

Linearization and quadratization approaches for non-linear 0-1 optimization

Elisabeth Rodriguez-Heck and Yves Crama

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

June 2015, TU Dortmund



Definitions

Definition: Pseudo-Boolean functions

A pseudo-Boolean function is a mapping $f : \{0, 1\}^n \rightarrow \mathbb{R}$.

Multilinear representation

Every pseudo-Boolean function f can be represented uniquely by a multilinear polynomial (Hammer, Rosenberg, Rudeanu [4]).

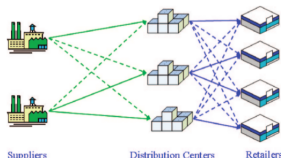
Example:

$$f(x_1, x_2, x_3) = 9x_1x_2x_3 + 8x_1x_2 - 6x_2x_3 + x_1 - 2x_2 + x_3$$

Applications



Computer vision: image restoration



Supply Chain Design with Stochastic Inventory Management (joint model of F. You, I. E. Grossman) [6]

Pseudo-Boolean Optimization

Many problems formulated as optimization of a pseudo-Boolean function

Pseudo-Boolean Optimization

$$\min_{x \in \{0,1\}^n} f(x)$$

- **Optimization is \mathcal{NP} -hard**, even if f is quadratic (MAX-2-SAT, MAX-CUT modelled by quadratic f).
- Approaches:
 - **Linearization**: standard approach to solve non-linear optimization.
 - **Quadratization**: Much progress has been done for the quadratic case (exact algorithms, heuristics, polyhedral results...).

Pseudo-Boolean Optimization

Many problems formulated as optimization of a pseudo-Boolean function

Pseudo-Boolean Optimization

$$\min_{x \in \{0,1\}^n} f(x)$$

- **Optimization is \mathcal{NP} -hard**, even if f is quadratic (MAX-2-SAT, MAX-CUT modelled by quadratic f).
- Approaches:
 - **Linearization**: standard approach to solve non-linear optimization.
 - **Quadratization**: Much progress has been done for the quadratic case (exact algorithms, heuristics, polyhedral results...).

Linearizations: reductions to the linear case

Standard linearization (SL)

$$\min_{\{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k,$$

$\mathcal{S} = \{S \subseteq \{1, \dots, n\} \mid a_S \neq 0\}$ (non-constant monomials)

1. Substitute monomials

$$\min \sum_{S \in \mathcal{S}} a_S z_S$$

$$\text{s.t. } z_S = \prod_{k \in S} x_k, \quad \forall S \in \mathcal{S}$$

$$z_S \in \{0, 1\}, \quad \forall S \in \mathcal{S}$$

$$x_k \in \{0, 1\}, \quad \forall k = 1, \dots, n$$

Standard linearization (SL)

$$\min_{\{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k,$$

$\mathcal{S} = \{S \subseteq \{1, \dots, n\} \mid a_S \neq 0\}$ (non-constant monomials)

1. Substitute monomials

$$\min \sum_{S \in \mathcal{S}} a_S z_S$$

$$\text{s.t. } z_S = \prod_{k \in S} x_k, \quad \forall S \in \mathcal{S}$$

$$z_S \in \{0, 1\}, \quad \forall S \in \mathcal{S}$$

$$x_k \in \{0, 1\}, \quad \forall k = 1, \dots, n$$

2. Linearize constraints

$$\min \sum_{S \in \mathcal{S}} a_S z_S$$

$$\text{s.t. } z_S \leq x_k, \quad \forall k \in S, \forall S \in \mathcal{S}$$

$$z_S \geq \sum_{k \in S} x_k - (|S| - 1), \quad \forall S \in \mathcal{S}$$

$$z_S \in \{0, 1\}, \quad \forall S \in \mathcal{S}$$

$$x_k \in \{0, 1\}, \quad \forall k = 1, \dots, n$$

Standard linearization (SL)

$$\min_{\{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k,$$

$\mathcal{S} = \{S \subseteq \{1, \dots, n\} \mid a_S \neq 0\}$ (non-constant monomials)

1. Substitute monomials

$$\min \sum_{S \in \mathcal{S}} a_S z_S$$

$$\text{s.t. } z_S = \prod_{k \in S} x_k, \quad \forall S \in \mathcal{S}$$

$$z_S \in \{0, 1\}, \quad \forall S \in \mathcal{S}$$

$$x_k \in \{0, 1\}, \quad \forall k = 1, \dots, n$$

2. Linearize constraints

$$\min \sum_{S \in \mathcal{S}} a_S z_S$$

$$\text{s.t. } z_S \leq x_k, \quad \forall k \in S, \forall S \in \mathcal{S}$$

$$z_S \geq \sum_{k \in S} x_k - (|S| - 1), \quad \forall S \in \mathcal{S}$$

$$z_S \in \{0, 1\}, \quad \forall S \in \mathcal{S}$$

$$x_k \in \{0, 1\}, \quad \forall k = 1, \dots, n$$

Standard linearization (SL)

$$\min_{\{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k,$$

$\mathcal{S} = \{S \subseteq \{1, \dots, n\} \mid a_S \neq 0\}$ (non-constant monomials)

1. Substitute monomials

$$\min \sum_{S \in \mathcal{S}} a_S z_S$$

$$\text{s.t. } z_S = \prod_{k \in S} x_k, \quad \forall S \in \mathcal{S}$$

$$z_S \in \{0, 1\}, \quad \forall S \in \mathcal{S}$$

$$x_k \in \{0, 1\}, \quad \forall k = 1, \dots, n$$

3. Linear relaxation

$$\min \sum_{S \in \mathcal{S}} a_S z_S$$

$$\text{s.t. } z_S \leq x_k, \quad \forall k \in S, \forall S \in \mathcal{S}$$

$$z_S \geq \sum_{k \in S} x_k - (|S| - 1), \quad \forall S \in \mathcal{S}$$

$$0 \leq z_S \leq 1, \quad \forall S \in \mathcal{S}$$

$$0 \leq x_k \leq 1, \quad \forall k = 1, \dots, n$$

Intermediate substitutions (IS) (one monomial)

SL substitution

$$z_S = \prod_{k \in S} x_k$$

SL linearization

$$\begin{aligned} z_S &\leq x_k, & \forall k \in S \\ z_S &\geq \sum_{k \in S} x_k - (|S| - 1) \end{aligned}$$

