R U T C O R

R E S E A R C H

R E P O R T

BLOCK LINEAR MAJORANTS IN QUADRATIC 0-1 OPTIMIZATION

Endre Boros[a] **Isabella Lari**[b] **Bruno Simeone**[b]

RRR 18-2000, March 2000

RUTCOR
Rutgers Center for Operations
Research, Rutgers University
640 Bartholomew Road
Piscataway, NJ 08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/~rrr

[a] RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854-8003, USA. (Email: boros@rutcor.rutgers.edu)
[b] Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università di Roma "La Sapienza", Piazzale Aldo Moro 5, 00185 Rome, Italy
(Emails: lari@rosd.sta.uniroma1.it and marsalis@rosd.sta.uniroma1.it)

RUTCOR **R**ESEARCH **R**EPORT
RRR 18-2000, March 2000

BLOCK LINEAR MAJORANTS IN QUADRATIC 0-1 OPTIMIZATION

Endre Boros, Isabella Lari and Bruno Simeone

Abstract. A usual technique to generate upper bounds on the optimum of a quadratic 0-1 maximization problem is to consider a linear majorant (LM) of the quadratic objective function f and then solve the corresponding linear relaxation. Several papers have considered LM's obtained by termwise bounding, but the possibility of bounding *groups* of terms simultaneously does not appear to have been explored so far. In the present paper we develop this idea by suggesting the following approach: First, a suitable collection of "elementary" quadratic functions of few variables (typically, 3 or 4) is generated. All the coefficients of any such function (*block*) are either 1 or –1, and agree in sign with the corresponding coefficients of the given quadratic function. Next, for each block, a tightest LM (i.e., one having the same value as the block in as many points as possible), or a closest LM (i.e., one minimizing the sum of slacks) is computed. This can be accomplished through the solution of a small mixed-integer program, or a small linear program, respectively. Finally, the objective function is written as a weighted sum of blocks, with non-negative weights. Replacing in this expression each block by the corresponding LM, an LM of f can be obtained. We shall choose the weights in this process so that the maximum value of the resulting linear function is as small as possible. This amounts to a large-scale linear program, which can be solved by column generation. The encouraging results of a preliminary set of numerical tests are presented.

Keywords: Quadratic 0-1 optimization, Linear Majorants, Roof duality, Column generation

Acknowledgements: Part of this work was done during two visits of the first Author to the Department of Statistics, La Sapienza University, that were made possible by the support of the Italian Ministry of University and by the National Reserch Council of Italy; and during a visit of the third Author to RUTCOR and DIMACS (LSDO Special Year), Rutgers University, supported by these two centers. The first Author also gratefully acknowledges the partial support by the Office of Naval Research (Grant N00014-92-J-1375) and by the National Science Foundation (Grant DMS 98-06389).

1. Introduction

The present paper deals with the problem of finding tight upper bounds of the optimum value z^* of the *quadratic 0-1 maximization problem*

$$\max_{x \in \mathbf{B}^n} x^T Q x \qquad (1)$$

where Q is an $n \times n$ upper triangular real matrix with null diagonal entries, and $\mathbf{B} = \{0,1\}$. Such problems are known to be NP-hard (Garey and Johnson, 1979), in general. Thus, it makes sense to try to find good bounds of z^* at a reasonable computational cost.
A usual method to generate upper bounds of z^* is to consider a *linear majorant* (abbr. LM), that is, a linear function $c_0 + cx$ such that $c_0 + cx \geq x^T Q x$ for all $x \in \mathbf{B}^n$. Then the optimal value of the *linear relaxation* of (1),

$$c_0 + \max_{x \in \mathbf{B}^n} c x \qquad (2)$$

is an upper bound of z^*. The simplest way to obtain an LM is to majorize each individual quadratic term $q_{ij} x_i x_j$ $(i < j)$ by a linear function of the form $a_{ij} x_i + b_{ij} x_j + c_{ij}$, where the parameters a_{ij}, b_{ij}, c_{ij} must be chosen so that

$$c_{ij} \geq 0, \quad a_{ij} + c_{ij} \geq 0, \quad b_{ij} + c_{ij} \geq 0, \quad a_{ij} + b_{ij} + c_{ij} \geq q_{ij}, \qquad (3)$$

and then add up all such linear terms. The LM's obtained in this way are called *paved upper planes* in Hansen,Lu, and Simeone, 1990. An important subclass are the *roofs*, obtained through the majorization of each term $q_{ij} x_i x_j$ by either $\lambda_{ij} x_i + (q_{ij} - \lambda_{ij}) x_j$ or $\lambda_{ij}(1 - x_i - x_j)$, where $0 \leq \lambda_{ij} \leq |q_{ij}|$, depending on whether q_{ij} is positive or negative. Among the paved upper planes, the roofs have the property that they minimize the sum of the slacks in the inequalities (3). The *roof–dual* problem consists in finding a *best* roof, i.e., one that makes the optimal value of the corresponding linear relaxation as small as possible. The theory of roof-duality was set forth in Hammer, Hansen, and Simeone, 1981, and further carried on in Hammer, Hansen, and Simeone, 1984; Lu and Williams, 1985; Adams and Dearing, 1988; Boros and Hammer, 1989; Hammer and Kalantari, 1989; Hammer and Simeone, 1989; Boros, Crama, and Hammer, 1990; Boros, Hammer and Sun, 1991j; Boros, Crama, and Hammer, 1992; Boros and Hammer, 1993; Hansen, Lu, and Simeone, 1990; Bourjolly, Hammer, Pulleyblank, and Simeone, 1992. Several characterizations of the roof-dual optimum were given: in particular, it was shown in Hansen, Lu, and Simeone, 1990 , that it coincides with the smallest possible value of (2) when $c_0 + cx$ is a paved upper plane.
However, the possibility of bounding *groups* of terms, rather than individual terms, by an LM does not appear to have been explored so far. We demonstrate the potential of this idea by the following simple example. Consider the quadratic function of three binary variables

$$x_1 x_2 + x_1 x_3 - x_2 x_3 . \qquad (4)$$

It is easy to check that x_1 is an LM, but not a roof, of (4). As a matter of fact, x_1 turns out to be an extremely tight majorant of (4), since its values coincide with those of (4) in 6, out of 8, points of $\mathbf{B}^3$, and provides an upper bound of value 1 for the maximum of (4), which is actually also equal to 1. For the sake of comparison, the best roof $1 + 2x_1 - x_2 - x_3$ coincides with (4) only in 4 points, and yields the much less tight bound of value 3.

