Strong Lift-and-project cutting planes for the stable set problem

Stefano Smriglio

University of L'Aquila
stefano.smriglio@univaq.it
joint work with
Monia Giandomenico and Fabrizio Rossi

MIP 2010 - Atlanta

Outline

- Review of known relaxations:
  - LP in the natural space: *clique*, (lifted) *odd-hole, web, antiweb, rank* inequalities, etc.
  - SDP:
    - the Lovász $\theta(G)$ relaxation
    - the Lovász-Schrijver lifting operator $\mathrm{M}_+(.)$
  - LP in the extended (quadratic) space:
    - the Lovász-Schrijver $\mathrm{M}(K, K)$ operator
  - comparison among upper bounds
- $\mathrm{N}_j(K, K)$: A strong LP relaxation in the natural space (!)
  - theoretical strength of the closure $\cap_{j \in V} \mathrm{N}_j(K, K)$
  - practical implementation: Benders cuts
  - computational experience

Outline

- Review of known relaxations:
  - LP in the natural space: *clique*, (lifted) *odd-hole, web, antiweb, rank* inequalities, etc.
  - SDP:
    - the Lovász $\theta(G)$ relaxation
    - the Lovász-Schrijver lifting operator $\mathrm{M}_+(.)$
  - LP in the extended (quadratic) space:
    - the Lovász-Schrijver $\mathrm{M}(K, K)$ operator
  - comparison among upper bounds
- $\mathrm{N}_j(K, K)$: A strong LP relaxation in the natural space (!)
  - theoretical strength of the closure $\cap_{j \in V} \mathrm{N}_j(K, K)$
  - practical implementation: Benders cuts
  - computational experience

Notation

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

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

The *stable set problem* calls for a stable set of *maximum cardinality* (*weight*, if a weight vector $w \in \mathbb{Q}^+$ is given)

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

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

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

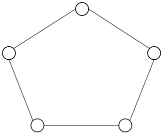
Valid inequalities

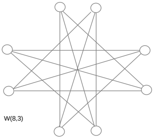
- clique inequalities: $\sum_{i \in C} x_i \leq 1$, $C \subseteq V$ induces a maximal clique

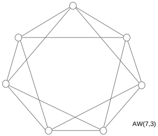
  $\mathrm{QSTAB}(G)$ polytope defined by clique and non-negativity inequalities

- rank inequalities: $\sum_{i \in W} x_i \leq \alpha(G[W])$, $W \subseteq V$

- special (structured) cases: W induces an *odd-hole, web, antiweb*, ...







- lifted versions (e.g. odd wheel inequalities)

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 $\|V\| \leq 150$

- clique inequalities are easy-to-manage and do close a relevant percentage of the gap
- effective separation heuristics for clique inequalities are important
- rank inequalities (and lifted odd-holes) can help especially on random graphs, but their contribution may be disappointing
- large enumeration trees are unavoidable to certify optimality (sparse graphs)

The Lovász theta relaxation

$$\begin{aligned} \theta(G) = \max \quad & \textstyle\sum_{i \in V} x_i \\ \text{s.t.} \quad & x_i = x_{ii} \qquad (i \in V) \\ & x_{ij} = 0 \qquad (\{i,j\} \in E) \\ & Y \in S_+^{n+1} \end{aligned}$$

where:

$$X = [x_{ij}]_{i=1,\ldots,n;j=1,\ldots,n}$$

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

S_+^{n+1} denotes the cone of real symmetric square **psd** matrices of order $n+1$

The Lovász theta relaxation

- polynomially solvable to arbitrary precision (Grötschel, Lovász and Schrijver, 1988)
- recent advances in practical implementations: Povh, Rendl, Wiegele '06; Malik, Povh, Rendl, Wiegele '07
- embedding $\theta(G)$ in branch-and-bound looks now viable (Wilson'09), even if reoptimization can still be too slow
- $\mathrm{STAB}(G) \subseteq TH(G) \subseteq \mathrm{QSTAB}(G)$ [equality iff G is perfect]
- $\theta(G)$ is very strong in practice: $[50-75]\%$ gap closed w.r.t. the bound reached by aggressive clique separation (DIMACS graphs $|V| \leq 300$)

Even stronger! The Lovász & Schrijver relaxation

The application of the Lovász & Schrijver M_+ operator to $\mathrm{FRAC}(G)$ is equivalent to adding to the Lovász (SDP) relaxation some linear inequalities:

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

$M_+(FRAC(G))$ computationally challenging (even if still polynomial): much harder than just computing θ (Burer & Vandenbussche, 2006)

Other attempts:

- Gruber & Rendl 03: add the ”remaining” triangle inequalities
- Dukanovich & Rendl ’07: weakened version

SDP vs. LP 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 (1996). Lift-and-project, L&S lifting operators: including clique inequalities help, while PSD cuts do not look convenient

The Lovász and Schrijver $\mathrm{M}(K, K)$ Operator

K given linear system in $0 - 1$ variables. For any pair of inequalities $\alpha x - \beta \geq 0$ and $\alpha^T x - \beta^T \geq 0$, compute the 'product inequality':

$$\begin{pmatrix} -\beta & \alpha^T \end{pmatrix} Y \begin{pmatrix} -\beta^T \\ \alpha^T \end{pmatrix} \geq 0.$$

where products $x_i x_j$, for $1 \leq i < j \leq n$, are replaced by variables x_{ij} and terms x_i^2, for $1 \leq i \leq n$, by x_i (binary)

$\Rightarrow$ an extended formulation provably stronger than the original (Lovász and Schrijver '91).

Apply to: (Ω set of all maximal cliques of G)

$$QSTAB(G) := \begin{cases} 1 - \sum_{i \in C} x_i \geq 0 & (C \in \Omega) \\ x_i \geq 0 & (i \in V) \end{cases}$$

The $\mathrm{M}(\mathrm{QSTAB}(G), \mathrm{QSTAB}(G))$ relaxation

$$
\begin{array}{rcl}
\max \sum_{i \in V} w_i x_i & & \\
\text{s.t.} & & \\
CVIs: & -x_i + \sum_{j \in C:\{i,j\} \in \bar{E}} x_{ij} \le 0 & (C \in \Omega, i \in V \setminus C) \\
CPIs: & \sum_{i \in C \cup C'} x_i - \sum_{\{i,j\} \in \bar{E}(C:C')} x_{ij} \le 1 & (C, C' \in \Omega) \\
 & x_{ij} = 0 & (\{i,j\} \in E) \\
 & x_{ij} \ge 0 & (\{i,j\} \in \bar{E})
\end{array}
$$

- **non-compact**: $n|\Omega|$ CVIs and $|\Omega|(|\Omega|-1)/2$ CPIs
- separation problems: both NP-hard (Giandomenico '06)

On the strength of $\mathrm{M}(K, K)$: theory

$\mathrm{N}(K, K)$ [$\mathrm{N}_+(K)$] projection of $\mathrm{M}(K, K)$ [M_+] onto the non-quadratic space.

Relaxation	implied inequalities
$TH(G)$ proj. of the θ relaxation convex but not polyhedral	clique [Grotschel, Lovász and Schrijver '88]
$\mathrm{N}_+(FRAC(G))$ polytope	clique, odd-cycle, odd-antihole and odd-wheel [Lovász and Schrijver '91] web [Giandomenico and Letchford '06]
$\mathrm{N}(QSTAB(G), QSTAB(G))$	web and antiweb and their lifted versions $\Rightarrow$ edge, clique, odd-hole, odd-antihole [Giandomenico, Letchford, Rossi, S. '07]