IS substitution

$$z_S = z_A \prod_{k \in S \setminus A} x_k$$

$$z_A = \prod_{k \in A} x_k$$

Intermediate substitutions (IS) (one monomial)

SL substitution

$$z_S = \prod_{k \in S} x_k$$

SL linearization

$$\begin{aligned} z_S &\leq x_k, & \forall k \in S \\ z_S &\geq \sum_{k \in S} x_k - (|S| - 1) \end{aligned}$$

IS substitution

$$\begin{aligned} z_S &= z_A \prod_{k \in S \setminus A} x_k \\ z_A &= \prod_{k \in A} x_k \end{aligned}$$

IS linearization

$$\begin{aligned} z_S &\leq x_k, & \forall k \in S \setminus A \\ z_S &\leq z_A, \\ z_S &\geq z_A + \sum_{k \in S \setminus A} x_k - |S \setminus A|, \\ z_A &\leq x_k, & \forall k \in A \\ z_A &\geq \sum_{k \in A} x_k - (|A| - 1). \end{aligned}$$

Intermediate substitutions (IS) (one monomial)

SL substitution

$$z_S = \prod_{k \in S} x_k$$

SL linearization

$$\begin{aligned} z_S &\leq x_k, & \forall k \in S \\ z_S &\geq \sum_{k \in S} x_k - (|S| - 1) \end{aligned}$$

IS substitution

$$\begin{aligned} z_S &= z_A \prod_{k \in S \setminus A} x_k \\ z_A &= \prod_{k \in A} x_k \end{aligned}$$

IS linearization

$$\begin{aligned} z_S &\leq x_k, & \forall k \in S \setminus A \\ z_S &\leq z_A, \\ z_S &\geq z_A + \sum_{k \in S \setminus A} x_k - |S \setminus A|, \\ z_A &\leq x_k, & \forall k \in A \\ z_A &\geq \sum_{k \in A} x_k - (|A| - 1). \end{aligned}$$

Intermediate Substitutions (IS) (one monomial)

Polytope $P_{SL,1} \subseteq \mathbb{R}^{n+1}$

$$z_S \leq x_k, \quad \forall k \in S$$

$$z_S \geq \sum_{k \in S} x_k - (|S| - 1)$$

$$0 \leq x_k \leq 1, \quad \forall k = 1, \dots, n$$

$$0 \leq z_S \leq 1, \quad \forall S \in \mathcal{S}$$

Polytope $P_{IS,1} \subseteq \mathbb{R}^{n+2}$

$$z_S \leq x_k, \quad \forall k \in S \setminus A$$

$$z_S \leq z_A,$$

$$z_S \geq z_A + \sum_{k \in S \setminus A} x_k - |S \setminus A|,$$

$$z_A \leq x_k, \quad \forall k \in A$$

$$z_A \geq \sum_{k \in A} x_k - (|A| - 1).$$

$$0 \leq x_k \leq 1, \quad \forall k = 1, \dots, n$$

$$0 \leq z_S \leq 1, \quad \forall S \in \mathcal{S}$$

Calculating projections: Fourier-Motzkin Elimination

Notation

$\mathbb{P}_{n,S}$: projection over the space of variables z_S and $x_k, k = 1, \dots, n$.

We calculate $\mathbb{P}_{n,S}(P_{IS,1})$ using the **Fourier-Motzkin Elimination**:

$$\begin{array}{lll} z_S \leq z_A & z_A \leq x_k, & \forall k \in A \\ \sum_{k \in A} x_k - (|A| - 1) \leq z_A & z_A \leq z_S - \sum_{k \in S \setminus A} x_k + |S \setminus A|. & \end{array}$$

We also take into account the inequalities of $P_{IS,1}$ that do not involve z_A

$$z_S \leq x_k, \forall k \in S \setminus A$$

Single monomials

Theorem

$$\mathbb{P}_{n,S}(P_{IS,1}) = P_{SL,1}$$

Theorem holds for *disjoint* several monomials:

$z_S = \prod_{k \in S} x_k$, $z_T = \prod_{k \in T} x_k$, take $A \subseteq S$, $B \subseteq T$.

$$z_S = z_A^S \prod_{k \in S \setminus A} x_k$$

$$z_A^S = \prod_{k \in A} x_k$$

$$z_T = z_B^T \prod_{k \in T \setminus B} x_k$$

$$z_B^T = \prod_{k \in B} x_k$$

Linearize, and apply Fourier-Motzkin as before (constraints never contain at the same time z_A^S and z_B^T).

Single monomials

Theorem

$$\mathbb{P}_{n,S}(P_{IS,1}) = P_{SL,1}$$

Theorem holds for *disjoint* several monomials:

$z_S = \prod_{k \in S} x_k$, $z_T = \prod_{k \in T} x_k$, take $A \subseteq S$, $B \subseteq T$.

$$z_S = z_A^S \prod_{k \in S \setminus A} x_k$$

$$z_A^S = \prod_{k \in A} x_k$$

$$z_T = z_B^T \prod_{k \in T \setminus B} x_k$$

$$z_B^T = \prod_{k \in B} x_k$$

Linearize, and apply Fourier-Motzkin as before (constraints never contain at the same time z_A^S and z_B^T).

Several monomials with common intersection

What happens with *non-disjoint* monomials? $A \subseteq S \cap T$, ($|A| \geq 2$).

$$z_S = z_A \prod_{k \in S \setminus A} x_k$$

$$z_T = z_A \prod_{k \in T \setminus A} x_k$$

$$z_A = \prod_{k \in A} x_k,$$

Several monomials with common intersection

What happens with *non-disjoint* monomials? $A \subseteq S \cap T$, ($|A| \geq 2$).

$$z_S = z_A \prod_{k \in S \setminus A} x_k$$

$$z_T = z_A \prod_{k \in T \setminus A} x_k$$

$$z_A = \prod_{k \in A} x_k,$$

$$z_S \leq x_k, \quad \forall k \in S \setminus A$$

$$z_S \leq z_A$$

$$z_S \geq z_A + \sum_{k \in S \setminus A} x_k - |S \setminus A|$$

$$z_T \leq x_k, \quad \forall k \in T \setminus A$$

$$z_T \leq z_A$$

$$z_T \geq z_A + \sum_{k \in T \setminus A} x_k - |T \setminus A|$$

$$z_A \leq x_k, \quad \forall k \in A$$

$$z_A \geq \sum_{k \in A} x_k - (|A| - 1).$$

Several monomials with common intersection

What happens with *non-disjoint* monomials? $A \subseteq S \cap T$, ($|A| \geq 2$).