In the present paper we develop this idea by pursuing the following approach:

1) A suitable collection B of "elementary" quadratic functions (*blocks*) of few variables (typically, 3 or 4) is generated. We shall consider blocks with coefficients ±1 according to the signs of the corresponding coefficients of the given function $f(x) = x^T Q x$.
2) For each block in B a tightest LM (where "tightest" means "having the same value as the block in as many points as possible") or a closest one (where "closest" means "minimizing the sum of slacks") is generated. This can be accomplished through the solution of a small mixed-integer program, or a small linear program, respectively. Let us observe that we need to solve such a program only once for each of the "template" blocks.
3) Then, the objective function $f(x)$ is written as a weighted sum (with non-negative weights) of these blocks. Clearly, if in this expression each block is replaced by the corresponding tightest, or closest LM, an LM of $f(x)$ of the form $c_0 + cx$ is obtained. We shall then choose the weights so that the maximum value (2) of the resulting linear function is as small as possible. This optimization problem amounts to a large-scale linear program (abbr. LP), which can be solved, for example, by column generation.

The paper is structured as follows. In Section 2 an appropriate notational framework is established, and basic definitions are given. Section 3 deals with properties that "good" collections of blocks are required to have. Section 4 discusses some properties of tightest and closest LM's, and presents techniques for producing them; the section ends with a full catalogue of tightest LM's for all possible template blocks with up to 4 variables. Several conjectures about tightest and closest LM's are also presented. Section 5 deals with the problem of choosing optimal weights, as indicated above. A column generation algorithm for the ensuing LP is described. Finally, in Section 6 we report on the encouraging results of a preliminary numerical experimentation on 140 randomly generated test problems of up to 80 variables.

1. Basic notation and definitions

Let $f(x) = x^T Q x$ be the given objective function to be maximized over $\mathbf{B}^n$. One can associate with f a signed undirected graph $G \equiv G_f = (V, E)$ as follows: the vertex-set V is the standard set $\{1, \ldots, n\}$ and the edge-set E is the union of two sets E^+ and E^-, where

$$E^+ = \{(i, j) : 1 \le i < j \le n, \quad q_{ij} > 0\}, \qquad E^- = \{(i, j) : 1 \le i < j \le n, \quad q_{ij} < 0\}.$$

The unordered pair $\{i, j\}$ is simply denoted as ij and identified with the ordered pair (i, j), $i<j$.
We denote by $m = |E|$ the number of edges of G. The *sign function* of G (or of f) is the function $\sigma \equiv \sigma(G) \equiv \sigma_f : E \to \{1,-1\}$ defined by

$$\sigma_{ij} = \begin{cases} 1, & \text{if } ij \in E^- \\ -1, & \text{if } ij \in E^+ \end{cases}$$

A *signed automorphism* of G is any automorphism of G that preserves edge signs.
A *block* is a signed subgraph $B = (V(B), E(B))$ of G, with $E(B) = E^+(B) \cup E^-(B)$. If $x \in \mathbf{B}^n$ and B is any block, we denote by x_B the vector $(x_{i_1}, \ldots, x_{i_r})$, where $i_1, \ldots i_r$ are the elements of $V(B)$ in

increasing order. For a block B, the associated *block function* (often also called a block for short) is the quadratic pseudo-boolean function

$$f_B(x_B) = \sum_{ij\in E(B)} \sigma_{ij} x_i x_j \quad . \qquad (5)$$

Next, we define the useful notion of "template block". Given the block B with vertices $i_1 < ... < i_r$, its *template block* is $B^t = (V(B^t), E(B^t))$, where $V(B^t) = \{1, ..., r\}$ and $E(B^t) = \{\rho(i)\rho(j)\colon ij\in E(B)\}$, with $\rho(i_k) = k$ for all $k = 1, ..., r$. Notice that in general B^t is a block of K_n, the complete graph on n vertices, but not a block of G. A similar definition applies to block functions. For example, the common template block of the two blocks $x_1x_5 + x_1x_4 - x_4x_5$ and $x_3x_4 + x_3x_6 - x_4x_6$ is $x_1x_2 + x_1x_3 - x_2x_3$.

Now let $g(x) = c_0 + cx$ be a linear LM of $f(x)$. The *contact* of f and g is the set

$$Cont(f,g) = \{x\in \mathbf{B}^n : \; g(x) = f(x)\} \; . \qquad (6)$$

The *excess* of g w.r.t. f is defined by

$$exc(f,g) = \sum_{x\in \mathbf{B}^n} \left(g(x) - f(x)\right) \qquad (7)$$

Let M be a family of LM's of f. The LM $g^* \in \mathrm{M}$ is said to be

- *tightest* (in M) if $|Cont(f, g^*)| = \max\{\,|Cont(f, g)| : g\in \mathrm{M}\}$;
- *closest* (in M) if $exc(f, g^*) = \min\{exc(f, g) : g\in \mathrm{M}\}$;
- *best* (in M) if

$$\max_{x\in \mathbf{B}^n} g^*(x) = \min_{g\in \mathrm{M}} \; \max_{x\in \mathbf{B}^n} g(x) \; . \qquad (8)$$

