Structure of Quadratizations of Pseudo-Boolean Functions

Aritanan Gruber

Assistant Professor
Center for Mathematics, Computing, and Cognition
Federal University of ABC
Santo André, SP, Brazil

Joint work with
Martin Anthony (LSE, London, UK)
Endre Boros (Rutgers University, NJ, USA)
Yves Crama (University of Liège, Belgium)
Elisabeth Rodriguez-Heck (RWTH Aachen, Germany)

Latin America School of Mathematics (ELAM)
Santo André, SP, Brazil – September 05, 2018

Agenda

- Introduction to pseudo-Boolean functions and optimization
- Present some frontier results on quadratizations of PBFs
- Discuss some open problems and future directions

Quadratizations of Pseudo-Boolean Functions

In a nutshell:

Quadratic reformulations (resorting to binary auxiliary variables) of unconstrained, nonlinear binary optimization problems specified through high-degree multilinear polynomials.

Objective:

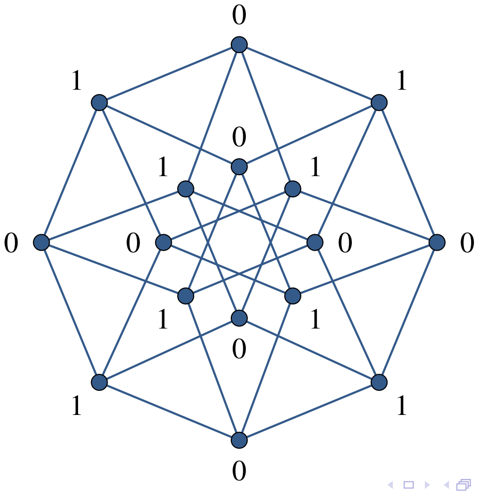
To allow their continuous relaxations to be handled though maximum flow techniques, so that:

- some variables can be fixed to their provably optimal values (*persistences*),
- and allow NP-hard problems to be decomposed into smaller, disjoint subproblems.

Boolean Functions

Binary-valued mappings of binary vectors:

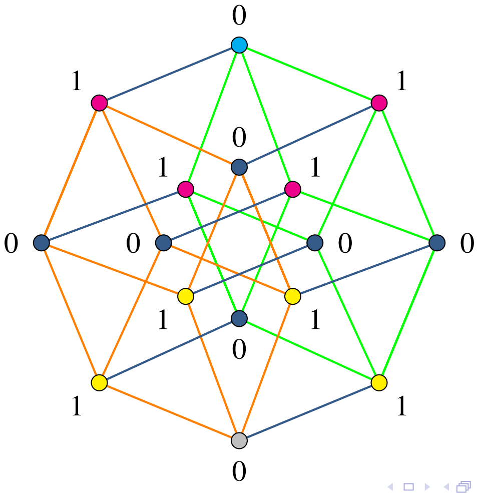
$$\mathscr{B}^n = \{0,1\}^{\{0,1\}^n} := \{\, h : \{0,1\}^n \to \{0,1\} \}$$



Boolean Functions

Parity function:

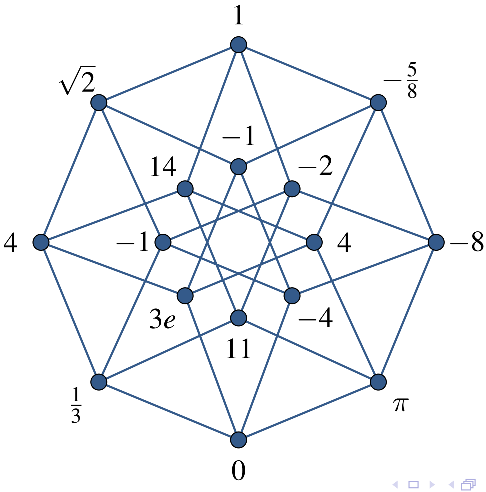
$$\bigoplus_4(x_1, x_2, x_3, x_4) = (x_1 + x_2 + x_3 + x_4) \pmod 2$$



Pseudo-Boolean Functions

Real-valued mappings of binary vectors:

$$\mathscr{P}^n = \mathbb{R}^{\{0,1\}^n} := \left\{ f : \{0,1\}^n \to \mathbb{R} \right\}$$



Pseudo-Boolean Functions

Theorem (Hammer and Rudeanu '68):

Each PBF $f : \{0,1\}^n \to \mathbb{R}$ can be uniquely represented as a *multilinear polynomial* in the binary variables $x_1, x_2, \ldots, x_n$:

$$f(x_1, x_2, \ldots, x_n) = \sum_{S \subseteq [n]} c_S \prod_{j \in S} x_j,$$

where $c_S \in \mathbb{R}$ for all $S \subseteq [n] := \{1, 2, \ldots, n\}$.

