	33.48	66.66	24.97
sanr200_09	42	60.04	66.66	49.31

- Balas, Ceria, Cornuejols and Pataki (1986). Lift-and-project cuts: including clique inequalities helps

To recap...

- clique inequalities are the only "plug-and-play" facet-inducing cuts while rank inequalities do not shutdown the gap
- these inequalities are sparse and well-managed by LP solvers
- overall, a state-of-the-art branch-and-cut based on clique inequalities is competitive
- standard lift-and-project methods are quite less effective than when applied to most $0-1$ problems
- the theta relaxation provides a nice compromise between tightness and computational tractability, but SDP solvers still create some difficulties
- using stronger relaxations does not look straightforward

Is it possible to "capture" the θ bound by linear programming?