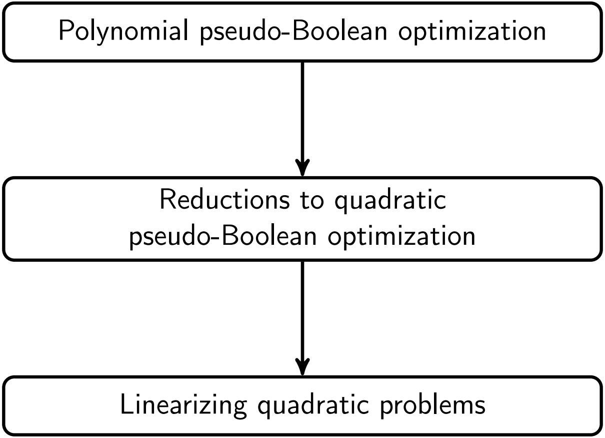
Quadratizations of pseudo-Boolean functions

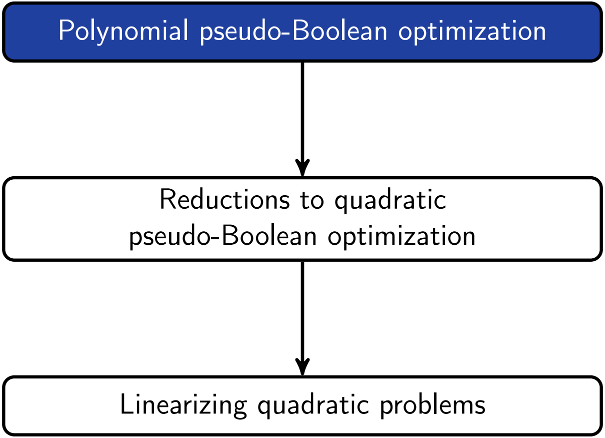
Elisabeth Rodriguez-Heck and Yves Crama

QuantOM, HEC Management School, University of Liège

Partially supported by Belspo - IAP Project COMEX

4th December 2014


Polynomial pseudo-Boolean optimization
Reductions to quadratic pseudo-Boolean optimization
Linearizing quadratic problems


Polynomial pseudo-Boolean optimization
Reductions to quadratic pseudo-Boolean optimization
Linearizing quadratic problems

Definitions

Definition: Pseudo-Boolean functions

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

Multilinear representation

Every pseudo-Boolean function f can be represented uniquely by a multilinear polynomial (Hammer, Rosenberg, Rudeanu [7]).

Definitions

Definition: Pseudo-Boolean functions

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

Multilinear representation

Every pseudo-Boolean function f can be represented uniquely by a multilinear polynomial (Hammer, Rosenberg, Rudeanu [7]).

Example:

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

Definitions

Definition: Pseudo-Boolean functions

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

Multilinear representation

Every pseudo-Boolean function f can be represented uniquely by a multilinear polynomial (Hammer, Rosenberg, Rudeanu [7]).

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: MAX-SAT

MAX-SAT problem

- INPUT: a set of Boolean clauses $C_k = \left(\vee_{i\in A_k}\bar{x}_i\right) \vee \left(\vee_{j\in B_k} x_j\right)$, for $k = 1, \ldots, m$, where $x_i \in \{0,1\}$, and $\bar{x}_i = 1 - x_i$.
- OBJECTIVE: find an assignment of the variables, $x^* \in \{0,1\}^n$ that maximizes the number of satisfied clauses.

Pseudo-Boolean formulation

$$\min \sum_{k=1}^{m} \left(\prod_{i\in A_k} x_i\right)\left(\prod_{j\in B_k} \bar{x}_j\right),$$

C_k takes value 1 *iff* the term $\prod_{i\in A_k} x_i \prod_{j\in B_k} \bar{x}_j$ takes value 0.

Applications: MAX-SAT

MAX-SAT problem

- INPUT: a set of Boolean clauses $C_k = \left(\vee_{i \in A_k} \bar{x}_i\right) \vee \left(\vee_{j \in B_k} x_j\right)$, for $k = 1, \ldots, m$, where $x_i \in \{0, 1\}$, and $\bar{x}_i = 1 - x_i$.
- OBJECTIVE: find an assignment of the variables, $x^* \in \{0, 1\}^n$ that maximizes the number of satisfied clauses.

Pseudo-Boolean formulation

$$\min \sum_{k=1}^{m} \left(\prod_{i \in A_k} x_i \right) \left(\prod_{j \in B_k} \bar{x}_j \right),$$

C_k takes value 1 *iff* the term $\prod_{i \in A_k} x_i \prod_{j \in B_k} \bar{x}_j$ takes value 0.

Applications: Computer Vision

Image restoration problems modelled as energy minimization

$$E(l) = \sum_{p \in \mathcal{P}} D_p(l_p) + \sum_{S \subseteq \mathcal{P}, |S| \geq 2} \sum_{p_1,\ldots,p_s \in S} V_{p_1,\ldots,p_s}(l_{p_1}, \ldots, l_{p_s}),$$

where $l_p \in \{0,1\} \;\; \forall p \in \mathcal{P}$.