On the strength of $\mathrm{M}(K,K)$: practice

Graph	V	$\lvert E\rvert$	$\alpha(G)$	UB_{clique}	$\theta(G)$	UB_{MKK}	DR07	BV06 $\mathrm{M}_+(\mathrm{FRAC}(G))$
brock200_1	200	5,066	21	38.20	**27.5**	30.25	*	27.98
brock200_2	200	10,024	12	21.53	**14.22**	16.09	*	17.08
brock200_3	200	7,852	15	27.73	**18.82**	21.16	*	20.79
brock200_4	200	6,811	17	30.84	**21.29**	23.80	*	22.8
C.125.9	125	787	34	43.06	37.89	**36.53**	*	*
C.250.9	250	3,141	44	71.50	**56.24**	59.96	*	*
c-fat200-1	200	18,336	12	12.53	12	12	*	14.97
c-fat200-5	200	11,427	58	66.67	60.34	**58**	*	58.17
DSJC125.1	125	736	34	43.15	38.39	**36.99**	*	*
DSJC125.5	125	3,891	10	15.6	11.47	11.41	**11.4**	*
DSJC125.9	125	6,961	4	4.72	4.00	4	4.06	*
mann_a9	45	72	16	18.50	17.47	**16.85**	*	17.17
mann_a27	378	702	126	135.00	132.76	**131.39**	*	*
hamming6-4	64	1,312	4	5.33	5.33	4	4	4.54
keller4	171	5,100	11	14.82	14.01	**13.17**	*	15.41
p_hat300_1	300	33,917	8	15.68	**10.1**	11.40	*	18.66
p_hat300_2	300	22,922	25	34.01	**27**	30.00	*	30.1
p_hat300_3	300	11,460	36	54.74	**41.16**	47.32	*	43.32
san200_0.7-2	200	5,970	18	21.14	18	18	*	20.01
san200_0.9-3	200	1,990	44	45.13	44	44	*	44.4
sanr200_07	200	6,032	18	33.48	**23.8**	26.12	*	24.97
sanr200_09	200	2,037	42	60.04	**49.3**	50.73	*	49.31

Benders reformulation $\Rightarrow \mathrm{N}(K,K)$

	$\max \sum_{i \in V} w_i x_i$ s. t.		$\max \sum_{i \in V} w_i x_i$ s. t.
cliques	$Ax \leq 1$	$\rightarrow$	$Ax \leq 1$
CVIs + CPIs	$Bx + Dy \leq d$		$v^T(d - Bx) \geq 0, \; v \in \mathrm{RAY}$ $\mathrm{RAY} :=$ {extreme rays of $\pi^T D \geq 0, \pi \geq 0$}
	$x \in \mathbb{R}^n_+, y \in \mathbb{R}^p_+$		$x \in \mathbb{R}^n_+$

Clever selection of Benders cuts accomplished by a *Cut Generating Linear Program* (CGLP) - Fischetti, Salvagnin, Zanette '08:

$$\min v^T(d - Bx^*)$$
$$v^T D \geq \mathbf{0}$$
$$\sum_{i=1,\dots,p} v_i = \mathbf{1}$$
$$v \geq \mathbf{0}$$

Benders reformulation $\Rightarrow \mathrm{N}(K, K)$

	$\max \sum_{i \in V} w_i x_i$ s. t.		$\max \sum_{i \in V} w_i x_i$ s. t.
cliques	$Ax \leq 1$	$\rightarrow$	$Ax \leq 1$
CVIs + CPIs	$Bx + Dy \leq d$		$v^T(d - Bx) \geq 0, \; v \in \mathrm{RAY}$ $\mathrm{RAY} :=$ {extreme rays of $\pi^T D \geq 0, \pi \geq 0$}
	$x \in \mathbb{R}^n_+, y \in \mathbb{R}^p_+$		$x \in \mathbb{R}^n_+$

Clever selection of Benders cuts accomplished by a *Cut Generating Linear Program* (CGLP) - Fischetti, Salvagnin, Zanette '08:

$$\min v^T(d - Bx^*)$$
$$v^T D \geq \mathbf{0}$$
$$\sum_{i=1,\ldots,p} v_i = \mathbf{1}$$
$$v \geq \mathbf{0}$$

Cut Generating LP

i vertex, H, K cliques.
r = index of the (clique-vertex) pair Ki
s = index of the (clique-clique) pair HK

	$\cdots u_{Ki} \cdots$	$\cdots\cdots v_{HK} \cdots\cdots$		
	$\cdots x_i^* \cdots$	$\cdots (1 - x^*(H \cup K)) \cdots$		
	$\cdots\cdots\cdots$	$\cdots\cdots\cdots$	$\geq$	0
	$\vdots$	$\vdots$	$\vdots$	
$(i,j) \in \bar{E}$	$a_{(i,j),(K,i)}$	$\cdots b_{(i,j)(H,K)} \cdots$	$\geq$	0
	$\vdots$	$\vdots$	$\vdots$	
	$\cdots\cdots\cdots$	$\cdots\cdots\cdots$	$\geq$	0
	$\cdots 1 \cdots$	$\cdots 1 \cdots$	$=$	1

$a_{(i,j),(K,i)} = 1$ if $j \in K$ and 0 otherwise;
$b_{(i,j)(H,K)} = -1$ if $i \in H$ and $j \in K$ and 0 otherwise