Example

$$f(x_1, x_2, x_3) = 3 + x_1 - 4x_2 + x_3 - 5x_1x_2 + 9x_1x_3 + 2x_2x_3 - 6x_1x_2x_3$$

Pseudo-Boolean Functions

Proof:

Existence by interpolation: for each $\mathbf{a} \in \{0,1\}^n$,

$$\chi_{\mathbf{a}}(\mathbf{x}) := \prod_{i:a_i=1} x_i \cdot \prod_{j:a_j=0} (1-x_j).$$

As $\chi_{\mathbf{a}}(\mathbf{x}) = 1 \iff \mathbf{x} = \mathbf{a}$,

$$f(\mathbf{x}) = P_f(\mathbf{x}) := \sum_{\mathbf{a}\in\{0,1\}^n} f(\mathbf{a})\chi_{\mathbf{a}}(\mathbf{x}) = \sum_{S\subseteq[n]} c_S \prod_{i\in S} x_i.$$

Uniqueness: $\{\chi_{\mathbf{a}} : \mathbf{a} \in \{0,1\}^n\}$ is a basis for $\mathscr{P}^n$. □

Degree of $f : \{0,1\} \to \mathbb{R}$:

$$\deg(f) = \deg(P_f) := \max\{|S| : S \subseteq [n], c_S \neq 0\}.$$

Pseudo-Boolean Functions

Proof:

Existence by interpolation: for each $\mathbf{a} \in \{0,1\}^n$,

$$\chi_{\mathbf{a}}(\mathbf{x}) := \prod_{i:a_i=1} x_i \cdot \prod_{j:a_j=0} (1 - x_j).$$

As $\chi_{\mathbf{a}}(\mathbf{x}) = 1 \iff \mathbf{x} = \mathbf{a}$,

$$f(\mathbf{x}) = P_f(\mathbf{x}) : = \sum_{\mathbf{a} \in \{0,1\}^n} f(\mathbf{a})\chi_{\mathbf{a}}(\mathbf{x}) = \sum_{S \subseteq [n]} c_S \prod_{i \in S} x_i.$$

Uniqueness: $\left\{\chi_{\mathbf{a}} : \mathbf{a} \in \{0,1\}^n\right\}$ is a basis for $\mathscr{P}^n$. □

Degree of $f : \{0,1\} \to \mathbb{R}$:

$$\deg(f) = \deg(P_f) := \max\left\{|S| : S \subseteq [n],\ c_S \neq 0\right\}.$$

Pseudo-Boolean Optimization

Unrestricted problem:

Determine

$$f(\mathbf{x}^*) = \min\{f(\mathbf{x}) : \mathbf{x} \in \{0,1\}^n\}$$

for a PBF $f : \{0,1\}^n \to \mathbb{R}$ in multilinear polynomial form.

Applications in many fields:

- Artificial Intelligence / Machine Learning.
- Combinatorics / Graph Theory / Coding Theory.
- Computational Complexity / Quantum Computing.
- Computer Vision / Integer Programming / VLSI Design.
- Operations Research / Game Theory / Social Choice.

Pseudo-Boolean Optimization

Unrestricted problem:

Determine

$$f(\mathbf{x}^*) = \min\{f(\mathbf{x}) : \mathbf{x} \in \{0,1\}^n\}$$

for a PBF $f : \{0,1\}^n \to \mathbb{R}$ in multilinear polynomial form.

Applications in many fields:

- Artificial Intelligence / Machine Learning.
- Combinatorics / Graph Theory / Coding Theory.
- Computational Complexity / Quantum Computing.
- Computer Vision / Integer Programming / VLSI Design.
- Operations Research / Game Theory / Social Choice.

Pseudo-Boolean Optimization

Complexity:

- Polytime(n) solution if f is:

  submodular, *unimodular*, *signed balanced*.

- NP-hard in general, even when f is quadratic:

  encompasses MAX-SAT, MAX-CUT, MAX-STABLESET.

- Constant-factor approximation if f is *supermodular*; no approximation possible in some settings, unless P = NP.

Pseudo-Boolean Optimization

Complexity:

- Polytime(n) solution if f is:

  submodular, *unimodular*, *signed balanced*.

- NP-hard in general, even when f is quadratic:

  encompasses MAX-SAT, MAX-CUT, MAX-STABLESET.

- Constant-factor approximation if f is *supermodular*; no approximation possible in some settings, unless P $=$ NP.

Pseudo-Boolean Optimization

Complexity:

- Polytime(n) solution if f is:

  submodular, *unimodular*, *signed balanced*.

- NP-hard in general, even when f is quadratic:

  encompasses MAX-SAT, MAX-CUT, MAX-STABLESET.

- Constant-factor approximation if f is *supermodular*; no approximation possible in some settings, unless P $=$ NP.

Pseudo-Boolean Optimization

Submodularity:

- PBF $f : \{0,1\}^n \to \mathbb{R}$ is submodular if for all $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$

$$f(\mathbf{x}) + f(\mathbf{y}) \geq f(\mathbf{x} \wedge \mathbf{y}) + f(\mathbf{x} \vee \mathbf{y}),$$

  where $\mathbf{x} \odot \mathbf{y} = (x_1 \odot y_1, \ldots, x_n \odot y_n)$, with $\odot \in \{\wedge, \vee\}$.