(We assume that all the above optima exist. This is certainly true, for example, if M – regarded as a subset of $\mathbf{R}^{n+1}$ - is closed). In particular, let B be a block , and let M be the family of all LM's of f_B . If

$$g_B(x_B) = c_0 + \sum_{i\in V(B)} c_i x_i \qquad (9)$$

is an LM (tightest or closest LM) in M, and if α is a signed automorphism of B, then also $c_0 + \sum_{i\in V(B)} c_{\alpha(i)} x_i$ is an LM (tightest or closest LM). If Γ is the set of all signed automorphisms of B, and if $\gamma_i = \frac{1}{|\Gamma|}\sum_{\alpha\in\Gamma} c_{\alpha(i)}$ for all $i \in V(B)$, then $\hat{g}_B(x_B) = c_0 + \sum_{i\in V(B)} \gamma_i x_i$ is also an LM. Such a linear function is called the *symmetrization* of g_B . If $g_B(x_B)$ is a closest LM, then $\hat{g}_B(x_B)$ is also closest. A similar property does not hold, however, for tightest LM's. For example, a tightest linear majorant (or a TLM, in short) of the block $-x_1x_2 - x_1x_3 - x_3x_4 + x_1x_4 + x_2x_3$ is $2 + x_2 - 2x_3 - x_4$. The only non-identical signed automorphism of the block is $\begin{pmatrix} 1\,2\,3\,4 \\ 3\,4\,1\,2 \end{pmatrix}$. Hence $2 - 2x_1 - x_2 + x_4$ is also a TLM . However, the symmetrization $2 - x_1 - x_3$ is not a TLM.

2. Covering and exhaustive families of blocks

In the present section the question of choosing "good" collections B of blocks is addressed. Surely, in order to carry out the approach outlined in the Introduction, the chosen family B must have the property that the objective function f can be expressed as a weighted sum of the block functions f_B, $B \in \mathrm{B}$:

$$f(x) = \sum_{B \in \mathrm{B}} w_B\, f_B(x_B), \quad x \in \mathbf{B}^n \quad , \qquad (10)$$

where $w_B \geq 0$ for $B \in \mathrm{B}$. When this property holds the family B is said to be *exhaustive* for f. One obvious example of exhaustive family is the collection of all the edges of G, taken with their sign.
For each $ij \in E$, let $\mathrm{B}_{ij} = \{B \in \mathrm{B} : ij \in E(B)\}$. In view of (5), B is exhaustive for f if and only if the system of $\binom{n}{2}$ linear equations in the unknowns w_B

$$\sum_{B \in \mathrm{B}_{ij}} w_B = |q_{ij}| \quad , \; ij \in E \qquad (11)$$

has a non-negative solution.

Prop. 3.1: The family B is exhaustive for f if and only if there is no edge-weighting $\boldsymbol{\lambda}$ such that

$$\begin{cases} \sum_{ij \in E(B)} \lambda_{ij} \geq 0\,, \quad B \in \mathrm{B} \\ \sum_{ij \in E} \lambda_{ij}\, |q_{ij}| < 0 \end{cases} \qquad (12)$$

Proof: Following Farkas's Lemma, the system (11) has a non-negative solution if and only if there is no edge-weighting $\boldsymbol{\lambda}$ satisfying (12) .
We shall say that B is a *covering family* for G if $\bigcup_{B \in \mathrm{B}} E^+(B) = E^+$ and $\bigcup_{B \in \mathrm{B}} E^-(B) = E^-$, i.e., if every edge of G belongs to some block of B . Obviously, a necessary condition for B to be an exhaustive family for f is that B is a covering family for G_f.

Given a small positive integer p , one can use the following "skimming" procedure to obtain an exhaustive family of blocks with at most p vertices in each of them.

Skimming Algorithm

Step 1: Build the graph $G \equiv G_f$ associated with the input quadratic function $f(x) = x^T Q x$;
Let $H = G$;
Let $R = Q$;
Let $\mathsf{B} = \varnothing$;
Step 2: Find a connected block B of H, such that $|V(B)| \le p$;
Step 3: Compute $\delta = \min \{ |r_{ij}| : ij \in E(B) \}$;
Step 4: For each $ij \in E(B)$:
Replace r_{ij} by $r_{ij} - \delta$;
If the updated r_{ij} is zero, then delete edge ij from H ;
Step 5: Add B to B;
Set $w_B = \delta$;
Step 6: If $E(H) = \varnothing$ then **stop** : output B and w_B , for all $B \in \mathsf{B}$;
otherwise go to Step 2;
end

Notice that each block added to B must be different from all the subsequent ones, since it contains at least one edge that is missing from all of them.
The above algorithm runs in $O(m)$ time for any fixed p. As a matter of fact, the number of iterations (additions of one block) is at most m , since at each iteration at least one edge is deleted from the residual graph H . Moreover, identifying each connected block B in Step 2 takes constant time for a fixed p .

3. Tightest and closest linear majorants of blocks

In this section we discuss some properties of tightest and closest linear majorants of a given block function, as well as techniques for generating them. Without loss of generality, we may assume that the block B under consideration is a template one (see Sec.2), and is connected. Several conjectures about properties of tightest linear majorants and closest linear majorants (abbr. CLM's) will be presented. In fact, the validity of each conjecture has been verified for all blocks with up to 4 vertices. A full catalogue of the TLM's of all template block functions with up to 4 variables will also be given.
Since B is template, we can write $x_B = x = (x_1, \ldots x_p)$, where $p = |V(B)|$. Accordingly, the block function (5) will be denoted simply by $h(x)$ and a TLM of it by $g(x) = t_0 + t\,x$.
Let $N = 2^p$, and let $x^1, \ldots, x^N$ be the complete list of the points of $\mathbf{B}^p$. Let us introduce the (unbounded) polyhedron

$$P = \left\{ (t_0, t) \in \mathbf{R}^{p+1} : \; t_0 + t x^k \ge h(x^k), \, k = 1, \ldots, N \right\}. \tag{13}$$

For any $(t_0, t) \in P$, let $\mathsf{A}(t_0, t) = \left\{ k : \; t_0 + t x^k = h(x^k) \right\}$.