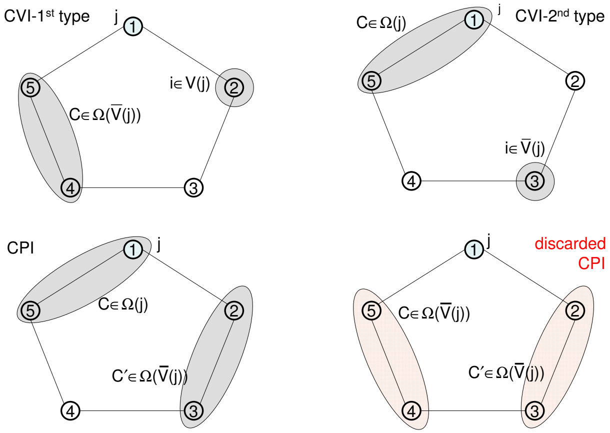
Cutting plane performance

- sparse graphs are tractable up to ~ 150 vertices (cut generation is too slow for $|V| = 200$)
- the integrality gap increases w.r.t. $M(K, K)$: the projection worsen about $[2, 10]\%$
- some cuts are sparse and ”clean”, but some other are quite dense
- sometimes cuts are helpful in branch-and-cut, but often confound it
- different normalization constraints may change things, but do not seen resolutive

Relaxing $\mathrm{M}(K, K)$

- pick a vertex j; $\Omega(j)$ maximal cliques containing j
- $\Omega(\bar{V}(j))$ maximal cliques with some vertex in $\bar{V}(j)$ (non adjacent to j)

$$V(1) = \{2, 5\},\ \bar{V}(1) = \{3, 4\},\ \Omega(1) = \{1, 2\}, \{1, 5\},\ \Omega(\bar{V}(1)) = \{2, 3\}, \{3, 4\}, \{4, 5\}$$



The $\mathrm{M}_j(K,K)$ relaxation

$$\max \sum_{i \in V} w_i x_i$$

s.t.

$$\begin{aligned}
-x_i + \sum_{j \in C : \{i,k\} \in \bar{E}} x_{ik} &\leq 0, & C \in \Omega(\bar{V}(j)), i \in j \cup V(j) \\
-x_i + \sum_{k \in C : \{i,k\} \in \bar{E}} x_{ik} &\leq 0, & C \in \Omega(j), i \in \bar{V}(j) \\
\sum_{i \in C \cup C'} x_i - \sum_{\{i,k\} \in \bar{E}(C:C')} x_{ik} &\leq 1, & C \in \Omega(j), C' \in \Omega(j) \cup \Omega(\bar{V}(j)) \\
x_{ik} &= 0, & \{i,k\} \in E \\
x_{ik} &\geq 0, & i \in j \cup V(j), k \in \bar{V}(j)
\end{aligned}$$

On the strength of $\mathrm{N}_j(K, K)$: theory

Relaxation	implied inequalities
$TH(G)$ proj. of the θ relaxation convex but not polyhedral	clique [Grotschel, Lovász and Schrijver '88]
$\mathrm{N}_+(FRAC(G))$ polytope	clique, odd-cycle, odd-antihole and odd-wheel [Lovász and Schrijver '91] web [Giandomenico and Letchford '06]
$\mathrm{N}(\mathrm{QSTAB}(G), \mathrm{QSTAB}(G))$	web and antiweb and their lifted versions $\Rightarrow$ edge, clique, odd-hole, odd-antihole [Giandomenico, Letchford, Rossi, S. '07]
$\cap_{j \in V} \mathrm{N}_j(\mathrm{QSTAB}(G), \mathrm{QSTAB}(G))$	odd-hole and antiweb and their lifted versions $\Rightarrow$ edge, clique, odd-antihole [Giandomenico, Rossi, S. '10] web? (we conjecture so)

Implementation

- (aggressive) clique separation heuristic
- $\Omega(j)$ filtered by slack: prefer cliques tight to the current fractional point
- **Cut generation (one round)** for each $i \in V$:

  STEP 1. build $\Omega(j), \Omega(\bar{V}(j)) \subseteq \Omega$, build CGLP
  STEP 2. Solve CGLP; add violated Benders cuts
  STEP 3. Run the clique separation heuristic;
  add violated clique cuts

- implemented in the IBM Cplex 11.2 framework; all default
- CGLPs solved by Cplex **hybaropt** option
- 2.4 GHz Pentium processor - 4 GB RAM