(Image from "Corel database" with additive Gaussian noise.)

Applications

- Constraint Satisfaction Problem
- Data mining, classification, learning theory...
- Graph theory
- Operations research
- Production management
- ...

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).

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).
- Much progress has been done for the quadratic case (exact algorithms, heuristics, polyhedral results...).

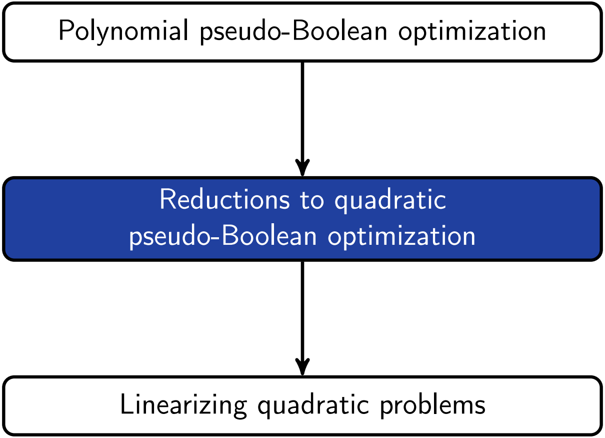
Pseudo-Boolean Optimization

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).
- Much progress has been done for the quadratic case (exact algorithms, heuristics, polyhedral results...).


Polynomial pseudo-Boolean optimization
Reductions to quadratic pseudo-Boolean optimization
Linearizing quadratic problems

Quadratizations

Definition: Quadratization

Given a pseudo-Boolean function $f(x)$ on $\{0,1\}^n$, we say that $g(x,y)$ is a *quadratization* of f if $g(x,y)$ is a quadratic polynomial depending on x and on m *auxiliary variables* $y_1,\dots,y_m$, such that

$$f(x) = \min\{g(x,y) : y \in \{0,1\}^m\} \quad \forall x \in \{0,1\}^n.$$

Then, $\min\{f(x) : x \in \{0,1\}^n\} = \min\{g(x,y) : x \in \{0,1\}^n, y \in \{0,1\}^m\}$.

Which quadratizations are "good"?

- Small number of auxiliary variables.
- Good optimization properties: submodularity.

Quadratizations

Definition: Quadratization

Given a pseudo-Boolean function $f(x)$ on $\{0,1\}^n$, we say that $g(x,y)$ is a *quadratization* of f if $g(x,y)$ is a quadratic polynomial depending on x and on *m auxiliary variables* $y_1,\ldots,y_m$, such that

$$f(x) = \min\{g(x,y) : y \in \{0,1\}^m\} \quad \forall x \in \{0,1\}^n.$$

Then, $\min\{f(x) : x \in \{0,1\}^n\} = \min\{g(x,y) : x \in \{0,1\}^n, y \in \{0,1\}^m\}$.

Which quadratizations are "good"?

- Small number of auxiliary variables.
- Good optimization properties: submodularity.
  - A quadratic pseudo-Boolean function is submodular *iff* all quadratic terms have non-positive coefficients.

Quadratizations

Definition: Quadratization

Given a pseudo-Boolean function $f(x)$ on $\{0,1\}^n$, we say that $g(x, y)$ is a *quadratization* of f if $g(x, y)$ is a quadratic polynomial depending on x and on m *auxiliary variables* $y_1, \ldots, y_m$, such that

$$f(x) = \min\{g(x, y) : y \in \{0,1\}^m\} \quad \forall x \in \{0,1\}^n.$$

Then, $\min\{f(x) : x \in \{0,1\}^n\} = \min\{g(x, y) : x \in \{0,1\}^n, y \in \{0,1\}^m\}$.

Which quadratizations are "good"?

- Small number of auxiliary variables.
- Good optimization properties: submodularity.
    - A quadratic pseudo-Boolean function is submodular *iff* all quadratic terms have non-positive coefficients.

Quadratic pseudo-Boolean optimization methods

Quadratic optimization is $\mathcal{NP}$-hard, but much work has been done:

- Algorithms based on **MAX-CUT** (which reduces to a polynomial MIN-CUT problem when f is submodular).

Quadratic pseudo-Boolean optimization methods

Quadratic optimization is $\mathcal{NP}$-hard, but much work has been done:

- Algorithms based on **MAX-CUT** (which reduces to a polynomial MIN-CUT problem when f is submodular).
- **Heuristics** such as local search.

Quadratic pseudo-Boolean optimization methods

Quadratic optimization is $\mathcal{NP}$-hard, but much work has been done:

- Algorithms based on **MAX-CUT** (which reduces to a polynomial MIN-CUT problem when f is submodular).
- **Heuristics** such as local search.
- In computer vision: approaches based on **Roof Duality** (1984) ([8]).