Now, it is immediate to see by the above definitions that

Prop. 4.1: If $g \in P$ is a nonnegative combination of $g_i \in P$ for $i=1,2,...$, then $\mathsf{A}(g) \subseteq \mathsf{A}(g_i)$ holds for all $i=1,2,....$.

Prop. 4.2: If $g, g' \in P$ are distinct vertices of P, then $\mathsf{A}(g) \neq \mathsf{A}(g')$.
Proof: Since the polyhedron P is clearly up-monotone and thus full-dimensional, the system of linear equations corresponding to $\mathsf{A}(g)$ is of full rank, and hence g is its unique solution. Similarly, g' is the unique solution of the full rank system of linear equations corresponding to $\mathsf{A}(g')$. Thus, $\mathsf{A}(g) \neq \mathsf{A}(g')$ follows by $g \neq g'$.

Corollary 4.3: If $g = t_0 + t\,x$ is a TLM of h, then g is an extreme point of P, and therefore $|Cont\,(h, g)| \geq p + 1$.

Conjecture 1: If g is a TLM of h and $|\,Cont\,(h, g)\,| = p + 1$, then g is a roof.

Prop. 4.4: If g is a roof of h such that $Cont\,(h, g) \neq \emptyset$, then
(i) there is a roof g' whose coefficients are integers and $Cont\,(h, g') = Cont\,(h, g)$;
(ii) there are no two points $x,y \in Cont\,(h, g)$ such that :
either $x_i = 1,\ x_j = 0$ *and* $y_i = 0,\ y_j = 1$ for some $ij \in E^+$,
or $x_i = 0,\ x_j = 0$ *and* $y_i = 1,\ y_j = 1$ for some $ij \in E^-$.
Proof: Since the roof

$$g(x) = \sum_{ij \in E^+} (\lambda_{ij}\, x_i + (1-\lambda_{ij})\, x_j) + \sum_{ij \in E^{-+}} \lambda_{ij}\,(1 - x_i - x_j)$$

where $0 \leq \lambda_{ij} \leq 1$ for all $ij \in E$, is obtained by termwise bounding, a point x belongs to $Cont\,(h, g)$ iff

$$\begin{aligned} x_i\, x_j &= \lambda_{ij}\, x_i + (1-\lambda_{ij})\, x_j\,, \quad ij \in E^+ \\ -x_i\, x_j &= \lambda_{ij}\,(1 - x_i - x_j)\,, \qquad ij \in E^- \end{aligned} \tag{14}$$

Hence if $ij \in E^+$

$x_i = 1,\ x_j = 0 \Rightarrow \lambda_{ij} = 0$
$x_i = 0,\ x_j = 1 \Rightarrow \lambda_{ij} = 1,$

and if $ij \in E^-$

$x_i = 0,\ x_j = 0 \Rightarrow \lambda_{ij} = 0$
$x_i = 1,\ x_j = 1 \Rightarrow \lambda_{ij} = 1.$

From these implications (ii) easily follows.

In all remaining cases the equalities (14) hold for arbitrary $0 \leq \lambda_{ij} \leq 1$, thus the contact does not change if one always chooses, say, $\lambda_{ij} = 1$. This proves (i).

Corollary 4.5: If some TLM of h is a roof, among the TLM's there is always a roof with integral coefficients.

Conjecture 2: There is always a TLM of h with integral coefficients.

Now we turn our attention to the generation of a TLM $t_0 + t\,x$ of the block h. This can be done by solving the following mixed-integer linear program:

$$\max \quad y_1 + \ldots + y_N$$

$$h(x^k) \le t_0 + t\,x^k \le h(x^k)\,y_k + M(1 - y_k)\ , \quad k = 1,\ldots,N \qquad (15)$$

$$(t_0\,,\,t) \in \mathbf{R}^{p+1}\,, \qquad y \in \mathbf{B}^N\,,$$

where M is a sufficiently large constant. In view of Cor. 4.3 and of Lemma 2.1 in Papadimitriou and Steiglitz 1982, one can choose $M = (p+1)(p+1)!\binom{p}{2}$, since $|\,h(x)\,|$ is bounded from above by $\binom{p}{2}$. The binary variable y_k is equal to 1 iff x^k belongs to the contact of h and $t_0 + t\,x$.

Notice that, since p is a small integer (typically, $p \le 4$), the size of the mixed-integer program (15) is small: for example, when $p = 4$ there are 32 constraints, 16 binary and 5 continous variables. Notice also that in practice we need to solve (15) only once for each of the template blocks with a small number of vertices. A full catalogue of TLM's for all template blocks with at most 4 vertices is given in Table 1, where the following symbols are used:

$\mathbf{K_n}$: complete graph with n vertices;
$\mathbf{K_{q,r}}$: complete bipartite graph with q and r vertices;
$\mathbf{P_n}$: path with n vertices;
$\mathbf{D_n}$: $\mathbf{K_n} \setminus e$, where e is an arbitrary edge;
$\mathbf{F_4}$: $\mathbf{K_{1,3}} \cup e$, where e is any nonexisting edge;

Table 1

BLOCKS WITH 3 VARIABLES

(**C** = number of contact points, **E** = excess of the TLM)

P_3 Blocks		**Tightest Linear Majorants**						
x_1x_2	x_2x_3	**const**	x_1	x_2	x_3	**C**	**E**	
-1	-1	2	-1	-2	-1	5	4	Roof
-1	1	1	-1	-1	1	5	4	Roof
1	1	0	0	2	0	5	4	Roof