$$z_S = z_A \prod_{k \in S \setminus A} x_k$$

$$z_T = z_A \prod_{k \in T \setminus A} x_k$$

$$z_A = \prod_{k \in A} x_k,$$

$$z_S \leq x_k, \quad \forall k \in S \setminus A$$

$$z_S \leq z_A$$

$$z_S \geq z_A + \sum_{k \in S \setminus A} x_k - |S \setminus A|$$

$$z_T \leq x_k, \quad \forall k \in T \setminus A$$

$$z_T \leq z_A$$

$$z_T \geq z_A + \sum_{k \in T \setminus A} x_k - |T \setminus A|$$

$$z_A \leq x_k, \quad \forall k \in A$$

$$z_A \geq \sum_{k \in A} x_k - (|A| - 1).$$

Several monomials with common intersection

What happens with *non-disjoint* monomials? $A \subseteq S \cap T$, ($|A| \geq 2$).

$$z_S = z_A \prod_{k \in S \setminus A} x_k$$

$$z_T = z_A \prod_{k \in T \setminus A} x_k$$

$$z_A = \prod_{k \in A} x_k,$$

$$z_S \leq x_k, \quad \forall k \in S \setminus A$$

$$z_S \leq z_A$$

$$z_S \geq z_A + \sum_{k \in S \setminus A} x_k - |S \setminus A|$$

$$z_T \leq x_k, \quad \forall k \in T \setminus A$$

$$z_T \leq z_A$$

$$z_T \geq z_A + \sum_{k \in T \setminus A} x_k - |T \setminus A|$$

$$z_A \leq x_k, \quad \forall k \in A$$

$$z_A \geq \sum_{k \in A} x_k - (|A| - 1).$$

Several monomials with common intersection

What happens with *non-disjoint* monomials? $A \subseteq S \cap T$, ($|A| \geq 2$).

$$z_S = z_A \prod_{k \in S \setminus A} x_k$$

$$z_T = z_A \prod_{k \in T \setminus A} x_k$$

$$z_A = \prod_{k \in A} x_k,$$

$$z_S \leq x_k, \quad \forall k \in S \setminus A$$

$$z_S \leq z_A$$

$$z_S \geq z_A + \sum_{k \in S \setminus A} x_k - |S \setminus A|$$

$$z_T \leq x_k, \quad \forall k \in T \setminus A$$

$$z_T \leq z_A$$

$$z_T \geq z_A + \sum_{k \in T \setminus A} x_k - |T \setminus A|$$

$$z_A \leq x_k, \quad \forall k \in A$$

$$z_A \geq \sum_{k \in A} x_k - (|A| - 1).$$

Several monomials with common intersection

Theorem

$$\mathbb{P}_{n,S,T}(P_{IS}) \subset P_{SL}$$

Proof:

- 1 Fourier-Motzkin gives:

$$z_S \leq z_T - \sum_{k \in T \setminus A} x_k + |T \setminus A|, \quad (1)$$

$$z_T \leq z_S - \sum_{k \in S \setminus A} x_k + |S \setminus A|, \quad (2)$$

- 2 $\mathbb{P}_{n,S,T}(P_{IS}) = P_{SL} \cap \{(x_k, z_S, z_T) \mid (1), (2) \text{ are satisfied}\}$
- 3 Point $x_k = 1$ for $k \notin A$, $x_k = \frac{1}{2}$ for $k \in A$, $z_S = 0$, $z_T = \frac{1}{2}$, is in P_{SL} but does not satisfy (2).

Larger subset substitutions are better

Consider $B \subset A \subseteq S \cap T$, $|B| \geq 2$.

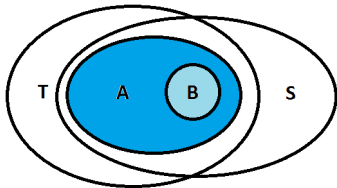
- 1 Take the first cut for both subsets:

$$z_S \leq z_T - \sum_{k \in T \setminus A} x_k + |T \setminus A|,$$

$$z_S \leq z_T - \sum_{k \in T \setminus B} x_k + |T \setminus B|,$$

2

$$\begin{aligned} z_S &\leq z_T - \sum_{k \in T \setminus A} x_k + |T \setminus A| \leq \\ &\leq z_T - \sum_{k \in T \setminus A} x_k + |T \setminus A| - \sum_{k \in A \setminus B} x_k + |A \setminus B| = \\ &= z_T - \sum_{k \in T \setminus B} x_k + |T \setminus B|. \end{aligned}$$



Larger subset substitutions are better

Theorem

$$\mathbb{P}_{n,S,T}(P_{IS}^A) \subset \mathbb{P}_{n,S,T}(P_{IS}^B).$$

(Point $x_k = 1$ for $k \notin A$, $x_k = \frac{1}{2}$ for $k \in A \setminus B$, $k \in B$, $z_T = 0$, $z_S = \frac{1}{2}$ satisfies cut for B but not for A .)

Corollary

Consider three monomials R, S, T , with intersections $R \cap S = A$, $S \cap T = B$, $R \cap T = C$, ($|A|, |B|, |C| \geq 2$). Then it is better to do intermediate substitutions of the two-by-two intersections, than a single intermediate substitution of the common intersection $A \cap B \cap C$.

Larger subset substitutions are better

Theorem

$$\mathbb{P}_{n,S,T}(P_{IS}^A) \subset \mathbb{P}_{n,S,T}(P_{IS}^B).$$

(Point $x_k = 1$ for $k \notin A$, $x_k = \frac{1}{2}$ for $k \in A \setminus B$, $k \in B$, $z_T = 0$, $z_S = \frac{1}{2}$ satisfies cut for B but not for A .)

Corollary

Consider three monomials R, S, T , with intersections $R \cap S = A$, $S \cap T = B$, $R \cap T = C$, ($|A|, |B|, |C| \geq 2$). Then it is better to do intermediate substitutions of the two-by-two intersections, than a single intermediate substitution of the common intersection $A \cap B \cap C$.