- In other words, f has the *diminishing returns* property.
- $\min f(\mathbf{x})$ available in strongly polytime:
  - Grötschel, Lováz, Schrijver '81: $O(n^8 + \gamma n^7)$,
  - Schrijver '00: $O(n^8 + \gamma n^7)$,
  - Iwata, Fleisher, Fujishige '01: $O(\gamma n^7 \log n)$,
  - Lee, Siford, Wong '15: $O(\gamma n^3 \log^2 n + n^4 \log^{O(1)} n)$.
- Chakrabarty, Lee, Siford, Wong '17: $O(\gamma n M^3 \log n) = o(n^2)$ for integer-valued f. Remains to be evaluated.

Pseudo-Boolean Optimization

Submodularity:

- PBF $f : \{0,1\}^n \to \mathbb{R}$ is submodular if for all $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$

$$f(\mathbf{x}) + f(\mathbf{y}) \geq f(\mathbf{x} \wedge \mathbf{y}) + f(\mathbf{x} \vee \mathbf{y}),$$

  where $\mathbf{x} \odot \mathbf{y} = (x_1 \odot y_1, \ldots, x_n \odot y_n)$, with $\odot \in \{\wedge, \vee\}$.

- In other words, f has the *diminishing returns* property.

- $\min f(\mathbf{x})$ available in strongly polytime:
  - Grötschel, Lováz, Schrijver '81: $O(n^8 + \gamma n^7)$,
  - Schrijver '00: $O(n^8 + \gamma n^7)$,
  - Iwata, Fleisher, Fujishige '01: $O(\gamma n^7 \log n)$,
  - Lee, Siford, Wong '15: $O(\gamma n^3 \log^2 n + n^4 \log^{O(1)} n)$.

- Chakrabarty, Lee, Siford, Wong '17: $O(\gamma n M^3 \log n) = o(n^2)$ for integer-valued f. Remains to be evaluated.

Pseudo-Boolean Optimization

Submodularity:

- PBF $f : \{0,1\}^n \to \mathbb{R}$ is submodular if for all $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$

$$f(\mathbf{x}) + f(\mathbf{y}) \geq f(\mathbf{x} \wedge \mathbf{y}) + f(\mathbf{x} \vee \mathbf{y}),$$

  where $\mathbf{x} \odot \mathbf{y} = (x_1 \odot y_1, \ldots, x_n \odot y_n)$, with $\odot \in \{\wedge, \vee\}$.

- In other words, f has the *diminishing returns* property.
- $\min f(\mathbf{x})$ available in strongly polytime:
  - Grötschel, Lováz, Schrijver '81: $O(n^8 + \gamma n^7)$,
  - Schrijver '00: $O(n^8 + \gamma n^7)$,
  - Iwata, Fleisher, Fujishige '01: $O(\gamma n^7 \log n)$,
  - Lee, Siford, Wong '15: $O(\gamma n^3 \log^2 n + n^4 \log^{O(1)} n)$.
- Chakrabarty, Lee, Siford, Wong '17: $O(\gamma n M^3 \log n) = o(n^2)$ for integer-valued f. Remains to be evaluated.

Pseudo-Boolean Optimization

Submodularity:

- PBF $f : \{0,1\}^n \to \mathbb{R}$ is submodular if for all $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$

$$f(\mathbf{x}) + f(\mathbf{y}) \geq f(\mathbf{x} \wedge \mathbf{y}) + f(\mathbf{x} \vee \mathbf{y}),$$

  where $\mathbf{x} \odot \mathbf{y} = (x_1 \odot y_1, \ldots, x_n \odot y_n)$, with $\odot \in \{\wedge, \vee\}$.

- In other words, f has the *diminishing returns* property.
- $\min f(\mathbf{x})$ available in strongly polytime:
  - Grötschel, Lováz, Schrijver '81: $O(n^8 + \gamma n^7)$,
  - Schrijver '00: $O(n^8 + \gamma n^7)$,
  - Iwata, Fleisher, Fujishige '01: $O(\gamma n^7 \log n)$,
  - Lee, Siford, Wong '15: $O(\gamma n^3 \log^2 n + n^4 \log^{O(1)} n)$.
- Chakrabarty, Lee, Siford, Wong '17: $O(\gamma n M^3 \log n) = o(n^2)$ for integer-valued f. Remains to be evaluated.

Quadratic Pseudo-Boolean Optimization

Submodularity:

- Quadratic PBF f is submodular if and only if all quadratic coefficients are nonpositive.
- Similar, albeit slightly more involved, result if f is cubic.
- co-NP-hard to decide if f is submodular when $\deg(f) \geq 4$.

Theorem (Hammer '65)): submodular QPBO $\in$ P

1-to-1 correspondence between values of f and (s,t)-cut values of a network G_f, whenever f is a submodular QPBF.

Rediscovered later inside the computer vision community: Boykov, Veksler, Zabih '99 and Kolmogorov, Zabih '02.