Quadratic pseudo-Boolean optimization methods

Quadratic optimization is $\mathcal{NP}$-hard, but much work has been done:

- Algorithms based on **MAX-CUT** (which reduces to a polynomial MIN-CUT problem when f is submodular).
- **Heuristics** such as local search.
- In computer vision: approaches based on **Roof Duality** (1984) ([8]).
- **Polyhedral results**: Isomorphism between *boolean quadric polytope* (associated to quadratic pseudo-Boolean optimization) and *cut polytope* (associated to MAX-CUT) (1990) ([4]), good separation algorithms...

Quadratic pseudo-Boolean optimization methods

Quadratic optimization is $\mathcal{NP}$-hard, but much work has been done:

- Algorithms based on **MAX-CUT** (which reduces to a polynomial MIN-CUT problem when f is submodular).
- **Heuristics** such as local search.
- In computer vision: approaches based on **Roof Duality** (1984) ([8]).
- **Polyhedral results**: Isomorphism between *boolean quadric polytope* (associated to quadratic pseudo-Boolean optimization) and *cut polytope* (associated to MAX-CUT) (1990) ([4]), good separation algorithms...

Rosenberg

Rosenberg (1975) [11]: 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.

Rosenberg

Rosenberg (1975) [11]: 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.

Rosenberg

Rosenberg (1975) [11]: 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.

Termwise quadratizations

Multilinear expression of a pseudo-Boolean function:

$$f(x) = \sum_{S \in 2^{[n]}} a_S \prod_{i \in S} x_i$$

Idea: quadratize monomial by monomial, using different sets of auxiliary variables for each monomial.

Termwise quadratizations: negative monomials

Kolmogorov and Zabih [10], Freedman and Drineas [6].

$$a \prod_{i=1}^{n} x_i = \min_{y \in \{0,1\}} ay \left(\sum_{i=1}^{n} x_i - (n-1) \right), a < 0.$$

- **Advantages:** one single auxiliary variable, submodular quadratization.

Termwise quadratizations: negative monomials

Kolmogorov and Zabih [10], Freedman and Drineas [6].

$$a \prod_{i=1}^{n} x_i = \min_{y \in \{0,1\}} ay \left(\sum_{i=1}^{n} x_i - (n-1) \right), a < 0.$$

- **Advantages:** one single auxiliary variable, submodular quadratization.

Termwise quadratizations: positive monomials

Ishikawa [9]

$$a \prod_{i=1}^{d} x_i = a \min_{y_1,\dots y_{n_d} \in \{0,1\}} \sum_{i=1}^{n_d} y_i(c_{i,d}(-S_1 + 2i) - 1) + aS_2,$$

S_1, S_2: elementary linear and quadratic symmetric polynomials in d variables, $\mathbf{n_d} = \lfloor \frac{\mathbf{d-1}}{\mathbf{2}} \rfloor$ and $c_{i,d} = \begin{cases} 1, \text{if } d \text{ is odd and } i = n_d, \\ 2, \text{otherwise.} \end{cases}$

- **Number of variables:** best known bound for positive monomials.
- **Submodularity:** $\binom{d}{2}$ positive quadratic terms, but very good computational results.

Termwise quadratizations: positive monomials

Ishikawa [9]

$$a\prod_{i=1}^{d} x_i = a \min_{y_1,\dots y_{n_d}\in\{0,1\}} \sum_{i=1}^{n_d} y_i(c_{i,d}(-S_1+2i)-1)+aS_2,$$

S_1, S_2: elementary linear and quadratic symmetric polynomials in d variables, $\mathbf{n_d} = \lfloor \frac{\mathbf{d-1}}{\mathbf{2}} \rfloor$ and $c_{i,d} = \begin{cases} 1, \text{if } d \text{ is odd and } i = n_d, \\ 2, \text{otherwise.} \end{cases}$

- **Number of variables:** best known bound for positive monomials.
- **Submodularity:** $\binom{d}{2}$ positive quadratic terms, but very good computational results.

Number of variables

Using termwise quadratizations:

- One variable per negative monomial and $\lfloor \frac{d-1}{2} \rfloor$ per positive monomial (d: degree of the monomial).
- Best known upper bounds: $O(n^d)$ variables for a polynomial of fixed degree d, $O(n2^n)$ for an arbitrary function.

Can we do better?

Tight upper and lower bounds *independent of the quadratization procedure* by Anthony, Boros, Crama and Gruber [1]

- $\Theta(2^{\frac{n}{2}})$ for a general pseudo-Boolean function.
- $\Theta(n^{\frac{d}{2}})$ for a fixed polynomial of degree d.

Number of variables

Using termwise quadratizations:

- One variable per negative monomial and $\lfloor \frac{d-1}{2} \rfloor$ per positive monomial (d: degree of the monomial).
- Best known upper bounds: $O(n^d)$ variables for a polynomial of fixed degree d, $O(n2^n)$ for an arbitrary function.

