A class of valid inequalities for multilinear 0–1 optimization problems

Elisabeth Rodríguez-Heck and Yves Crama

QuantOM, HEC Management School, University of Liège
Partially supported by Belspo - IAP Project COMEX

EUROPT Workshop, Montréal, July 13, 2017
Multilinear 0-1 optimization

\[
\begin{align*}
\min & \quad \sum_{S \in S} a_S \prod_{i \in S} x_i + l(x) \\
\text{s. t.} & \quad x_i \in \{0, 1\} \quad i = 1, \ldots, n
\end{align*}
\]

- \( S \): subsets of \( \{1, \ldots, n\} \) with \( a_S \neq 0 \) and \( |S| \geq 2 \),
- \( l(x) \) linear part.
Standard Linearization (SL)

Standard Linearization

\[
\min \sum_{S \in S} a_S y_S + l(x) \\
\text{s. t. } y_S \leq x_i \quad \forall i \in S, \forall S \in S \\
y_S \geq \sum_{i \in S} x_i - (|S| - 1) \quad \forall S \in S \\
(y_S = \prod_{i \in S} x_i)
\]

- for variables \(x_i, y_S \in \{0, 1\}\), the convex hull of feasible solutions is \(P^*_SL\),
- for continous variables \(x_i, y_S \in [0, 1]\), the set of feasible solutions is \(P_{SL}\).

SL drawback: The continuous relaxation given by the SL is very weak!
The 2-link inequalities

Definition

For \( S, T \in \mathcal{S} \) and \( y_S, y_T \) such that \( y_S = \prod_{i \in S} x_i \), \( y_T = \prod_{i \in T} x_i \),

- the 2-link associated with \((S, T)\) is the linear inequality

\[
    y_S \leq y_T - \sum_{i \in T \setminus S} x_i + |T \setminus S|
\]

- \( P_{SL}^{2\text{links}} \) is the polytope defined by the SL inequalities and the 2-links.

Interpretation

\( y_S = 1 \Rightarrow \forall i \in S, x_i = 1 \)

\( y_T = 0 \) and \( y_S = 1 \Rightarrow \exists j \in T \setminus S, x_j = 0 \)
Theoretical contributions

Theorem 1: A complete description for the case of two monomials

For the case of two nonlinear monomials, $P_{SL}^* = P_{SL}^{2\text{links}}$, i.e., the standard linearization and the 2-links provide a complete description of $P_{SL}^*$. 
Proof idea (Theorem 1):

- Consider the extended formulation (with variables in \([0, 1]\))

\[
\begin{align*}
    y_{S \cap T} & \leq x_i, \quad \forall i \in S \cap T, \\
    y_{S \cap T} & \geq \sum_{i \in S \cap T} x_i - (|S \cap T| - 1), \\
    y_S & \leq y_{S \cap T}, \\
    y_S & \leq x_i, \quad \forall i \in S \setminus T, \\
    y_S & \geq \sum_{i \in S \setminus T} x_i + y_{S \cap T} - |S \setminus T|, \\
    y_T & \leq y_{S \cap T}, \\
    y_T & \leq x_i, \quad \forall i \in T \setminus S, \\
    y_T & \geq \sum_{i \in T \setminus S} x_i + y_{S \cap T} - |T \setminus S|.
\end{align*}
\]

- Notice that the two polytopes \(P^0\) and \(P^1\) obtained by fixing variable \(y_{S \cap T}\) to 0 and 1, resp., are integral.

- Compute \(\text{conv}(P^0 \cup P^1)\) using Balas (1974) and see that it is \(P_{SL}^{2\text{links}}\).
Theoretical contributions

Theorem 2: Facet-defining inequalities for the case of two monomials
For the case of two nonlinear monomials defined by $S$, $T$ with $|S \cap T| \geq 2$, the 2-links are facet-defining for $P^*_{SL}$.

Proof idea (Theorem 2): Since $P^*_{SL}$ is full-dimensional (dim $n + 2$), find $n + 1$ affinely independent points in the faces defined by the 2-links.
Computational experiments: are the 2-links helpful for the general case?

Objectives

- compare the **bounds** obtained when optimizing over $P_{SL}$ and $P_{SL}^{2\text{links}}$,
- compare the **computational performance** of **exact resolution methods**.

Software used: CPLEX 12.06.

Inequalities used

- SL: standard linearization (model),
- cplex: CPLEX automatic cuts,
- 2L: 2-links.
Problem definition
The 2-link inequalities
Perspectives

Random instances: bound improvement

LP relax. gap (%) fixed degree

Fixed degree:

<table>
<thead>
<tr>
<th>inst.</th>
<th>d</th>
<th>n</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>rf-a</td>
<td>3</td>
<td>400</td>
<td>800</td>
</tr>
<tr>
<td>rf-b</td>
<td>3</td>
<td>400</td>
<td>900</td>
</tr>
<tr>
<td>rf-c</td>
<td>3</td>
<td>600</td>
<td>1100</td>
</tr>
<tr>
<td>rf-d</td>
<td>3</td>
<td>600</td>
<td>1200</td>
</tr>
<tr>
<td>rf-e</td>
<td>4</td>
<td>400</td>
<td>550</td>
</tr>
<tr>
<td>rf-f</td>
<td>4</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>rf-g</td>
<td>4</td>
<td>600</td>
<td>750</td>
</tr>
<tr>
<td>rf-h</td>
<td>4</td>
<td>600</td>
<td>800</td>
</tr>
</tbody>
</table>
Random instances: computation times results