Quadratic Pseudo-Boolean Optimization

Submodularity:

- Quadratic PBF f is submodular if and only if all quadratic coefficients are nonpositive.
- Similar, albeit slightly more involved, result if f is cubic.
- co-NP-hard to decide if f is submodular when $\deg(f) \geq 4$.

Theorem (Hammer '65)): submodular QPBO $\in$ P

1-to-1 correspondence between values of f and (s,t)-cut values of a network G_f, whenever f is a submodular QPBF.

Rediscovered later inside the computer vision community: Boykov, Veksler, Zabih '99 and Kolmogorov, Zabih '02.

Quadratic Pseudo-Boolean Optimization

Submodularity:

- Quadratic PBF f is submodular if and only if all quadratic coefficients are nonpositive.
- Similar, albeit slightly more involved, result if f is cubic.
- co-NP-hard to decide if f is submodular when $\deg(f) \geq 4$.

Theorem (Hammer '65)): submodular QPBO $\in$ P

1-to-1 correspondence between values of f and (s, t)-cut values of a network G_f, whenever f is a submodular QPBF.

Rediscovered later inside the computer vision community: Boykov, Veksler, Zabih '99 and Kolmogorov, Zabih '02.

Quadratic Pseudo-Boolean Optimization

Submodularity:

- Quadratic PBF f is submodular if and only if all quadratic coefficients are nonpositive.
- Similar, albeit slightly more involved, result if f is cubic.
- co-NP-hard to decide if f is submodular when $\deg(f) \geq 4$.

Theorem (Hammer '65)): submodular QPBO $\in$ P

1-to-1 correspondence between values of f and (s, t)-cut values of a network G_f, whenever f is a submodular QPBF.

Rediscovered later inside the computer vision community: Boykov, Veksler, Zabih '99 and Kolmogorov, Zabih '02.

Quadratic Pseudo-Boolean Optimization

Submodularity:

- Quadratic PBF f is submodular if and only if all quadratic coefficients are nonpositive.
- Similar, albeit slightly more involved, result if f is cubic.
- co-NP-hard to decide if f is submodular when $\deg(f) \geq 4$.

Theorem (Hammer '65)): submodular QPBO $\in$ P

1-to-1 correspondence between values of f and (s, t)-cut values of a network G_f, whenever f is a submodular QPBF.

Rediscovered later inside the computer vision community: Boykov, Veksler, Zabih '99 and Kolmogorov, Zabih '02.

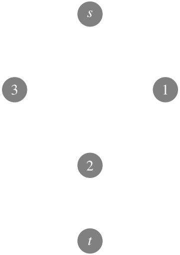
Network Model for Submodular QPBO

- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

$$=$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO



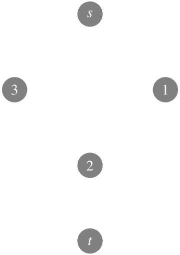
- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

$$\mathbf{f} = 4 - \mathbf{x}_1 + 7\mathbf{x}_2 + \mathbf{x}_3 - 3\mathbf{x}_1\mathbf{x}_2 - \mathbf{x}_1\mathbf{x}_3 - 2\mathbf{x}_2\mathbf{x}_3$$

$$=$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO



- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

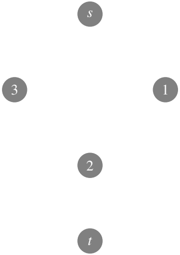
$$\mathbf{f} = 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3$$

$$= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

$$=$$

- There is a 1-to-1 correspondence between values of f and (s, t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO



- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

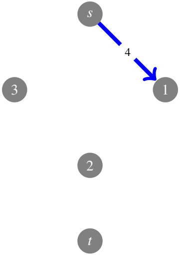
$$\mathbf{f = 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3}$$

$$\mathbf{= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3}$$

$$\mathbf{= 4s\bar{x}_1 + 5x_2\bar{t} + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3}$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO



- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

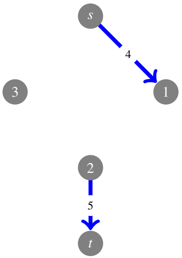
$$\mathbf{f} = 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3$$

$$= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

$$= 4s\bar{x}_1 + 5x_2\bar{t} + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. *(Hammer '65)*

Network Model for Submodular QPBO



- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

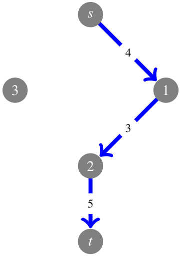
$$\mathbf{f} = 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3$$

$$= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

$$= 4s\bar{x}_1 + 5x_2\bar{t} + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO



- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

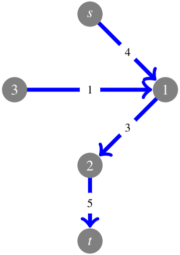
$$\mathbf{f} = 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3$$

$$= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

$$= 4s\bar{x}_1 + 5x_2\bar{t} + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO

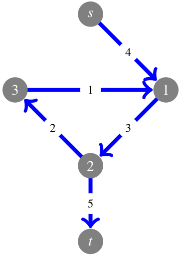


- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

$$
\begin{aligned}
\mathbf{f} &= 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3 \\
&= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3 \\
&= 4s\bar{x}_1 + 5x_2\bar{t} + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3
\end{aligned}
$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO



- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

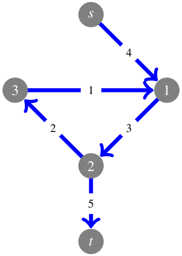
$$\mathbf{f} = 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3$$

$$= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

$$= 4s\bar{x}_1 + 5x_2\bar{t} + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3$$

- There is a 1-to-1 correspondence between values of f and (s, t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO

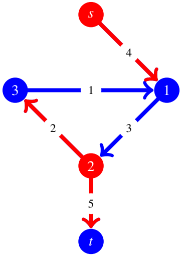


- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

$$\begin{aligned} \mathbf{f} &= 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3 \\ &= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3 \\ &= 4s\bar{x}_1 + 5x_2\bar{t} + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3 \end{aligned}$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. (*Hammer '65*)

Network Model for Submodular QPBO



- A QPBF is submodular IFF all quadratic coefficients are nonpositive.
- To a submodular QPBF f associate a network G_f as follows:

$$\mathbf{f = 4 - x_1 + 7x_2 + x_3 - 3x_1x_2 - x_1x_3 - 2x_2x_3}$$

$$\mathbf{= 4\bar{x}_1 + 5x_2 + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3}$$

$$\mathbf{= 4s\bar{x}_1 + 5x_2\bar{t} + 3x_1\bar{x}_2 + \bar{x}_1x_3 + 2x_2\bar{x}_3}$$

- There is a 1-to-1 correspondence between values of f and (s,t)-cut values of G_f. (*Hammer '65*)

$$\mathbf{f(0,1,0) \;=\; \partial(\{s,2\},\{1,3,t\}) \;=\; 11}$$

Network Model for Non-submodular QPBO

When f is non-submodular:

- Implication Networks: Boros and Hammer '89, Boros, Hammer, Sun '91, Boros, Hammer, Sun, Tavares '06, '08.
- Associates a different network H_f to a QPBF f.
  - Binary vectors $\mathbf{x}$ correspond to cuts in H_f with value $= f(\mathbf{x})$.
  - Max-Flow provides lower bounds for f: *roof duality*.
  - Residual network provides optimal values for some of the variables (*persistencies*) and a decomposition of the residual problem.
- Strongly polynomial preprocessing fixes $60 - 90\%$ of the variables in many applications (especially in computer vision).
- Fast QPBO implementations: Tavares '06; Rother, Kolmogorov, Lempitsky, Szummer '07.

Network Model for Non-submodular QPBO

When f is non-submodular:

- Implication Networks: Boros and Hammer '89, Boros, Hammer, Sun '91, Boros, Hammer, Sun, Tavares '06, '08.
- Associates a different network H_f to a QPBF f.
  - Binary vectors $\mathbf{x}$ correspond to cuts in H_f with value $= f(\mathbf{x})$.
  - Max-Flow provides lower bounds for f: *roof duality*.
  - Residual network provides optimal values for some of the variables (*persistencies*) and a decomposition of the residual problem.
- Strongly polynomial preprocessing fixes $60 - 90\%$ of the variables in many applications (especially in computer vision).
- Fast QPBO implementations: Tavares '06; Rother, Kolmogorov, Lempitsky, Szummer '07.

Network Model for Non-submodular QPBO

When f is non-submodular:

- Implication Networks: Boros and Hammer '89, Boros, Hammer, Sun '91, Boros, Hammer, Sun, Tavares '06, '08.
- Associates a different network H_f to a QPBF f.
  - Binary vectors $\mathbf{x}$ correspond to cuts in H_f with value $= f(\mathbf{x})$.
  - Max-Flow provides lower bounds for f: *roof duality*.
  - Residual network provides optimal values for some of the variables (*persistencies*) and a decomposition of the residual problem.
- Strongly polynomial preprocessing fixes $60 - 90\%$ of the variables in many applications (especially in computer vision).
- Fast QPBO implementations: Tavares '06; Rother, Kolmogorov, Lempitsky, Szummer '07.

Network Model for Non-submodular QPBO

When f is non-submodular:

- Implication Networks: Boros and Hammer '89, Boros, Hammer, Sun '91, Boros, Hammer, Sun, Tavares '06, '08.
- Associates a different network H_f to a QPBF f.
  - Binary vectors $\mathbf{x}$ correspond to cuts in H_f with value $= f(\mathbf{x})$.
  - Max-Flow provides lower bounds for f: *roof duality*.
  - Residual network provides optimal values for some of the variables (*persistencies*) and a decomposition of the residual problem.
- Strongly polynomial preprocessing fixes $60 - 90\%$ of the variables in many applications (especially in computer vision).
- Fast QPBO implementations: Tavares '06; Rother, Kolmogorov, Lempitsky, Szummer '07.