Can we do better?

Tight upper and lower bounds *independent of the quadratization procedure* by Anthony, Boros, Crama and Gruber [1]

- $\Theta(2^{\frac{n}{2}})$ for a general pseudo-Boolean function.
- $\Theta(n^{\frac{d}{2}})$ for a fixed polynomial of degree d.

Note: bounds are polynomial in the *size of the input*:
$f(x) = \sum_{S \in 2^{[n]}} a_S \prod_{i \in S} x_i$ can have up to 2^n monomials.

Number of variables

Using termwise quadratizations:

- One variable per negative monomial and $\lfloor \frac{d-1}{2} \rfloor$ per positive monomial (d: degree of the monomial).
- Best known upper bounds: $O(n^d)$ variables for a polynomial of fixed degree d, $O(n2^n)$ for an arbitrary function.

Can we do better?

Tight upper and lower bounds *independent of the quadratization procedure* by Anthony, Boros, Crama and Gruber [1]

- $\Theta(2^{\frac{n}{2}})$ for a general pseudo-Boolean function.
- $\Theta(n^{\frac{d}{2}})$ for a fixed polynomial of degree d.

Note: bounds are polynomial in the *size of the input*:
$f(x) = \sum_{S \in 2^{[n]}} a_S \prod_{i \in S} x_i$ can have up to 2^n monomials.

Summary: known quadratization techniques

What has been done until now?

Specific quadratization procedures, experimental evaluations...

- Several quadratizations known for monomials.

Summary: known quadratization techniques

What has been done until now?

Specific quadratization procedures, experimental evaluations...

- Several quadratizations known for monomials.
  - Case of negative monomials well-solved (one auxiliary variable, submodular).

Summary: known quadratization techniques

What has been done until now?

Specific quadratization procedures, experimental evaluations...

- Several quadratizations known for monomials.
  - Case of negative monomials well-solved (one auxiliary variable, submodular).
  - Improvements can perhaps be done for positive monomials.

Summary: known quadratization techniques

What has been done until now?

Specific quadratization procedures, experimental evaluations...

- Several quadratizations known for monomials.
  - Case of negative monomials well-solved (one auxiliary variable, submodular).
  - Improvements can perhaps be done for positive monomials.
- Non-termwise quadratization techniques: reduce degree of several terms at once (Fix, Gruber, Boros, Zabih [5]).

Summary: known quadratization techniques

What has been done until now?

Specific quadratization procedures, experimental evaluations...

- Several quadratizations known for monomials.
  - Case of negative monomials well-solved (one auxiliary variable, submodular).
  - Improvements can perhaps be done for positive monomials.
- Non-termwise quadratization techniques: reduce degree of several terms at once (Fix, Gruber, Boros, Zabih [5]).

Objective: Global perspective

Systematic study of quadratizations, understand properties and structure.

Summary: known quadratization techniques

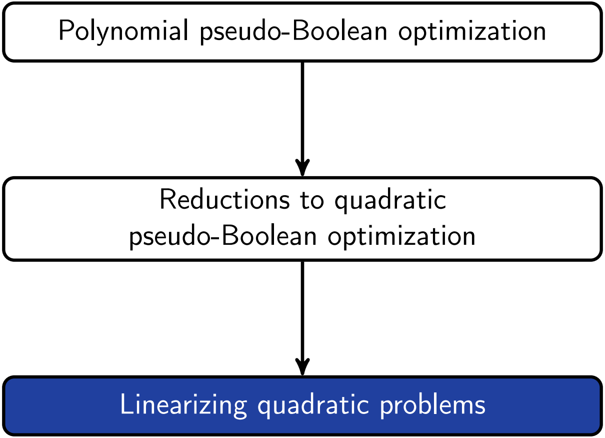
What has been done until now?

Specific quadratization procedures, experimental evaluations...

- Several quadratizations known for monomials.
  - Case of negative monomials well-solved (one auxiliary variable, submodular).
  - Improvements can perhaps be done for positive monomials.
- Non-termwise quadratization techniques: reduce degree of several terms at once (Fix, Gruber, Boros, Zabih [5]).

Objective: Global perspective

Systematic study of quadratizations, understand properties and structure.


Polynomial pseudo-Boolean optimization
Reductions to quadratic pseudo-Boolean optimization
Linearizing quadratic problems

Standard linearization

Original polynomial problem

$$\min_{x \in \{0,1\}^n} f(x) = \sum_{S \in \mathcal{S}} a_S \prod_{i \in S} x_i$$

$\mathcal{S}$: set of non-constant monomials.

1. Substitute monomials

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

$$\text{s.t. } z_S = \prod_{i \in S} z_i, \ \forall S \in \mathcal{S}$$

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

Standard linearization

Original polynomial problem

$$\min_{x\in\{0,1\}^n} f(x) = \sum_{S\in\mathcal{S}} a_S \prod_{i\in S} x_i$$

$\mathcal{S}$: set of non-constant monomials.