Improving the SL formulation: 2-intersection-cuts

SL relaxation with 2-intersection-cuts

$$\begin{aligned}
 & \min \sum_{S \in \mathcal{S}} a_S z_S \\
 \text{s.t. } & z_S \leq x_k, & \forall k \in S, \forall S \in \mathcal{S} \\
 & z_S \geq \sum_{k \in S} x_k - (|S| - 1), & \forall S \in \mathcal{S} \\
 & z_S \leq z_T - \sum_{k \in T \setminus S} x_k + |T \setminus S| & \forall S, T, |S \cap T| \geq 2 \\
 & z_T \leq z_S - \sum_{k \in S \setminus T} x_k + |S \setminus T| & \forall S, T, |S \cap T| \geq 2 \\
 & 0 \leq z_S \leq 1, & \forall S \in \mathcal{S} \\
 & 0 \leq x_k \leq 1 & \forall k = 1, \dots, n
 \end{aligned}$$

How strong are the 2-intersection-cuts?

Consider the *standard linearization polytope*:

$$\begin{aligned} P_{SL}^{conv} &= \text{conv}\{(x, y_S) \in \{0, 1\}^{n+|\mathcal{S}|} \mid y_S = \prod_{i \in S} x_i, \forall S \in \mathcal{S}\} \\ &= \text{conv}\{(x, y_S) \in \{0, 1\}^{n+|\mathcal{S}|} \mid y_S \leq x_i, y_S \geq \sum_{i \in S} x_i - (|S| - 1), \forall S \in \mathcal{S}\}, \end{aligned}$$

and its linear relaxation

$$P_{SL} = \{(x, y_S) \in [0, 1]^{n+|\mathcal{S}|} \mid y_S \leq x_i, y_S \geq \sum_{i \in S} x_i - (|S| - 1), \forall S \in \mathcal{S}\}$$

- **Question 1:** Are the 2-intersection-cuts *facet-defining* for P_{SL}^{conv} ?

How strong are the 2-intersection-cuts?

Consider the *standard linearization polytope*:

$$\begin{aligned} P_{SL}^{conv} &= \text{conv}\{(x, y_S) \in \{0, 1\}^{n+|S|} \mid y_S = \prod_{i \in S} x_i, \forall S \in \mathcal{S}\} \\ &= \text{conv}\{(x, y_S) \in \{0, 1\}^{n+|S|} \mid y_S \leq x_i, y_S \geq \sum_{i \in S} x_i - (|S| - 1), \forall S \in \mathcal{S}\}, \end{aligned}$$

and its linear relaxation

$$P_{SL} = \{(x, y_S) \in [0, 1]^{n+|S|} \mid y_S \leq x_i, y_S \geq \sum_{i \in S} x_i - (|S| - 1), \forall S \in \mathcal{S}\}$$

- **Question 1:** Are the 2-intersection-cuts *facet-defining* for P_{SL}^{conv} ?
- **Question 2:** Is there some case for which we obtain the *convex hull* P_{SL}^{conv} when adding the 2-intersection-cuts to P_{SL} ?

How strong are the 2-intersection-cuts?

Consider the *standard linearization polytope*:

$$\begin{aligned} P_{SL}^{conv} &= \text{conv}\{(x, y_S) \in \{0, 1\}^{n+|S|} \mid y_S = \prod_{i \in S} x_i, \forall S \in \mathcal{S}\} \\ &= \text{conv}\{(x, y_S) \in \{0, 1\}^{n+|S|} \mid y_S \leq x_i, y_S \geq \sum_{i \in S} x_i - (|S| - 1), \forall S \in \mathcal{S}\}, \end{aligned}$$

and its linear relaxation

$$P_{SL} = \{(x, y_S) \in [0, 1]^{n+|S|} \mid y_S \leq x_i, y_S \geq \sum_{i \in S} x_i - (|S| - 1), \forall S \in \mathcal{S}\}$$

- **Question 1:** Are the 2-intersection-cuts *facet-defining* for P_{SL}^{conv} ?
- **Question 2:** Is there some case for which we obtain the *convex hull* P_{SL}^{conv} when adding the 2-intersection-cuts to P_{SL} ?

Facet-defining cuts (2 monomials)

Theorem: 2-term objective function

The 2-intersection-cuts are facet-defining for $P_{SL,2}^{conv}$:

$$z_S \leq z_T - \sum_{k \in T \setminus S} x_k + |T \setminus S|$$

$$z_T \leq z_S - \sum_{k \in S \setminus T} x_k + |S \setminus T|$$

Facet-defining cuts (2 monomials)

Special forms of the cuts in some cases:

- ① If $S \subseteq T$,

$$z_S \leq z_T - \sum_{k \in T \setminus S} x_k + |T \setminus S|$$

$$z_T \leq z_S$$

- ② If $T = \emptyset$ (and setting by definition $z_\emptyset = 1$),

$$z_S \leq 1$$

$$1 \leq z_S - \sum_{i \in S} x_i + |S|$$

Conjecture on the convex hull (2 monomials)

Conjecture

Consider a pseudo-Boolean function consisting of two terms, its standard linearization polytope $P_{SL,2}^{conv}$ and its linear relaxation $P_{SL,2}$. Then,

$$P_{SL,2}^{conv} = P_{SL,2} \cap \{(x, y_S, y_T) \in [0, 1]^{n+2} \mid \text{2-intersection-cuts are satisfied}\}.$$

Facet-defining cuts (nested monomials)

Theorem: Nested sequence of terms

Consider a pseudo-Boolean function $f(x) = \sum_{l \in L} a_{S(l)} \prod_{i \in S(l)} x_i$, such that $S^{(1)} \subseteq S^{(2)} \subseteq \dots \subseteq S^{(|L|)}$, and its standard linearization polytope $P_{SL, nest}^{conv}$.
The 2-intersection-cuts

$$z_{S^{(l)}} \leq z_{S^{(l+1)}} - \sum_{k \in S^{(l+1)} \setminus S^{(l)}} x_k + |S^{(l+1)} \setminus S^{(l)}|$$

$$z_{S^{(l+1)}} \leq z_{S^{(l)}},$$

are facet-defining for $P_{SL, nest}^{conv}$ for two consecutive monomials in the nest (and cuts are redundant for non-consecutive monomials).