Network Model for Non-submodular QPBO

When f is non-submodular:

- Implication Networks: Boros and Hammer '89, Boros, Hammer, Sun '91, Boros, Hammer, Sun, Tavares '06, '08.
- Associates a different network H_f to a QPBF f.
  - Binary vectors $\mathbf{x}$ correspond to cuts in H_f with value $= f(\mathbf{x})$.
  - Max-Flow provides lower bounds for f: *roof duality*.
  - Residual network provides optimal values for some of the variables (*persistencies*) and a decomposition of the residual problem.
- Strongly polynomial preprocessing fixes $60 - 90\%$ of the variables in many applications (especially in computer vision).
- Fast QPBO implementations: Tavares '06; Rother, Kolmogorov, Lempitsky, Szummer '07.

Network Model for Non-submodular QPBO

When f is non-submodular:

- Implication Networks: Boros and Hammer '89, Boros, Hammer, Sun '91, Boros, Hammer, Sun, Tavares '06, '08.
- Associates a different network H_f to a QPBF f.
  - Binary vectors $\mathbf{x}$ correspond to cuts in H_f with value $= f(\mathbf{x})$.
  - Max-Flow provides lower bounds for f: *roof duality*.
  - Residual network provides optimal values for some of the variables (*persistencies*) and a decomposition of the residual problem.
- Strongly polynomial preprocessing fixes $60 - 90\%$ of the variables in many applications (especially in computer vision).
- Fast QPBO implementations: Tavares '06; Rother, Kolmogorov, Lempitsky, Szummer '07.

Network Model for Non-submodular QPBO

When f is non-submodular:

- Implication Networks: Boros and Hammer '89, Boros, Hammer, Sun '91, Boros, Hammer, Sun, Tavares '06, '08.
- Associates a different network H_f to a QPBF f.
    - Binary vectors $\mathbf{x}$ correspond to cuts in H_f with value $= f(\mathbf{x})$.
    - Max-Flow provides lower bounds for f: *roof duality*.
    - Residual network provides optimal values for some of the variables (*persistencies*) and a decomposition of the residual problem.
- Strongly polynomial preprocessing fixes $60-90\%$ of the variables in many applications (especially in computer vision).
- Fast QPBO implementations: Tavares '06; Rother, Kolmogorov, Lempitsky, Szummer '07.

Pseudo-Boolean Optimization

For higher degree PBFs:

No combinatorial algorithms for degree larger than 3, and no persistence results for degree larger than 2.

Quadratizations of Pseudo-Boolean Functions

Quadratizations:

PBF $g : \{0,1\}^{n+m} \to \mathbb{R}$ is an *m-quadratization* of $f : \{0,1\}^n \to \mathbb{R}$ if g is **quadratic** and

$$f(x) = \min\{g(x,y) : y \in \{0,1\}^m\} \quad \text{for all} \quad \mathbf{x} \in \{0,1\}^n;$$

g is *y-linear* if there are no monomials of the form

$$y_i y_j \quad \text{for} \quad i,j \in [m],\ i \neq j.$$

The y-variables are called *auxiliary* variables.

Example

$$-x_1x_2x_3 = \min_{y_1}\{2y_1 - x_1y_1 - x_2y_1 - x_3y_1\}$$

$$= \min_{y_1,y_2}\{x_3 + 2y_2 - x_1x_3 + x_1y_1 - x_1y_2 - x_2y_2 - x_3y_1 - y_1y_2\}$$

Quadratizations of Pseudo-Boolean Functions

Quadratizations:

PBF $g : \{0,1\}^{n+m} \to \mathbb{R}$ is an *m-quadratization* of $f : \{0,1\}^n \to \mathbb{R}$ if g is **quadratic** and

$$f(x) = \min\big\{g(x,y) : y \in \{0,1\}^m\big\} \quad \text{for all} \quad \mathbf{x} \in \{0,1\}^n;$$

g is *y-linear* if there are no monomials of the form

$$y_i y_j \quad \text{for} \quad i,j \in [m],\ i \neq j.$$

The y-variables are called *auxiliary* variables.

Example

$$\begin{aligned} -x_1x_2x_3 &= \min_{y_1}\big\{2y_1 - x_1y_1 - x_2y_1 - x_3y_1\big\} \\ &= \min_{y_1,y_2}\big\{x_3 + 2y_2 - x_1x_3 + x_1y_1 - x_1y_2 - x_2y_2 - x_3y_1 - y_1y_2\big\} \end{aligned}$$

Quadratizations of Pseudo-Boolean Functions

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

Goals:

- Keep m small!
- Have g "as submodular as possible!"
- With no large coefficients!

- Rosenberg '75: All PBFs have quadratizations.

Quadratizations of Pseudo-Boolean Functions

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

Goals:

- Keep m small!
- Have g "as submodular as possible!"
- With no large coefficients!

- Rosenberg '75: All PBFs have quadratizations.

Quadratizations of Pseudo-Boolean Functions

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

Goals:

- Keep m small!
- Have g "as submodular as possible!"
- With no large coefficients!