1. Substitute monomials

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

$$\text{s.t. } z_S = \prod_{i\in S} z_i, \ \forall S \in \mathcal{S}$$

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

2. Linearize constraints

$$\min_{z_S} \sum_{S\in\mathcal{S}} a_S z_S \qquad \text{(A)}$$

$$\text{s.t. } z_S \leq z_i, \ \forall i \in S, \forall S \in \mathcal{S}$$

$$z_S \geq \sum_{i\in S} z_i - |S| + 1, \ \forall S \in \mathcal{S}$$

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

Standard linearization

Original polynomial problem

$$\min_{x\in\{0,1\}^n} f(x) = \sum_{S\in\mathcal{S}} a_S \prod_{i\in S} x_i$$

$\mathcal{S}$: set of non-constant monomials.

1. Substitute monomials

$$\begin{aligned} \min_{z_S} & \sum_{S\in\mathcal{S}} a_S z_S \\ \text{s.t. } & z_S = \prod_{i\in S} z_i, \ \forall S \in \mathcal{S} \\ & z_S \in \{0,1\}, \ \forall S \in \mathcal{S} \end{aligned}$$

2. Linearize constraints

$$\min_{z_S} \sum_{S\in\mathcal{S}} a_S z_S \qquad \text{(A)}$$

$$\begin{aligned} \text{s.t. } & z_S \le z_i, \ \forall i \in S, \forall S \in \mathcal{S} \\ & z_S \ge \sum_{i\in S} z_i - |S| + 1, \ \forall S \in \mathcal{S} \\ & z_S \in \{0,1\}, \ \forall S \in \mathcal{S} \end{aligned}$$

Linearization of a quadratization

Original polynomial problem

$$\min_{x\in\{0,1\}^n} f(x) = \sum_{S\in\mathcal{S}} a_S \prod_{i\in S} x_i$$

$\mathcal{S}$: set of non-constant monomials

1. Define quadratization for f

$$\min_{x\in\{0,1\}^{n+m}} \sum_{Q\in\mathcal{Q}} b_Q \prod_{i\in Q} x_i$$

where $\mathcal{Q}$ is the set of non-constant monomials in the original $\{x_1,\ldots,x_n\}$ and the auxiliary $\{x_{n+1},\ldots,x_{n+m}\}$ variables and all $Q\in\mathcal{Q}$ have degree at most 2.

Linearization of a quadratization

Original polynomial problem

$$\min_{x\in\{0,1\}^n} f(x) = \sum_{S\in\mathcal{S}} a_S \prod_{i\in S} x_i$$

$\mathcal{S}$: set of non-constant monomials

1. Define quadratization for f

$$\min_{x\in\{0,1\}^{n+m}} \sum_{Q\in\mathcal{Q}} b_Q \prod_{i\in Q} x_i$$

where $\mathcal{Q}$ is the set of non-constant monomials in the original $\{x_1, \ldots, x_n\}$ and the auxiliary $\{x_{n+1}, \ldots, x_{n+m}\}$ variables and all $Q \in \mathcal{Q}$ have degree at most 2.

Linearization of a quadratization

2. Substitute monomials

$$\min_{w_Q} \sum_{Q \in \mathcal{Q}} b_Q w_Q$$

$$\text{s.t. } w_Q = \prod_{i \in Q} w_i, \ \forall Q \in \mathcal{Q}$$

$$w_Q \in \{0, 1\}, \ \forall Q \in \mathcal{Q}$$

3. Linearize constraints

$$\min_{w_Q} \sum_{Q \in \mathcal{Q}} b_Q w_Q \quad \text{(B)}$$

$$\text{s.t. } w_Q \leq w_i, \ \forall i \in Q, \forall Q \in \mathcal{Q}$$

$$w_Q \geq \sum_{i \in Q} w_i - |Q| + 1, \ \forall Q \in \mathcal{Q}$$

$$w_Q \in \{0, 1\}, \ \forall Q \in \mathcal{Q}$$

Linearization of a quadratization

2. Substitute monomials

$$\min_{w_Q} \sum_{Q \in \mathcal{Q}} b_Q w_Q$$

$$\text{s.t. } w_Q = \prod_{i \in Q} w_i, \ \forall Q \in \mathcal{Q}$$

$$w_Q \in \{0,1\}, \ \forall Q \in \mathcal{Q}$$

3. Linearize constraints

$$\min_{w_Q} \sum_{Q \in \mathcal{Q}} b_Q w_Q \quad \text{(B)}$$

$$\text{s.t. } w_Q \leq w_i, \ \forall i \in Q, \forall Q \in \mathcal{Q}$$

$$w_Q \geq \sum_{i \in Q} w_i - |Q| + 1, \ \forall Q \in \mathcal{Q}$$

$$w_Q \in \{0,1\}, \ \forall Q \in \mathcal{Q}$$

Comparing relaxations of linearizations

Relaxation of standard linearization (A)