Conjectures for m monomials

Conjecture: facet-defining

The 2-intersection-cuts are facet-defining for the case of m monomials.

Convex-hull for the general case

The 2-intersection-cuts and standard linearization inequalities are **not** enough to define the convex hull P_{SL}^{conv} (otherwise we could solve an \mathcal{NP} -hard problem efficiently...).

Conjectures for m monomials

Conjecture: facet-defining

The 2-intersection-cuts are facet-defining for the case of m monomials.

Convex-hull for the general case

The 2-intersection-cuts and standard linearization inequalities are **not** enough to define the convex hull P_{SL}^{conv} (otherwise we could solve an \mathcal{NP} -hard problem efficiently...).

- $m = 3$, set of 3 monomials for which there exists an objective function which has a fractional optimal solution on $P_{SL} \cap \{2\text{-intersection-cuts}\}$:
 $\{x_1 x_2 x_4, x_1 x_3 x_4, x_1 x_2 x_3\}$

Conjectures for m monomials

Conjecture: facet-defining

The 2-intersection-cuts are facet-defining for the case of m monomials.

Convex-hull for the general case

The 2-intersection-cuts and standard linearization inequalities are **not** enough to define the convex hull P_{SL}^{conv} (otherwise we could solve an \mathcal{NP} -hard problem efficiently...).

- $m = 3$, set of 3 monomials for which there exists an objective function which has a fractional optimal solution on $P_{SL} \cap \{2\text{-intersection-cuts}\}$:
 $\{x_1 x_2 x_4, x_1 x_3 x_4, x_1 x_2 x_3\}$

A (vague) idea of the convex hull for the general case

Idea of the convex hull

$$\min \sum_{S \in \mathcal{S}} a_S z_S$$

s.t. SL-constraints: linking a term with its variables

2-intersection inequalities: linking terms 2 by 2

3-intersection inequalities: linking terms 3 by 3

...

$$0 \leq z_S \leq 1,$$

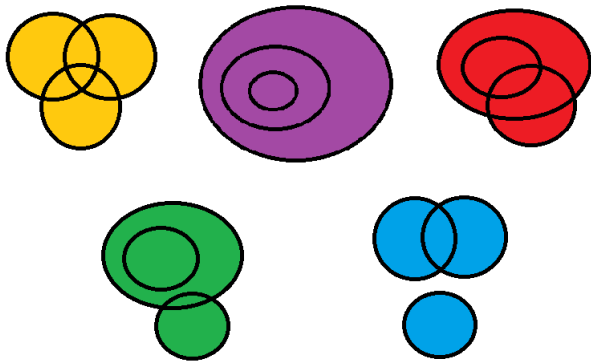
$$0 \leq x_k \leq 1$$

$$\forall S \in \mathcal{S}$$

$$\forall k = 1, \dots, n$$

One way of viewing the difficulty of the convex hull

For 3 monomials we already have many different possible ways for them to intersect:



A short summary of the linearizations part and some ideas

- We have obtained interesting cuts for P_{SL} by applying intermediate substitutions for subsets of size ≥ 2 .
- We could apply iteratively these intermediate substitutions, the last substitution step has only quadratic constraints

$$z_{ij} = x_i x_j,$$

$$z_{iJ} = x_i z_J,$$

$$z_{IJ} = z_I z_J,$$

x : original variables, z : variables that are already substitutions of other subsets.

A short summary of the linearizations part and some ideas

- We have obtained interesting cuts for P_{SL} by applying intermediate substitutions for subsets of size ≥ 2 .
- We could apply iteratively these intermediate substitutions, the last substitution step has only quadratic constraints

$$z_{ij} = x_i x_j,$$

$$z_{iJ} = x_i z_J,$$

$$z_{IJ} = z_I z_J,$$

x : original variables, z : variables that are already substitutions of other subsets.

- **Open questions:**
 - How many intermediate substitutions provide practical improvements?
 - Relationship of this quadratically constrained program with quadratizations.

A short summary of the linearizations part and some ideas

- We have obtained interesting cuts for P_{SL} by applying intermediate substitutions for subsets of size ≥ 2 .
- We could apply iteratively these intermediate substitutions, the last substitution step has only quadratic constraints

$$z_{ij} = x_i x_j,$$

$$z_{iJ} = x_i z_J,$$

$$z_{IJ} = z_I z_J,$$

x : original variables, z : variables that are already substitutions of other subsets.

- **Open questions:**
 - How many intermediate substitutions provide practical improvements?
 - Relationship of this quadratically constrained program with quadratizations.

Quadratizations: reductions to the quadratic case

Linearizations vs. quadratizations

Initial objective function (polynomial and binary):

$$f(x) = \sum_{S \in \mathcal{S}} a_S \prod_{i \in S} x_i$$

Linearizations

Introduce **new variables** to obtain an equivalent **linear** problem.

Linearizations vs. quadratizations

Initial objective function (polynomial and binary):

$$f(x) = \sum_{S \in \mathcal{S}} a_S \prod_{i \in S} x_i$$

Linearizations

Introduce **new variables** to obtain an equivalent **linear** problem.

Quadratizations

Introduce **new variables** to obtain an equivalent **quadratic** problem.

Linearizations vs. quadratizations

Initial objective function (polynomial and binary):

$$f(x) = \sum_{S \in \mathcal{S}} a_S \prod_{i \in S} x_i$$

Linearizations

Introduce **new variables** to obtain an equivalent **linear** problem.

Quadratizations

Introduce **new variables** to obtain an equivalent **quadratic** problem.

- Quadratic binary optimization is \mathcal{NP} -hard.

Linearizations vs. quadratizations

Initial objective function (polynomial and binary):

$$f(x) = \sum_{S \in \mathcal{S}} a_S \prod_{i \in S} x_i$$

Linearizations

Introduce **new variables** to obtain an equivalent **linear** problem.

Quadratizations

Introduce **new variables** to obtain an equivalent **quadratic** problem.

- Quadratic binary optimization is \mathcal{NP} -hard.
- Many techniques known for the quadratic case (exact algorithms, heuristics, cuts...).

Linearizations vs. quadratizations

Initial objective function (polynomial and binary):

$$f(x) = \sum_{S \in \mathcal{S}} a_S \prod_{i \in S} x_i$$

Linearizations

Introduce **new variables** to obtain an equivalent **linear** problem.