K_3 Blocks			Tightest Linear Majorants						
x_1x_2	x_1x_3	x_2x_3	const	x_1	x_2	x_3	C	E	
-1	-1	-1	1	-1	-1	-1	6	2	Trunc. Roof
-1	1	1	0	0	0	1	6	2	Trunc.Roof
-1	1	-1	2	-1	-2	0	4	6	Roof
1	1	1	0	0	1	2	4	6	Roof

BLOCKS WITH 4 VARIABLES

P_4 Blocks			Tightest Linear Majorants							
x_1x_2	x_2x_3	x_3x_4	const	x_1	x_2	x_3	x_4	C	E	
-1	-1	-1	3	-1	-2	-2	-1	8	12	Roof
-1	-1	1	2	-1	-2	-1	1	8	12	Roof
-1	1	-1	1	-1	-1	1	0	8	12	Roof
-1	1	1	1	-1	-1	2	0	8	12	Roof
1	-1	1	1	1	-1	-1	1	8	12	Roof
1	1	1	0	1	0	2	0	8	12	Roof

$K_{1,3}$ Blocks			Tightest Linear Majorants							
x_1x_2	x_1x_3	x_1x_4	const	x_1	x_2	x_3	x_4	C	E	
-1	-1	-1	0	0	0	0	0	9	12	Roof
-1	-1	1	0	1	0	0	0	9	12	Roof
-1	1	1	0	2	0	0	0	9	12	Roof
1	1	1	0	3	0	0	0	9	12	Roof

F_4 Blocks				Tightest Linear Majorants							
x_1x_2	x_1x_3	x_1x_4	x_2x_3	const	x_1	x_2	x_3	x_4	C	E	
-1	-1	-1	-1	1	-1	-1	-1	0	9	8	Trunc.Roof
-1	-1	1	-1	1	-1	-1	-1	1	9	8	Trunc.Roof
-1	1	-1	1	0	0	0	1	0	9	8	Trunc.Roof
-1	1	1	1	0	0	0	1	1	9	8	Trunc.Roof
1	1	-1	-1	1	0	0	0	-1	9	8	Trunc.Roof
1	1	1	-1	0	1	0	0	1	9	8	Trunc.Roof
-1	-1	-1	1	0	0	0	1	0	7	16	Roof
-1	-1	1	1	0	1	1	0	0	7	16	Roof
-1	1	-1	-1	0	1	0	0	0	7	16	Roof
-1	1	1	-1	1	2	-1	-1	0	7	16	Roof

F_4 Blocks				Tightest Linear Majorants							
1	1	-1	1	0	2	1	0	0	7	16	Roof
1	1	1	1	0	3	0	1	0	7	16	Roof

$K_{2,2}$ Blocks				Tightest Linear Majorants							
x_1x_2	x_1x_4	x_2x_3	x_3x_4	const	x_1	x_2	x_3	x_4	C	E	
-1	1	-1	-1	1	0	-1	-1	0	8	8	Trunc. Roof
-1	1	1	1	0	0	0	1	1	8	8	Trunc. Roof
-1	-1	-1	-1	4	-2	-2	-2	-2	7	16	Roof
-1	1	-1	1	2	-1	-2	-1	2	7	16	Roof
-1	1	1	-1	1	-1	-1	1	1	7	16	Roof
1	1	1	1	0	0	2	0	2	7	16	Roof

D_4 Blocks					Tightest Linear Majorants							
x_1x_2	x_1x_3	x_1x_4	x_2x_3	x_3x_4	const	x_1	x_2	x_3	x_4	C	E	
-1	-1	1	-1	-1	2	0	-1	-2	-1	8	12	Trunc. Roof
-1	1	1	-1	-1	2	0	-2	-1	0	8	12	Trunc. Roof
-1	-1	1	1	1	1	-1	-1	1	1	8	12	Trunc. Roof
-1	1	1	1	1	0	0	0	1	2	8	12	Trunc. Roof
-1	-1	-1	-1	-1	3	-2	-2	-2	-1	7	12	Trunc. Roof
-1	-1	1	-1	1	2	-1	-2	-1	1	7	12	Trunc. Roof
-1	1	1	1	-1	0	0	0	1	1	7	12	Trunc. Roof
-1	1	-1	1	1	1	-1	-1	2	0	7	12	Trunc. Roof
1	-1	1	1	1	0	0	1	0	2	7	12	Trunc. Roof
-1	1	-1	-1	-1	4	-2	-2	-1	-2	6	20	Roof
-1	1	1	-1	1	2	-1	-2	0	2	6	20	Roof
-1	-1	1	1	-1	2	0	1	-2	-1	6	20	Roof
-1	-1	-1	1	1	3	-3	-1	1	-1	6	20	Roof
1	1	1	1	1	0	2	0	3	0	6	20	Roof

K4 Blocks						Tightest Linear Majorants							
x_1x_2	x_1x_3	x_1x_4	x_2x_3	x_2x_4	x_3x_4	const	x_1	x_2	x_3	x_4	C	E	
-1	-1	-1	-1	-1	-1	1	-1	-1	-1	-1	10	8	Dried Roof
-1	-1	1	-1	1	1	0	0	0	0	1	10	8	Dried Roof
-1	1	1	1	1	-1	0	1	1	0	0	10	8	Dried Roof
-1	-1	-1	-1	1	-1	3	-2	-2	-2	0	7	16	Trunc.Roof
-1	-1	1	-1	1	-1	3	-1	-1	-3	0	7	16	Trunc.Roof
-1	1	1	-1	1	-1	2	1	-1	-2	0	7	16	Trunc.Roof
-1	1	1	-1	1	1	0	2	0	0	1	7	16	Trunc.Roof
-1	1	1	1	1	1	0	0	0	1	3	7	16	Trunc.Roof
-1	1	-1	-1	1	-1	4	-2	-2	-1	-1	5	24	Roof
-1	1	1	-1	-1	1	3	-1	-3	0	1	5	24	Roof
1	1	1	1	1	1	0	1	0	2	3	5	24	Roof