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

$$\text{s.t. } z_S \leq z_i, \; \forall i \in S, \forall S \in \mathcal{S}$$

$$z_S \geq \sum_{i \in S} z_i - |S| + 1, \; \forall S \in \mathcal{S}$$

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

Relaxation of linearized quadratization (B)

$$\min_{w_Q} \sum_{Q \in \mathcal{Q}} b_Q w_Q$$

$$\text{s.t. } w_Q \leq w_i, \; \forall i \in Q, \forall Q \in \mathcal{Q}$$

$$w_Q \geq \sum_{i \in Q} w_i - |Q| + 1, \; \forall Q \in \mathcal{Q}$$

$$0 \leq w_Q \leq 1, \; \forall Q \in \mathcal{Q}$$

Comparing relaxations of linearizations

Questions:

- Which relaxation is tighter?

Comparing relaxations of linearizations

Questions:

- Which relaxation is tighter?
- Relaxation (B) also depends on the chosen quadratization method, which quadratization gives better relaxations?

Comparing relaxations of linearizations

Questions:

- Which relaxation is tighter?
- Relaxation (B) also depends on the chosen quadratization method, which quadratization gives better relaxations?
- What happens if we intersect the constraints of the relaxed linearizations of *all* quadratizations of f?

Comparing relaxations of linearizations

Questions:

- Which relaxation is tighter?
- Relaxation (B) also depends on the chosen quadratization method, which quadratization gives better relaxations?
- What happens if we intersect the constraints of the relaxed linearizations of *all* quadratizations of f?

Comparing polytopes at substitution step

Standard linearization (A)

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

$$\text{s.t. } z_S = \prod_{i \in S} z_i, \ \forall S \in \mathcal{S}$$

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

Quadratization (B)

$$\min_{w_Q} \sum_{Q \in \mathcal{Q}} b_Q w_Q$$

$$\text{s.t. } w_Q = \prod_{i \in Q} w_i, \ \forall Q \in \mathcal{Q}$$

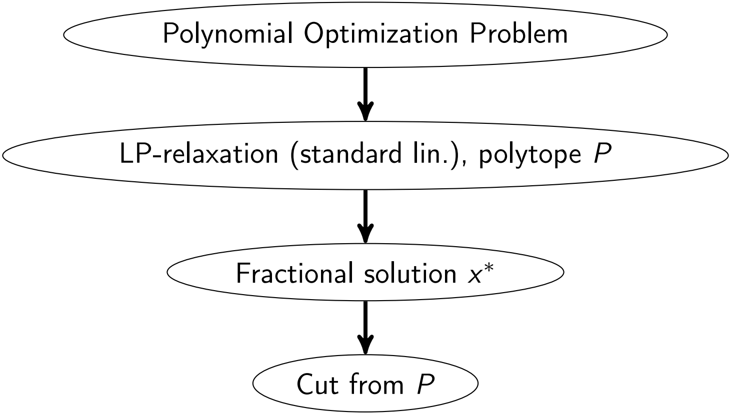
$$w_Q \in \{0, 1\}, \ \forall Q \in \mathcal{Q}$$

Comparing polytopes at substitution step

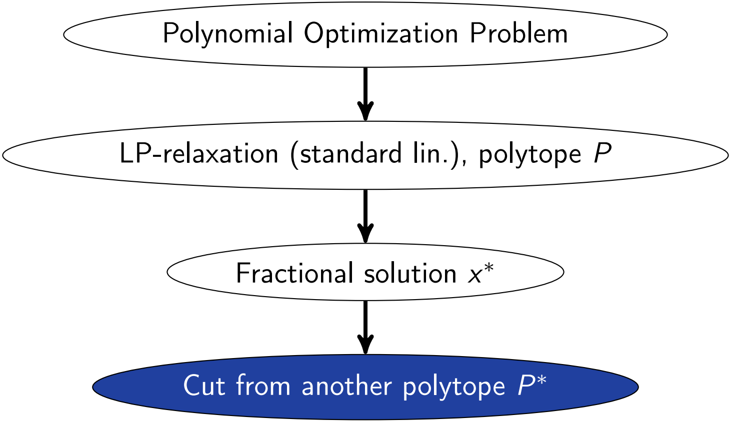
Questions:

- Do we have a better knowledge of one of the convex hull of feasible solutions of one of these problems? (i.e., polyhedral description, good separation algorithms...).

Buchheim and Rinaldi's approach



Buchheim and Rinaldi's approach



Buchheim and Rinaldi's approach: Polytope P^*?

Presented in [2], [3].

Original polynomial problem

$$\min_{x \in \{0,1\}^n} f(x) = \sum_{S \in \mathcal{S}} a_S \prod_{i \in S} x_i$$

$\mathcal{S}$: set of non-constant monomials

- **Assumption:** every $S \in \mathcal{S}$ can be written as the union of two *other* monomials $S_l, S_r \in \mathcal{S}$.