Upper bounds comparison

Graph	$\alpha(G)$	UB_{clique}	$\theta(G)$	UB $\mathrm{M}(K,K)$	time $\mathrm{M}(K,K)$	UB $\cap \mathrm{N}_j$	# cuts	time $\cap N_j$
brock200_1	21	38.20	27.50	30.25	17,670	33.71	240	405
brock200_2	12	21.53	14.22	16.09	26,501	18.27	25	343
brock200_3	15	27.73	18.82	21.16	22,386	23.57	64	413
brock200_4	17	30.84	21.29	23.80	25,362	26.77	144	446
C.125.9	34	43.06	**37.89**	36.53	227	**37.83**	624	506
C.250.9	44	71.50	56.24	59.96	9,397	63.95	500	8,908
c-fat200-5	58	66.67	60.34	58	265	58	68	45
DSJC125.1	34	43.15	**38.39**	36.99	274	**38.22**	623	335
DSJC125.5	10	15.6	11.47	11.41	377	13.93	99	27
mann_a27	126	135	**132.76**	131.39	393	**132.44**	378	79
mann_a45	345	360	358.99*	-	-	**355.79**	1,034	1,895
hamming6-4	4	5.33	**5.33**	4	4	**4.64**	123	31
keller4	11	14.82	14.01	13.82	15,324	14.55	171	14,073
p_hat300_2	25	34.01	27.00	30.00	24,337	31.55	128	958
p_hat300_3	36	54.74	41.16	47.32	46,408	49.79	407	2,419
san200_0.7-2	18	21.14	18	18	300	18	30	117
sanr200_07	18	33.48	23.8	26.12	9,971	29.45	297	762
sanr200_09	42	60.04	49.3	50.73	8,483	54.52	200	1254

$\cap_j N_j$ **outperforms** $\theta(G)$

Branch-and-cut-results

Graph	Cplex		Clique			$\cap_j N_j$		
	time	#sub	time	#sub	#clique cuts	time (sep. time)	#sub	#NKK cuts
brock200_1	1,784.78	270,169	1,474.88	91,543	1,492	**888.96** (15.36)	**60,248**	19
brock200_2	**83.16**	6,102	201.49	3,684	5,798	180.33 (25.77)	**1,875**	6
brock200_3	459.02	52,173	382.28	10,906	3,260	**303.19** (8.78)	**5,982**	9
brock200_4	735.52	85,134	724.14	24,835	2,584	**454.77** (10.74)	**13,381**	15
C125.9	5.63	3,049	4.16	2,357	192	**3.97** (1.39)	**2,287**	31
DSJC125.1	6.13	4,743	**5.11**	3,494	162	7.9 (3.3)	**2,640**	54
DSJC125.5	**6.97**	1,138	8.3	504	2,088	15.8 (6.7)	**436**	26
keller4	**21.88**	4,856	28.25	**2,730**	991	69.32 (41.72)	3,066	5
p_hat300-2	194.27	7,150	124.95	1,788	2,638	124.1 (7.3)	1,458	6
p_hat300-3	10,270.53	398,516	5,569.74	93,641	2,513	**3,469.52** (134.56)	**66,109**	38
sanr200_0.7	**827.52**	88,931	1,196.41	47,591	2,177	888.61 (37.19)	**30,965**	38
sanr200_0.9	2,650.55	635,496	1,529.2	355,867	366	**1,259.15** (30.46)	**267,708**	97
rand150_5	2.58	982.6	**2.15**	2,186.2	111.2	7.76 (5.62)	**798.6**	74.3
rand150_10	73.23	37,990.4	52.20	25,593.0	242.8	**48.13** (3.07)	**23,625.6**	81.3
rand150_20	108.90	31,646.4	**70.32**	12,284.0	584.60	82.67 (13.98)	**11,091.6**	95.0
rand200_5	705.03	320,921.0	545.76	320,362.4	202.0	**514.57** (17.09)	**186,450.3**	114.0
rand200_10	7,381.99	2,097,029.2	4,603.14	1,185,897.4	345.6	**3,160.56** (44.05)	**736,072.3**	120.6
rand200_20	4,261.68	670,879.4	2,598.74	246,003.4	1,007.4	**1,845.21** (207.94)	**137,387.6**	133.0

initial formulation: (good) collection of clique inequalities covering all edges
times are in seconds

random: average of 5 graphs