It turns out that, at least for $p \le 4$, TLM's of blocks are related to roofs in a very simple way. We need two preliminary definitions. Given an LM $g(x) = c_0 + c\,x$ of the (template) block h, let $\Delta = \min \{g(x) - h(x) : x \in \mathbf{B}^p\}$. The *dried LM* $\tilde{g}(x) = (c_0 - \Delta) + c\,x$ is again an LM of h. If $c_0 - \lfloor c_0 \rfloor \le \Delta$, then the *truncated LM* $\hat{g}(x) = \lfloor c_0 \rfloor + c\,x$ is also an LM of h.
In the simplest case, $p = 2$, the TLM's of the block $x_i\,x_j$ are x_i and x_j, those of the block $-x_i\,x_j$ are the constant 0 and $x_i + x_j - 1$. In both cases the contact consists of 3 points out of 4, and the TLM's are roofs with parameters $\lambda_{ij} = 0$ or 1.

As a matter of fact, it turns out that all TLM's of template block functions with $p \le 4$ variables are dried roofs. Actually, in 33 cases out of 60 they happen to be roofs; in 24 cases out of 60 they are truncated roofs; in the remaining 3 cases they are dried roofs with $\Delta = 1$.
However, for $p = 5$ there are blocks whose TLM's are not dried roofs.

Conjecture 3: For any block h, one can obtain every TLM of h by taking a suitable convex combination of TLM's of sub-blocks of h, and then drying the resulting LM.

A closest linear majorant (CLM) $u_0 + u\,x$ of a template block h can be obtained through the solution of the following LP:

$$\begin{aligned} &\min\ s_1 + \dots s_N \\ &u_0 + u\,x^k - s_k = h(x^k), \quad k = 1, \dots, N \qquad (16) \\ &s_k \ge 0, \quad k = 1, \dots, N \\ &(u_0, u) \in \mathbf{R}^{p+1}, \ s \in \mathbf{R}^N \end{aligned}$$

Again, since p is typically a small integer, in practice the size of this LP is small.

Conjecture 4: Every TLM of a block is a CLM.

4. Finding optimal weights

Given an exhaustive family $\mathcal{B}$ for f, let us denote by w a non-negative $|\mathcal{B}|$-vector whose components w_B satisfy the equations (11), and let W denote the set of all such vectors w. For each $B \in \mathcal{B}$ let $t_0(B) + t(B)\, x_B$ be a "good" linear majorant (e.g., a TLM or a CLM) of the block function $f_B(x_B)$ **.** Then, for any $w \in W$, $g(w;x) = \sum_{B\in\mathcal{B}} w_B\,(t_0(B) + t(B)\,x_B)$ is an LM of f.

Our aim is to choose $w^* \in W$ such that $g(w^*; x)$ is best among all LM's $g(w; x)$, $w \in W$, i.e., the optimal value $L(w)$ of the corresponding linear relaxation (2) is as small as possible. Let $\mathcal{B}_i = \{B \in \mathcal{B} : i \in V(B)\}$, $i = 1, \dots, n$, and as customary let $z^+ = \max\{z, 0\}$, $z \in \mathbf{R}$. Then we have

$$L(w) = \sum_{B\in\mathcal{B}} w_B\, t_0(B) + \max_{x\in\mathbf{B}^n} \sum_{B\in\mathcal{B}} w_B \sum_{i\in V(B)} t_i(B)\, x_i = \sum_{B\in\mathcal{B}} w_B\, t_0(B) + \sum_{i=1}^{n} \left(\sum_{B\in\mathcal{B}_i} w_B\, t_i(B) \right)^+ . \quad (17)$$

Therefore, we can compute an optimal weight vector w^* by solving the LP

$$\min \quad \sum_{B\in\mathcal{B}} w_B\, t_0(B) + \sum_{i=1}^{n} u_i$$

(L)

$$u_i \ge \sum_{B\in\mathcal{B}_i} w_B\, t_i(B), \quad i = 1, \dots, n \quad (18)$$

$$\sum_{B\in\mathcal{B}_{ij}} w_B = |q_{ij}| \quad , \; ij \in E \quad (19)$$

$$w \ge 0, \; u \ge 0 .$$

First of all, let us consider two special cases of (L).

1. Let us choose $\mathcal{B}$ as the family of all edges with their appropriate sign. For each block $x_i x_j$ we take as LM either the TLM x_i or the TLM x_j; for each block $-x_i x_j$ we take as LM the TLM given by the constant 0. Then the optimal value of (L) is $\sum_{ij\in E^+} q_{ij}$, a rather weak bound.

2. Let us choose again $\mathcal{B}$ as above. Let

$$g^*(x) = \sum_{ij\in E^+} (\lambda_{ij}^*\, x_i + (1-\lambda_{ij}^*)\, x_j) + \sum_{ij\in E^{-+}} \lambda_{ij}^*\, (1 - x_i - x_j)$$

be a best roof, let $\bar{\lambda}_{ij} = \lambda_{ij}^* / |q_{ij}|$, $ij \in E$, and let us take

$\bar{\lambda}_{ij}\, x_i + (1-\bar{\lambda}_{ij})\, x_j$ as the LM of the block $x_i x_j$;

$\bar{\lambda}_{ij}\, (1 - x_i - x_j)$ as the LM of the block $-x_i x_j$ **.**

In this case the optimal value of (L) coincides with the roof-dual optimum.

Coming back to the general version of ($\mathcal{L}$), the total number of constraints is $n + m$, a manageable number, while the number of variables is $n + |\mathcal{B}|$, which may be quite large. Therefore, it is quite natural to think of a column generation approach to its solution.