Buchheim and Rinaldi's approach: Polytope P^*?

Presented in [2], [3].

Original polynomial problem

$$\min_{x \in \{0,1\}^n} f(x) = \sum_{S \in \mathcal{S}} a_S \prod_{i \in S} x_i$$

$\mathcal{S}$: set of non-constant monomials

- **Assumption:** every $S \in \mathcal{S}$ can be written as the union of two *other* monomials $S_l, S_r \in \mathcal{S}$.
- For a given S, there might be several pairs of subsets S_l, S_r such that $S_l \cup S_r = S$.

Buchheim and Rinaldi's approach: Polytope P^*?

Presented in [2], [3].

Original polynomial problem

$$\min_{x\in\{0,1\}^n} f(x) = \sum_{S\in\mathcal{S}} a_S \prod_{i\in S} x_i$$

$\mathcal{S}$: set of non-constant monomials

- **Assumption:** every $S \in \mathcal{S}$ can be written as the union of two *other* monomials $S_l, S_r \in \mathcal{S}$.
- For a given S, there might be several pairs of subsets S_l, S_r such that $S_l \cup S_r = S$.
- Set of monomials can be "completed" heuristically if necessary.

Buchheim and Rinaldi's approach: Polytope P^*?

Presented in [2], [3].

Original polynomial problem

$$\min_{x\in\{0,1\}^n} f(x) = \sum_{S\in\mathcal{S}} a_S \prod_{i\in S} x_i$$

$\mathcal{S}$: set of non-constant monomials

- **Assumption:** every $S \in \mathcal{S}$ can be written as the union of two *other* monomials $S_l, S_r \in \mathcal{S}$.
- For a given S, there might be several pairs of subsets S_l, S_r such that $S_l \cup S_r = S$.
- Set of monomials can be "completed" heuristically if necessary.

Buchheim and Rinaldi's approach: Polytope P^*?

Buchheim and Rinaldi's formulation over a quadric polytope

Consider the set $\mathcal{S}^* = \{\{S, T\} \mid S, T \in \mathcal{S} \text{ and } S \cup T \in \mathcal{S}\}$.

$$\min_{y_{\{S\}}} \sum_{S \in \mathcal{S}} a_S y_{\{S\}} \tag{C}$$

$$\text{s.t. } y_{\{S,T\}} = y_{\{S\}} y_{\{T\}}, \ \forall \{S, T\} \in \mathcal{S}^*$$

$$y_{\{S,T\}} \in \{0, 1\}, \ \forall \{S, T\} \in \mathcal{S}^*$$

P^*: convex hull of feasible solutions of problem (C)

Observation: P^* is a boolean quadric polytope.

Buchheim and Rinaldi's approach: polytope P^*?

Polytope P^*

- Is isomorph to a Cut-Polytope (efficient separation algorithms are known).
- Lives in a higher-dimensional space (introducing auxiliary variables).

Buchheim and Rinaldi's approach: polytope P^*?

Polytope P^*

- Is isomorph to a Cut-Polytope (efficient separation algorithms are known).
- Lives in a higher-dimensional space (introducing auxiliary variables).
- P is isomorph to a face of P^* (imposing some additional constraints to P^*): **cutting on P^* is equivalent to cutting on P.**

Buchheim and Rinaldi's approach: polytope P^*?

Polytope P^*

- Is isomorph to a Cut-Polytope (efficient separation algorithms are known).
- Lives in a higher-dimensional space (introducing auxiliary variables).
- P is isomorph to a face of P^* (imposing some additional constraints to P^*): **cutting on P^* is equivalent to cutting on P**.

Objective:

We are interested in understanding the quadratic problem that is used to define P^*:

Buchheim and Rinaldi's approach: polytope P^*?

Polytope P^*

- Is isomorph to a Cut-Polytope (efficient separation algorithms are known).
- Lives in a higher-dimensional space (introducing auxiliary variables).
- P is isomorph to a face of P^* (imposing some additional constraints to P^*): **cutting on P^* is equivalent to cutting on P**.

Objective:

We are interested in understanding the quadratic problem that is used to define P^*:

- Is it a quadratization in our sense?

Buchheim and Rinaldi's approach: polytope P^*?

Polytope P^*

- Is isomorph to a Cut-Polytope (efficient separation algorithms are known).
- Lives in a higher-dimensional space (introducing auxiliary variables).
- P is isomorph to a face of P^* (imposing some additional constraints to P^*): **cutting on P^* is equivalent to cutting on P**.

Objective:

We are interested in understanding the quadratic problem that is used to define P^*:

- Is it a quadratization in our sense?