- Rosenberg '75: All PBFs have quadratizations.

Quadratizations of Pseudo-Boolean Functions

$$f(\mathbf{x}) = \min\left\{g(\mathbf{x},\mathbf{y}) : \mathbf{y} \in \{0,1\}^m\right\} \quad \text{for all} \quad \mathbf{x} \in \{0,1\}^n.$$

Goals:

- Keep m small!
- Have g "as submodular as possible!"
- With no large coefficients!

- Rosenberg '75: All PBFs have quadratizations.

Quadratizations of Pseudo-Boolean Functions

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

Goals:

- Keep m small!
- Have g “as submodular as possible!”
- With no large coefficients!

- Rosenberg ’75: All PBFs have quadratizations.

Rosenberg's Penalty Function Method '75

- Let $p(x, y, w) \;:=\; xy - 2xw - 2yw + 3w \begin{cases} = 0 & \text{if } w = xy, \\ \geq 1 & \text{if } w \neq xy. \end{cases}$
- For $M > 0$ large enough,

$$f(x, y, \ldots) \;=\; xyA + B \;=\; \min_{w \in \{0,1\}} wA + B + Mp(x, y, w)$$

Example

$$\begin{aligned} f(x_1, x_2, x_3, x_4, x_5) &:= 5x_1x_2 - 7x_1x_2x_3x_4 + 2x_1x_2x_3x_5 \\ &= 45y_1 + 45y_2 + 20x_1x_2 - 30x_1x_6 - 30x_2x_6 \\ &+ 15x_3y_1 - 30x_3y_2 - 7x_4y_2 + 2x_5y_2 - 30y_1y_2 \\ &=: g(x_1, x_2, x_3, x_4, x_5, y_1, y_2) \end{aligned}$$

$$\min f(x) = \min g(x, y) = -2$$

Rosenberg's Penalty Function Method '75

- Let $p(x,y,w) \;:=\; xy - 2xw - 2yw + 3w \begin{cases} = 0 & \text{if } w = xy, \\ \geq 1 & \text{if } w \neq xy. \end{cases}$
- For $M > 0$ large enough,

$$f(x,y,\ldots) \;=\; xyA + B \;=\; \min_{w\in\{0,1\}} wA + B + Mp(x,y,w)$$

Example

$$\begin{aligned} f(x_1,x_2,x_3,x_4,x_5) &:= 5x_1x_2 - 7x_1x_2x_3x_4 + 2x_1x_2x_3x_5 \\ &= 45y_1 + 45y_2 + 20x_1x_2 - 30x_1x_6 - 30x_2x_6 \\ &\quad + 15x_3y_1 - 30x_3y_2 - 7x_4y_2 + 2x_5y_2 - 30y_1y_2 \\ &=: g(x_1,x_2,x_3,x_4,x_5,y_1,y_2) \end{aligned}$$

$$\min f(x) = \min g(x,y) = -2$$

Rosenberg's Penalty Function Method '75

- Let $p(x,y,w) \;:=\; xy - 2xw - 2yw + 3w \begin{cases} = 0 & \text{if } w = xy, \\ \geq 1 & \text{if } w \neq xy. \end{cases}$
- For $M > 0$ large enough,

$$f(x,y,\ldots) \;=\; xyA + B \;=\; \min_{w\in\{0,1\}} wA + B + Mp(x,y,w)$$

Example

$$\begin{aligned} f(x_1,x_2,x_3,x_4,x_5) &:= 5x_1x_2 - 7x_1x_2x_3x_4 + 2x_1x_2x_3x_5 \\ &= 45y_1 + 45y_2 + 20x_1x_2 - 30x_1x_6 - 30x_2x_6 \\ &+ 15x_3y_1 - 30x_3y_2 - 7x_4y_2 + 2x_5y_2 - 30y_1y_2 \\ &=: g(x_1,x_2,x_3,x_4,x_5,y_1,y_2) \end{aligned}$$

$\min f(x) = \min g(x,y) = -2$

Rosenberg's Penalty Function Method '75

Possible drawbacks:

- Many positive quadratic terms with large coefficients, even if the input f is submodular!
- Does not generalize for the substition of 3 or more variables at once.
- Introduces many auxiliary variables in the worst case: $O(n^n)$.
- Can we find better estimates? $O(n^n)$ seems ludicrous!
- NP-hard to find a quadratization with a minimum number of auxiliary variables (reduction from MAX-STABLESET).

Rosenberg's Penalty Function Method '75

Possible drawbacks:

- Many positive quadratic terms with large coefficients, even if the input f is submodular!
- Does not generalize for the substition of 3 or more variables at once.
- Introduces many auxiliary variables in the worst case: $O(n^n)$.
- Can we find better estimates? $O(n^n)$ seems ludicrous!
- NP-hard to find a quadratization with a minimum number of auxiliary variables (reduction from MAX-STABLESET).