Quadratizations

Introduce **new variables** to obtain an equivalent **quadratic** problem.

- Quadratic binary optimization is \mathcal{NP} -hard.
- Many techniques known for the quadratic case (exact algorithms, heuristics, cuts...).

Quadratization methods: Rosenberg (Example)

$$\min_{x \in \{0,1\}^4} 3x_1x_2x_3x_4 + 2x_1x_2x_5 - 5x_1x_2 + 6x_3x_4 - x_1 + x_2 - x_3 + x_4$$

Rosenberg (penalties): Iteration 1

$$\min_{x \in \{0,1\}^4, y_{12} \in \{0,1\}} 3y_{12}x_3x_4 + 2y_{12}x_5 - 5y_{12} + 6x_3x_4 - x_1 + x_2 - x_3 + x_4 \\ + M_1(x_1x_2 - 2x_1y_{12} - 2x_2y_{12} + 3y_{12})$$

Quadratization methods: Rosenberg (Example)

$$\min_{x \in \{0,1\}^4} 3x_1x_2x_3x_4 + 2x_1x_2x_5 - 5x_1x_2 + 6x_3x_4 - x_1 + x_2 - x_3 + x_4$$

Rosenberg (penalties): Iteration 1

$$\min_{x \in \{0,1\}^4, y_{12} \in \{0,1\}} 3y_{12}x_3x_4 + 2y_{12}x_5 - 5y_{12} + 6x_3x_4 - x_1 + x_2 - x_3 + x_4 \\ + M_1(x_1x_2 - 2x_1y_{12} - 2x_2y_{12} + 3y_{12})$$

The penalty vanishes if $y_{12} = x_1x_2$:

- $y_{12} = 0$ and $x_1 = 0$ (or $x_2 = 0$): all terms vanish,
- $y_{12} = 1$ and both $x_1 = 1, x_2 = 1$: $M_1 - 2M_1 - 2M_1 + 3M_1 = 0$.

Quadratization methods: Rosenberg (Example)

$$\min_{x \in \{0,1\}^4} 3x_1x_2x_3x_4 + 2x_1x_2x_5 - 5x_1x_2 + 6x_3x_4 - x_1 + x_2 - x_3 + x_4$$

Rosenberg (penalties): Iteration 1

$$\min_{x \in \{0,1\}^4, y_{12} \in \{0,1\}} 3y_{12}x_3x_4 + 2y_{12}x_5 - 5y_{12} + 6x_3x_4 - x_1 + x_2 - x_3 + x_4 \\ + M_1(x_1x_2 - 2x_1y_{12} - 2x_2y_{12} + 3y_{12})$$

The penalty vanishes if $y_{12} = x_1x_2$:

- $y_{12} = 0$ and $x_1 = 0$ (or $x_2 = 0$): all terms vanish,
- $y_{12} = 1$ and both $x_1 = 1, x_2 = 1$: $M_1 - 2M_1 - 2M_1 + 3M_1 = 0$.

A penalty M_1 is incurred at least one time if $y_{12} \neq x_1x_2$:

- $y_{12} = 1$ and $x_1 = 0$: $-2M_1 + 3M_1 = M_1$ (idem if $x_2 = 0$),
- $y_{12} = 0$ and both $x_1 = 1, x_2 = 1$: all terms vanish but the first one, M_1 is incurred.

Quadratization methods: Rosenberg (Example)

$$\min_{x \in \{0,1\}^4} 3x_1x_2x_3x_4 + 2x_1x_2x_5 - 5x_1x_2 + 6x_3x_4 - x_1 + x_2 - x_3 + x_4$$

Rosenberg (penalties): Iteration 1

$$\min_{x \in \{0,1\}^4, y_{12} \in \{0,1\}} 3y_{12}x_3x_4 + 2y_{12}x_5 - 5y_{12} + 6x_3x_4 - x_1 + x_2 - x_3 + x_4 \\ + M_1(x_1x_2 - 2x_1y_{12} - 2x_2y_{12} + 3y_{12})$$

The penalty vanishes if $y_{12} = x_1x_2$:

- $y_{12} = 0$ and $x_1 = 0$ (or $x_2 = 0$): all terms vanish,
- $y_{12} = 1$ and both $x_1 = 1, x_2 = 1$: $M_1 - 2M_1 - 2M_1 + 3M_1 = 0$.

A penalty M_1 is incurred at least one time if $y_{12} \neq x_1x_2$:

- $y_{12} = 1$ and $x_1 = 0$: $-2M_1 + 3M_1 = M_1$ (idem if $x_2 = 0$),
- $y_{12} = 0$ and both $x_1 = 1, x_2 = 1$: all terms vanish but the first one, M_1 is incurred.

Quadratization methods: Rosenberg (Example)

$$\min_{x \in \{0,1\}^4} 3x_1x_2x_3x_4 + 2x_1x_2x_5 - 5x_1x_2 + 6x_3x_4 - x_1 + x_2 - x_3 + x_4$$

Rosenberg (penalties): Iteration 1

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12} \in \{0,1\}} & 3y_{12}x_3x_4 + 2y_{12}x_5 - 5y_{12} + 6x_3x_4 - x_1 + x_2 - x_3 + x_4 \\ & + \mathbf{M}_1(x_1x_2 - 2x_1y_{12} - 2x_2y_{12} + 3y_{12}) \end{aligned}$$

Rosenberg (penalties): Iteration 2

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12}, y_{34} \in \{0,1\}} & 3y_{12}y_{34} + 2y_{12}x_5 - 5y_{12} + 6y_{34} - x_1 + x_2 - x_3 + x_4 \\ & + \mathbf{M}_1(x_1x_2 - 2x_1y_{12} - 2x_2y_{12} + 3y_{12}) \\ & + \mathbf{M}_2(x_3x_4 - 2x_3y_{34} - 2x_4y_{34} + 3y_{34}) \end{aligned}$$

Quadratization methods: Rosenberg (Example)

$$\min_{x \in \{0,1\}^4} 3x_1x_2x_3x_4 + 2x_1x_2x_5 - 5x_1x_2 + 6x_3x_4 - x_1 + x_2 - x_3 + x_4$$