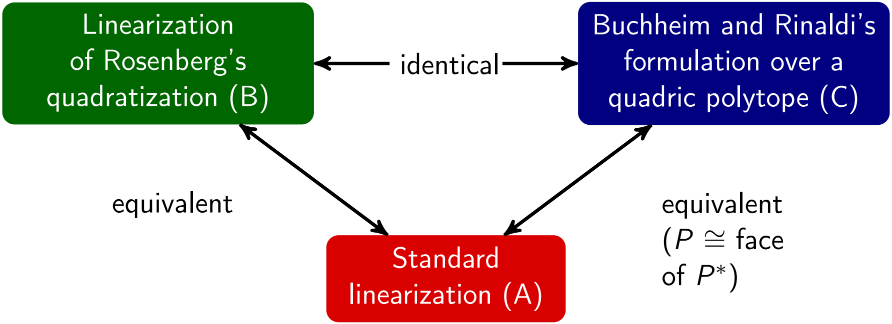
Buchheim and Rinaldi's formulation and Rosenberg's quadratization

Theorem

Buchheim and Rinaldi's formulation over a quadric polytope can be obtained (up to elimination of redundant constraints) by linearizing a variant of Rosenberg's quadratization where:

- the order of substituting variables is induced by the decomposition S_l, S_r of each monomial S, and
- when substituting a product by a variable, we do not impose $y_{ij} = x_i x_j$ with a penalty, but with a constraint.

Assumption: every $S \in \mathcal{S}$ can be written as the union of two *other* monomials $S_l, S_r \in \mathcal{S}$.


Linearization of Rosenberg's quadratization (B)
identical
Buchheim and Rinaldi's formulation over a quadric polytope (C)
equivalent
Standard linearization (A)
equivalent ($P \cong$ face of P^*)

Conclusions

- Understand quadratization methods from a global perspective:
  - Properties and structure, which quadratizations are "better"?

Conclusions

- Understand quadratization methods from a global perspective:
  - Properties and structure, which quadratizations are “better”?
  - Describe *all* quadratizations of f.

Conclusions

- Understand quadratization methods from a global perspective:
  - Properties and structure, which quadratizations are "better"?
  - Describe *all* quadratizations of f.
- Linearizing quadratizations:

Conclusions

- Understand quadratization methods from a global perspective:
  - Properties and structure, which quadratizations are "better"?
  - Describe *all* quadratizations of f.
- Linearizing quadratizations:
  - How does the tightness of a relaxation depend on the quadratization?

Conclusions

- Understand quadratization methods from a global perspective:
  - Properties and structure, which quadratizations are "better"?
  - Describe *all* quadratizations of f.
- Linearizing quadratizations:
  - How does the tightness of a relaxation depend on the quadratization?
  - How do relaxations of linearized quadratizations relate to other methods (e.g. standard linearization)?

Conclusions

- Understand quadratization methods from a global perspective:
  - Properties and structure, which quadratizations are "better"?
  - Describe *all* quadratizations of f.
- Linearizing quadratizations:
  - How does the tightness of a relaxation depend on the quadratization?
  - How do relaxations of linearized quadratizations relate to other methods (e.g. standard linearization)?
  - Can we use the knowledge about the polytopes associated to linearizations? (e.g. cut polytope separation techniques...)?

Conclusions

- Understand quadratization methods from a global perspective:
  - Properties and structure, which quadratizations are "better"?
  - Describe *all* quadratizations of f.
- Linearizing quadratizations:
  - How does the tightness of a relaxation depend on the quadratization?
  - How do relaxations of linearized quadratizations relate to other methods (e.g. standard linearization)?
  - Can we use the knowledge about the polytopes associated to linearizations? (e.g. cut polytope separation techniques...)?

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. Buchheim and G. Rinaldi. Terse integer linear programs for boolean optimization. Journal on Satisfiability, Boolean Modeling and Computation, 6:121–139, 2009.
- C. De Simone. The cut polytope and the boolean quadric polytope. Discrete Mathematics, 79(1):71–75, 1990.
- A. Fix, A. Gruber, E. Boros, and R. Zabih. A hypergraph-based reduction for higher-order markov random fields. Working paper, submitted to PAMI?, 2014.

Some references II

- D. Freedman and P. Drineas. Energy minimization via graph cuts: settling what is possible. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 2, pages 939–946, June 2005.
- 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.
- P. L. Hammer, P. Hansen, and B. Simeone. Roof duality, complementation and persistency in quadratic 0-1 optimization. Mathematical Programming, 28(2):121–155, 1984.
- H. Ishikawa. Transformation of general binary mrf minimization to the first-order case. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 33(6):1234–1249, June 2011.

Some references III

- V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? Pattern Analysis and Machine Intelligence, IEEE Transactions on, 26(2):147–159, Feb 2004.
- I. G. Rosenberg. Reduction of bivalent maximization to the quadratic case. Cahiers du Centre d'Etudes de Recherche Opérationnelle, 17:71–74, 1975.
- S. Živnỳ, D. A. Cohen, and P. G. Jeavons. The expressive power of binary submodular functions. Discrete Applied Mathematics, 157(15):3347–3358, 2009.