We shall assume that the given family $\mathcal{B}$ always contains all the blocks consisting of single signed edges, and that the LM's corresponding to these blocks are defined as in 2. These choices ensure that:

(i) $\mathcal{B}$ is exhaustive;

(ii) The optimal value of ($\mathcal{L}$) is at least as good as the roof-dual optimum.

Given a feasible basis for ($\mathcal{L}$), let μ_i and π_{ij} be the associated multipliers of the constraints (18) and (19), respectively. Then, the reduced cost of the variable w_B, $B \in \mathcal{B}$ can be written as

$$\bar{c}(B) = t_0(B) + \sum_{i \in V(B)} t_i(B)\,\mu_i - \sum_{ij \in E(B)} \pi_{ij} \ . \qquad (20)$$

Let B^t be the template block of B, and let us denote by $B(i)$ the i-th smallest element of B, $i = 1, \ldots, n$. Then we can re-write (20) as

$$\bar{c}(B) = t_0(B^t) + \sum_{i=1}^{|V(B)|} t_i(B^t)\,\mu_{B(i)} - \sum_{ij \in E(B)} \pi_{ij} \ . \qquad (21)$$

In view of (21), only the LM's corresponding to template blocks are needed to be stored.

A formal description of the algorithm follows.

Column generation algorithm

Step 1: Let $k = 0$;

Initially, choose $\mathcal{B} = \mathcal{B}_0$, where $\mathcal{B}_0$ is the family of blocks defined as in 2.;

Solve the roof-dual of (1) to get a best roof $g^*(x)$;

For each block B in $\mathcal{B}_0$, get an LM $t_0(B) + t(B)\,x_B$ from $g^*(x)$ as explained in 2.;

Solve ($\mathcal{L}$) with $\mathcal{B}$ replaced by $\mathcal{B}_0$;

{As noticed above, the optimum z_0 of this initial LP is equal to the roof-dual optimum. }

Step 2: At the k-th iteration, let $\mathcal{B}_k \subseteq \mathcal{B}$ be the current family of blocks;

Solve ($\mathcal{L}$) with $\mathcal{B}$ replaced by $\mathcal{B}_k$;

Let (u^k, w^k) be an optimal solution; let μ_i^k and π_{ij}^k be the optimal multipliers associated with the constraints (18) and (19), respectively.

Step 3: For each $B \in \mathcal{B} \setminus \mathcal{B}_k$, compute the reduced cost of the variable w_B according to (21):

$$\bar{c}(B) = t_0(B^t) + \sum_{i=1}^{|V(B)|} t_i(B^t)\,\mu_{B(i)}^k - \sum_{ij \in E(B)} \pi_{ij}^k \ ;$$

compute $\bar{c}(B^*) = \min\left\{\bar{c}(B) : B \in \mathcal{B} \setminus \mathcal{B}_k\right\}$;

Step 4: If $\bar{c}(B^*) \geq 0$, then **stop**:

$$w_B^* = \begin{cases} w_B^k, & B \in \mathrm{B}_k \\ 0, & \text{else} \end{cases}$$

is an optimal weight vector; otherwise go to Step 5;

Step 5: Set $\mathrm{B}_{k+1} = \mathrm{B}_k \cup \{B^*\}$;

increase k by 1 and go to Step 2;

end

5. Computational experiments

The sharpness of the upper bounds given by the solution of problem (L) and the performance of the column generation procedure have been evaluated in a very preliminary set of 140 test problems whose number of variables is betwen 20 and 80. Each experiment is defined by a pair:

TEST = (n, *density*)

where $n \in \{20, 80\}$ and $density \in (19, 70)$.

Here, n is the number of variables and *density* is the density of G, that is, $density = p$ means that in the test problems the average number of edges of G is p% of $\binom{n}{2}$.

For each fixed TEST, 10 test problems were randomly generated with q_{ij} in the interval (-10,10), for each $(i,j) \in E$.

We have implemented the column generation procedure using CPLEX via AMPL.

In order to obtain a family of blocks B containing a reasonable number of blocks we made use of the following procedure:

0. let d_{ij} = (degree of vertex i)*(degree of vertex j) for each $(i,j) \in E$ and let α be a real number in the interval (0,1) ;
1. let $\mathrm{B} = \varnothing$;
2. for the types of blocks K_4, D_4, $K_{2,2}$, F_4 and K_3 (in this order) do

   selecting the blocks in decreasing order of number of contact points of the TLM (see Table 1), add to B a signed subgraph B of G only if each edge ij of B is contained in less than $\alpha\, d_{ij}$ blocks of B.

Notice that, since the column generation procedure starts from the exaustive family given by all the edges of G, we did not consider the blocks whose TLM is a roof. In the experiments we let $\alpha = 0.2$.

In a first set of experiments convergence was very slow. In order to improve on the convergence of the column generation procedure, in Step 5 we added to B_k a set B’ of blocks $B \in \mathsf{B} \setminus \mathsf{B}_k$ such that each edge of G belongs to at most one block of B’ and

$$\bar{c}\,(B) \le \delta\,\bar{c}\,(B^*) + (1-\delta)\bar{c}\,(B^{\max})$$

where $B^{\max}$ is a block of $\mathsf{B} \setminus \mathsf{B}_k$ with maximum negative reduced cost and $0 < \delta < 1$. In the experiments we let $\delta = 0.5$. When this method for updating the current set of columns was used, the final number of columns did not significantly increase and the convergence improved by a factor of 3 or 4 .