Rosenberg (constraints): Iteration 1

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12} \in \{0,1\}} & 3y_{12}x_3x_4 + 2y_{12}x_5 - 5y_{12} + 6x_3x_4 - x_1 + x_2 - x_3 + x_4 \\ \text{s.t. } & y_{12} = x_1x_2 \end{aligned}$$

Rosenberg (penalties): Iteration 2

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12}, y_{34} \in \{0,1\}} & 3y_{12}y_{34} + 2y_{12}x_5 - 5y_{12} + 6y_{34} - x_1 + x_2 - x_3 + x_4 \\ & + M_1(x_1x_2 - 2x_1y_{12} - 2x_2y_{12} + 3y_{12}) \\ & + M_2(x_3x_4 - 2x_3y_{34} - 2x_4y_{34} + 3y_{34}) \end{aligned}$$

Quadratization methods: Rosenberg (Example)

$$\min_{x \in \{0,1\}^4} 3x_1x_2x_3x_4 + 2x_1x_2x_5 - 5x_1x_2 + 6x_3x_4 - x_1 + x_2 - x_3 + x_4$$

Rosenberg (constraints): Iteration 1

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12} \in \{0,1\}} & 3y_{12}x_3x_4 + 2y_{12}x_5 - 5y_{12} + 6x_3x_4 - x_1 + x_2 - x_3 + x_4 \\ \text{s.t.} & y_{12} = x_1x_2 \end{aligned}$$

Rosenberg (constraints): Iteration 2

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12}, y_{34} \in \{0,1\}} & 3y_{12}y_{34} + 2y_{12}x_5 - 5y_{12} + 6y_{34} - x_1 + x_2 - x_3 + x_4 \\ \text{s.t.} & y_{12} = x_1x_2 \\ & y_{34} = x_3x_4 \end{aligned}$$

Quadratization methods: Rosenberg

Rosenberg (1975) [5]: first quadratization method.

- 1 Take a product $x_i x_j$ from a highest-degree monomial of f and substitute it by a new variable y_{ij} .

Quadratization methods: Rosenberg

Rosenberg (1975) [5]: first quadratization method.

- 1 Take a product $x_i x_j$ from a highest-degree monomial of f and substitute it by a new variable y_{ij} .
- 2 Add a penalty term $M(x_i x_j - 2x_i y_{ij} - 2x_j y_{ij} + 3y_{ij})$ (M large enough) to the objective function to force $y_{ij} = x_i x_j$ at all optimal solutions.

Quadratization methods: Rosenberg

Rosenberg (1975) [5]: first quadratization method.

- 1 Take a product $x_i x_j$ from a highest-degree monomial of f and substitute it by a new variable y_{ij} .
- 2 Add a penalty term $M(x_i x_j - 2x_i y_{ij} - 2x_j y_{ij} + 3y_{ij})$ (M large enough) to the objective function to force $y_{ij} = x_i x_j$ at all optimal solutions.
- 3 Iterate until obtaining a quadratic function.

Quadratization methods: Rosenberg

Rosenberg (1975) [5]: first quadratization method.

- 1 Take a product $x_i x_j$ from a highest-degree monomial of f and substitute it by a new variable y_{ij} .
 - 2 Add a penalty term $M(x_i x_j - 2x_i y_{ij} - 2x_j y_{ij} + 3y_{ij})$ (M large enough) to the objective function to force $y_{ij} = x_i x_j$ at all optimal solutions.
 - 3 Iterate until obtaining a quadratic function.
- **Advantages:**
 - Can be applied to any pseudo-Boolean function f .
 - The transformation is polynomial in the size of the input.

Quadratization methods: Rosenberg

Rosenberg (1975) [5]: first quadratization method.

- 1 Take a product $x_i x_j$ from a highest-degree monomial of f and substitute it by a new variable y_{ij} .
- 2 Add a penalty term $M(x_i x_j - 2x_i y_{ij} - 2x_j y_{ij} + 3y_{ij})$ (M large enough) to the objective function to force $y_{ij} = x_i x_j$ at all optimal solutions.
- 3 Iterate until obtaining a quadratic function.

- **Advantages:**

- Can be applied to any pseudo-Boolean function f .
- The transformation is polynomial in the size of the input.

- **Drawbacks:**

- The obtained quadratization is highly non-submodular.
- Big M .

Quadratization methods: Rosenberg

Rosenberg (1975) [5]: first quadratization method.

- 1 Take a product $x_i x_j$ from a highest-degree monomial of f and substitute it by a new variable y_{ij} .
- 2 Add a penalty term $M(x_i x_j - 2x_i y_{ij} - 2x_j y_{ij} + 3y_{ij})$ (M large enough) to the objective function to force $y_{ij} = x_i x_j$ at all optimal solutions.
- 3 Iterate until obtaining a quadratic function.

- **Advantages:**

- Can be applied to any pseudo-Boolean function f .
- The transformation is polynomial in the size of the input.

- **Drawbacks:**

- The obtained quadratization is highly non-submodular.
- Big M .

Quadratization methods: Rosenberg with constraints

Rosenberg (1975) [5]: first quadratization method (**Variant**).

- ① Take a product $x_i x_j$ from a highest-degree monomial of f and substitute it by a new variable y_{ij} .
- ② ~~Add a penalty term $M(x_i x_j - 2x_i y_{ij} - 2x_j y_{ij} + 3y_{ij})$ (M large enough) to the objective function...~~ **Add a constraint** $y_{ij} = x_i x_j$.
- ③ Iterate until obtaining a quadratic function.

- **Advantages:**

- Can be applied to any pseudo-Boolean function f .
- The transformation is polynomial in the size of the input.