Run times (in sec.) fixed degree

Fixed degree:

<table>
<thead>
<tr>
<th>inst.</th>
<th>d</th>
<th>n</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>rf-a</td>
<td>3</td>
<td>400</td>
<td>800</td>
</tr>
<tr>
<td>rf-b</td>
<td>3</td>
<td>400</td>
<td>900</td>
</tr>
<tr>
<td>rf-c</td>
<td>3</td>
<td>600</td>
<td>1100</td>
</tr>
<tr>
<td>rf-d</td>
<td>3</td>
<td>600</td>
<td>1200</td>
</tr>
<tr>
<td>rf-e</td>
<td>4</td>
<td>400</td>
<td>550</td>
</tr>
<tr>
<td>rf-f</td>
<td>4</td>
<td>400</td>
<td>600</td>
</tr>
<tr>
<td>rf-g</td>
<td>4</td>
<td>600</td>
<td>750</td>
</tr>
<tr>
<td>rf-h</td>
<td>4</td>
<td>600</td>
<td>800</td>
</tr>
</tbody>
</table>
## Instances inspired from image restoration: definition

**Image restoration**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>n1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Base images:
- top left rect. (tl),
- centre rect. (cr),
- cross (cx).

### Perturbations:
- none (n),
- low (l),
- high (h).

Up to $n = 225$ variables and $m = 1598$ terms
Image restoration instances: bounds results

LP relax. gap (%) 10x15 images

SL SL&2L

tl-n tl-l tl-h cr-n cr-l cr-h cx-n cx-l cx-h

0 500 1,000 1,500 2,000
Image restoration instances: bounds results

LP relax. gap (%) 15x15 images

<table>
<thead>
<tr>
<th>SL</th>
<th>SL&amp;2L</th>
</tr>
</thead>
<tbody>
<tr>
<td>tl-n</td>
<td>cr-n</td>
</tr>
<tr>
<td>0</td>
<td>200</td>
</tr>
</tbody>
</table>

12/15
Image restoration instances: computation times results

Run times (in sec.) 10x15 images

- cx-h
- cx-l
- cx-n
- cr-h
- cr-l
- cr-n
- tl-h
- tl-l
- tl-n

- SL&cplex
- SL&cplex&2L
Image restoration instances: computation times results

Run times (in sec.) 15x15 images

- cx-h
- cx-l
- cx-n
- cr-h
- cr-l
- cr-n
- tl-h
- tl-l
- tl-n

SL&cplex
SL&cplex&2L

<table>
<thead>
<tr>
<th>Run times (in sec.)</th>
<th>100</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>15x15 images</td>
<td>14/15</td>
<td></td>
</tr>
</tbody>
</table>
Summary

- SL + 2-links = a complete description (two nonlinear monomials).
- 2-links help computationally for the general case.

Extensions and perspectives

- Characterization of when $P_{SL}$ has integer vertices (joint work with C. Buchheim, discovered independently by A. del Pia and A. Khajavirad).
- “3-link”, “4-link” inequalities... worth?
- Other reduction methods, to the quadratic case.
Some references I


Some references II


When is SL a complete description?

Summary

- SL + 2-links = a complete description (two nonlinear monomials).
- 2-links help computationally for the general case.

Question:
Can we characterize when the SL alone is a complete description of the convex hull $P_{SL}^*$?

Joint work with C. Buchheim.
Characterization independently discovered by A. Del Pia and A. Khajavirad.
Multilinear 0-1 optimization

\[
\begin{align*}
\min & \quad \sum_{S \in S} a_S \prod_{i \in S} x_i + l(x) \\
\text{s. t.} & \quad x_i \in \{0, 1\}, \quad i = 1, \ldots, n
\end{align*}
\]

Subsets \( S \) define a hypergraph \( H \). We write \( P_{SL} = P_{SL}^{(H)} \).

Matrix of constraints \( M_H \).

Standard linearization constraints

\[
\begin{align*}
\min & \quad \sum_{S \in S} a_S y_S + l(x) \\
\text{s. t.} & \quad y_S \leq x_i, \quad \forall i \in S, \forall S \in \mathcal{S} \\
& \quad y_S \geq \sum_{i \in S} x_i - (|S| - 1), \quad \forall S \in \mathcal{S}
\end{align*}
\]
Theorem 3
Given a hypergraph $H$, the following statements are equivalent:

(a) $P_{SL}^{(H)}$ is an integer polytope.
(b) $M_H$ is balanced.
(c) $H$ is Berge-acyclic.

Derived from a more general result taking into account the sign pattern of the monomials.
Theorem 4
Given a hypergraph $H = (V, E)$ and a sign pattern $s \in \{-1, 1\}^E$, the following statements are equivalent:

(a) For all $f \in \mathcal{P}(H)$ with sign pattern $s$, every vertex of $P_H$ maximizing $L_f$ is integer.

(b) $M_{H(s)}$ is balanced.

(c) $H(s)$ has no negative special cycle.

(d) $P_{H(s)}$ is an integer polytope.

$P_{H(s)}$ is defined by constraints

\begin{align*}
y_S & \leq x_i & \forall i \in S, \forall S \in S, \text{sgn}(a_S) = +1 \\
y_S & \geq \sum_{i \in S} x_i - (|S| - 1) & \forall S \in S, \text{sgn}(a_S) = -1
\end{align*}