For a given test problem, let L^* be the bound obtained by the column generation procedure and let RD and Opt be the roof dual optimum and the optimum of (1), respectively. Moreover, let

$$E_{L^*} = \frac{L^* - Opt}{Opt}$$

$$E_{RD} = \frac{RD - Opt}{Opt}$$

$$E_{RD,L^*} = \frac{RD - L^*}{RD}$$

For each experiment we computed the average values, over the 10 test problems, of the following performance indicators:

- E_{L^*} , E_{RD} , E_{RD,L^*} ;
- the cardinality of B;
- the number of columns at the end of the procedure;
- the number of iterations of the column generation procedure;
- the running times.

In Table 2 the experimental results are shown. The upper bound given by the solution to problem (L) is always very close to the optimum and the (often substantial) improvement with respect to the roof dual value increases as *density* increases and, for fixed *density,* as n increases. On the other hand the number of iterations and the running times are large. Further research is needed in order to speed up the column generation algorithm through a suitable stabilization procedure (Du Merle et al., 1997). Moreover, experiments should be tried with smaller families of blocks. For example, one can decrease α or consider only K_4 or K_3 blocks or only the blocks with the maximum numbers of contact points.

Table 2

Number of vertices	Average density (%)	Running times (sec.)	Number of iterations	$\|B \setminus B_0\|$	$\|B_k \setminus B_0\|$ at the last iteration	E_{L^*}	E_{RD}	E_{RD,L^*}
20	19	1	1.1	12.2	0.1	0.000	0.002	0.001
20	35	3	2.6	102.7	4.5	0.001	0.021	0.018
20	48	14	9.8	241.2	40.9	0.002	0.226	0.180
20	57	23	14.2	430.3	61.6	0.012	0.272	0.188
20	65	34	19.1	664.4	86.1	0.021	0.443	0.279
40	19	3	1.7	298.7	10.8	0.000	0.025	0.024
40	34	73	19.1	1576.1	214.5	0.012	0.376	0.260
40	46	295	33.9	4201.8	433.6	0.027	0.580	0.342
40	56	836	53.9	7386.5	690.2	0.101	1.010	0.448
40	65	1092	48.0	9231.6	768.4	0.038	0.893	0.443
60	19	31	8.8	1529.1	143.6	n.a.	n.a.	0.173
60	34	2231	48.5	8262.4	961.6	n.a.	n.a.	0.381
60	47*	2457	38.7	4995.5	922.1	n.a.	n.a.	0.390
80	18*	235	18.7	862.2	433.9	n.a.	n.a.	0.264

* For the experiments with (n = 60, *density* = 47) and (n = 80 , *density* = 19) the parameter used in the column generation procedure has been changed as follows: the STOP condition was $\bar{c}(B^*) \geq 0.1$ and α = 0.05.

References

Adams, W.P. and P.M. Dearing (1988): ``On the equivalence between roof duality and Lagrangian duality for unconstrained 0-1 quadratic programming problems", Technical Report No. URI-061, Clemson University.

E. Boros, P.L. Hammer (1989): “A max-flow approach to improved roof-duality in quadratic 0-1 minimization”, Technical Report RRR 15-89, RUTCOR, Rutgers University

E. Boros, Y. Crama, P.L. Hammer (1990): “Upper bounds for quadratic 0-1 maximization”, *Operations Research Letters* 9, 73-79

E. Boros, Y. Crama and P.L. Hammer (1992): ``Chvátal cuts and odd cycle inequalities in quadratic 0-1 optimization", *SIAM Journal on Discrete Mathematics* 5, 163-177.

E. Boros and P.L. Hammer (1993): ``Cut-Polytopes, Boolean quadric polytopes and nonnegative quadratic pseudo-Boolean functions", *Mathematics of Operations Research* 18, 245-253.

E. Boros, P.L. Hammer and X. Sun (1991): ``Network flows and minimization of quadratic pseudo-Boolean functions," RUTCOR Research Report RRR 17-91, RUTCOR, Rutgers University.

J.-M. Bourjolly, P.L. Hammer, W.R. Pulleyblank, B. Simeone (1992): "Boolean-combinatorial bounding of maximum 2-satisfiability", in: O. Balci, R. Sharda, S.A. Zenios (eds.) : *Computer Science and Operations Research: New Developments in their Interfaces* (Pergamon Press, London) 23-42

O. Du Merle, D. Villeneuve, J. Desrosiers, P. Hansen (1997) : "Stabilisation dans la cadre de la génération de colonnes", Les Cahiers du GERAD, G-97-08

M.R. Garey, D.S. Johnson (1979): *Computers and Intractability: a Guide to the Theory of NP-completeness* (Freeman, S. Francisco)

P.L. Hammer, P. Hansen, B. Simeone (1981): "Upper planes of quadratic 0-1 functions and stability in graphs", in: O. Mangasarian, R.R. Meyer, S.M. Robinson (eds.): *Nonlinear Programming* 4 (Academic Press, New York) 395-414

P.L. Hammer, P. Hansen, B. Simeone (1984): "Roof-duality, complementation, and persistency in quadratic 0-1 optimization", *Mathematical Programming* 28, 121-155

P.L. Hammer, B. Kalantari (1989): "A bound on the roof-duality gap", in: B. Simeone (ed.): *Combinatorial Optimization, Lecture Notes in Mathematics*, vol. 1403 (Springer, Heidelberg)

P.L. Hammer , B. Simeone (1989): "Quadratic functions of binary variables", in: B. Simeone (ed.): *Combinatorial Optimization, Lecture Notes in Mathematics*, vol. 1403 (Springer, Heidelberg)

P. Hansen, S.H. Lu, B. Simeone (1990): "On the equivalence of paved duality and standard linearization in nonlinear 0-1 optimization", *Discrete Applied Mathematics* 29, 187-193

S.H. Lu, A.C. Williams (1985): "Roof-duality for nonlinear 0-1 programming", Technical Report RRR 2-85, RUTCOR, Rutgers University

C.H. Papadimitriou, K. Steiglitz (1982): *Combinatorial Optimization* (Prentice-Hall, Englewood Cliffs, NJ)