- **Drawbacks:**

- ~~The obtained quadratization is highly non-submodular.~~
- ~~Big M .~~

Rosenberg and intermediate substitutions

Rosenberg (constraints): Iteration 2

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12}, y_{34} \in \{0,1\}} \quad & 3y_{12}y_{34} + 2y_{12}x_5 - 5y_{12} + 6y_{34} - x_1 + x_2 - x_3 + x_4 \\ \text{s.t.} \quad & y_{12} = x_1x_2 \\ & y_{34} = x_3x_4 \end{aligned}$$

Linearization of Rosenberg

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12}, y_{34}, y_{1234}, y_{125} \in \{0,1\}} \quad & 3y_{1234} + 2y_{125} - 5y_{12} + 6y_{34} - x_1 + x_2 - x_3 + x_4 \\ \text{s.t.} \quad & y_{12} = x_1x_2 \\ & y_{34} = x_3x_4 \\ & y_{1234} = y_{12}y_{34} \\ & y_{125} = y_{12}x_5 \end{aligned}$$

Rosenberg and intermediate substitutions

Rosenberg (constraints): Iteration 2

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12}, y_{34} \in \{0,1\}} \quad & 3y_{12}y_{34} + 2y_{12}x_5 - 5y_{12} + 6y_{34} - x_1 + x_2 - x_3 + x_4 \\ \text{s.t.} \quad & y_{12} = x_1x_2 \\ & y_{34} = x_3x_4 \end{aligned}$$

Linearization of Rosenberg

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12}, y_{34}, y_{1234}, y_{125} \in \{0,1\}} \quad & 3y_{1234} + 2y_{125} - 5y_{12} + 6y_{34} - x_1 + x_2 - x_3 + x_4 \\ \text{s.t.} \quad & y_{12} = x_1x_2 \\ & y_{34} = x_3x_4 \\ & y_{1234} = y_{12}y_{34} \\ & y_{125} = y_{12}x_5 \end{aligned}$$

Rosenberg and intermediate substitutions

Iterated intermediate substitutions

$$z_{ij} = x_i x_j,$$

$$z_{iJ} = x_i z_J,$$

$$z_{IJ} = z_I z_J,$$

- How many intermediate substitutions provide practical improvements?
- Relationship of this quadratically constrained program with quadratizations.

Linearization of Rosenberg

$$\begin{aligned} \min_{x \in \{0,1\}^4, y_{12}, y_{34}, y_{1234}, y_{125} \in \{0,1\}} \quad & 3y_{1234} + 2y_{125} - 5y_{12} + 6y_{34} - x_1 + x_2 - x_3 + x_4 \\ \text{s.t.} \quad & y_{12} = x_1 x_2 \\ & y_{34} = x_3 x_4 \\ & y_{1234} = y_{12} y_{34} \\ & y_{125} = y_{12} x_5 \end{aligned}$$

Linearization of quadratic vs. polynomial functions

Buchheim and Rinaldi's result [2]:

Consider linearization of a polynomial function:

$$P_{SL}^{conv} = \text{conv}\{(x, y_S) \in \{0, 1\}^{n+|S|} \mid y_S = \prod_{i \in S} x_i, \forall S \in \mathcal{S}\}$$

Define an extended quadratic formulation and linearize it:

$$P^* = \text{conv}\{y_{\{S, T\}} \in \{0, 1\} \mid y_{\{S, T\}} = y_S y_T, \forall \{S, T\} \text{ where } S, T, S \cup T \in \mathcal{S}\}$$

If we know P^* then we can construct P_{SL}^{conv} .

Questions: If instead of having P^* , we have a relaxation,

- what do we know about P_{SL}^{conv} ?
- for the standard linearization, 2-intersection, 3-intersection (up to m -intersection) cuts help to obtain information about P_{SL}^{conv} , what about other relaxations?

Perspectives

Perspectives

- Several conjectures concerning the strength of the intersection-cuts that we generate with intermediate substitutions
 - Case of 2 monomials: convex hull of standard linearization polytope using intersection-cuts and linearization inequalities.

Perspectives

- Several conjectures concerning the strength of the intersection-cuts that we generate with intermediate substitutions
 - Case of 2 monomials: convex hull of standard linearization polytope using intersection-cuts and linearization inequalities.
 - Case of m monomials: intersection-cuts are facet-defining.

Perspectives

- Several conjectures concerning the strength of the intersection-cuts that we generate with intermediate substitutions
 - Case of 2 monomials: convex hull of standard linearization polytope using intersection-cuts and linearization inequalities.
 - Case of m monomials: intersection-cuts are facet-defining.
 - Idea on the general form of the convex hull for m terms.

Perspectives

- Several conjectures concerning the strength of the intersection-cuts that we generate with intermediate substitutions
 - Case of 2 monomials: convex hull of standard linearization polytope using intersection-cuts and linearization inequalities.
 - Case of m monomials: intersection-cuts are facet-defining.
 - Idea on the general form of the convex hull for m terms.
- Run computational experiments to measure improvements of the intersection-cuts.






Perspectives

- Several conjectures concerning the strength of the intersection-cuts that we generate with intermediate substitutions
 - Case of 2 monomials: convex hull of standard linearization polytope using intersection-cuts and linearization inequalities.
 - Case of m monomials: intersection-cuts are facet-defining.
 - Idea on the general form of the convex hull for m terms.
- Run computational experiments to measure improvements of the intersection-cuts.
- Information of the quadratic case that can be used to improve knowledge on linearizations of polynomials.

Perspectives

- Several conjectures concerning the strength of the intersection-cuts that we generate with intermediate substitutions
 - Case of 2 monomials: convex hull of standard linearization polytope using intersection-cuts and linearization inequalities.
 - Case of m monomials: intersection-cuts are facet-defining.
 - Idea on the general form of the convex hull for m terms.
- Run computational experiments to measure improvements of the intersection-cuts.
- Information of the quadratic case that can be used to improve knowledge on linearizations of polynomials.

Some references I

-  M. Anthony, E. Boros, Y. Crama, and A. Gruber. Quadratic reformulations of nonlinear binary optimization problems. Working paper, 2014.
-  C. Buchheim and G. Rinaldi. Efficient reduction of polynomial zero-one optimization to the quadratic case. *SIAM Journal on Optimization*, 18(4):1398–1413, 2007.
-  C. De Simone. The cut polytope and the boolean quadric polytope. *Discrete Mathematics*, 79(1):71–75, 1990.
-  P. L. Hammer, I. Rosenberg, and S. Rudeanu. On the determination of the minima of pseudo-boolean functions. *Studii si Cercetari Matematice*, 14:359–364, 1963. in Romanian.
-  I. G. Rosenberg. Reduction of bivalent maximization to the quadratic case. *Cahiers du Centre d'Etudes de Recherche Opérationnelle*, 17:71–74, 1975.

Some references II



Fengqi You, Ignacio E. Grossman, Mixed-Integer Nonlinear Programming Models and Algorithms for Large-Scale Supply Chain Design with Stochastic Inventory Management, **2008**, *Industrial & Engineering Chemistry Research*, 47 (20), pp 7802–7817.