Rosenberg's Penalty Function Method '75

Possible drawbacks:

- Many positive quadratic terms with large coefficients, even if the input f is submodular!
- Does not generalize for the substition of 3 or more variables at once.
- Introduces many auxiliary variables in the worst case: $O(n^n)$.
- Can we find better estimates? $O(n^n)$ seems ludicrous!
- NP-hard to find a quadratization with a minimum number of auxiliary variables (reduction from MAX-STABLESET).

Rosenberg's Penalty Function Method '75

Possible drawbacks:

- Many positive quadratic terms with large coefficients, even if the input f is submodular!
- Does not generalize for the substition of 3 or more variables at once.
- Introduces many auxiliary variables in the worst case: $O(n^n)$.
- Can we find better estimates? $O(n^n)$ seems ludicrous!
- NP-hard to find a quadratization with a minimum number of auxiliary variables (reduction from MAX-STABLESET).

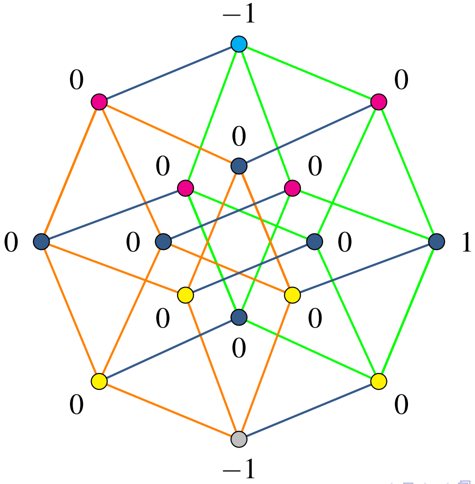
Rosenberg's Penalty Function Method '75

Possible drawbacks:

- Many positive quadratic terms with large coefficients, even if the input f is submodular!
- Does not generalize for the substition of 3 or more variables at once.
- Introduces many auxiliary variables in the worst case: $O(n^n)$.
- Can we find better estimates? $O(n^n)$ seems ludicrous!
- NP-hard to find a quadratization with a minimum number of auxiliary variables (reduction from MAX-STABLESET).

Živný, Cohen, Jeavons '09

Not all submodular PBFs have submodular quadratizations!
Quasi-indecomposable function $\theta_t(w, x, y, z)$ for $t = (1, 1, 0, 0)$.



Higher-Degree Substitution for Negative Terms

Kolmogorov and Zabih '04, Fredman and Drineas '05:

Only one auxiliary variable, no positive quadratic term, no large coefficient:

$$N(\mathbf{x}) := -\prod_{i=1}^{n} x_i = \min_{y\in\{0,1\}} y\left(n-1-\sum_{i=1}^{n} x_i\right).$$

Proof:

It holds that

$$-x_1x_2\cdots x_n \leq 0 \quad \text{and} \quad -x_1x_2\cdots x_n \leq (n-1)-\sum_{j=1}^{n} x_j.$$

RHS correspond to $y=0$ and $y=1$ above, respectively, and so $y=\prod_{j=1}^{d} x_j$ at a minimum. □

Higher-Degree Substitution for Negative Terms

Negating variables:

For binary variables $x_1, x_2, \ldots, x_n$ and y, it holds that

$$-\bar{x}_1\bar{x}_2\cdots\bar{x}_n = \min_{y\in\{0,1\}} y\left((n-1) - \sum_{j=1}^{n}\bar{x}_j\right)$$

$$= -1 + \sum_{j=1}^{n} x_j + \min_{\bar{y}\in\{0,1\}} \bar{y}\left(1 - \sum_{j=1}^{n} x_j\right).$$

Rother, Kohli, Feng and Jia '09:

$$-\prod_{j\in N}\bar{x}_j\prod_{j\in P}x_j = \min_{u,v\in\{0,1\}} -uv + u\sum_{j\in N}x_j + v\sum_{j\in P}\bar{x}_j$$

Two auxiliary variables, $|N|$ positive quadratic terms, no large coefficients.

Higher-Degree Substitution for Negative Terms

Negating variables:

For binary variables $x_1, x_2, \ldots, x_n$ and y, it holds that

$$-\bar{x}_1\bar{x}_2\cdots\bar{x}_n = \min_{y\in\{0,1\}} y\left((n-1) - \sum_{j=1}^{n}\bar{x}_j\right)$$
$$= -1 + \sum_{j=1}^{n} x_j + \min_{\bar{y}\in\{0,1\}} \bar{y}\left(1 - \sum_{j=1}^{n} x_j\right).$$

Rother, Kohli, Feng and Jia '09:

$$-\prod_{j\in N}\bar{x}_j\prod_{j\in P}x_j = \min_{u,v\in\{0,1\}} -uv + u\sum_{j\in N}x_j + v\sum_{j\in P}\bar{x}_j$$

Two auxiliary variables, $|N|$ positive quadratic terms, no large coefficients.

Characterizing Negative Terms

Theorem: 1-Quadratizations of negative monomials

Up to a permutation of the $\mathbf{x}$-variables and to a switch of the y-variable,

$$s_n = \left(n - 1 - \sum_{i=1}^{n} x_i\right) y$$

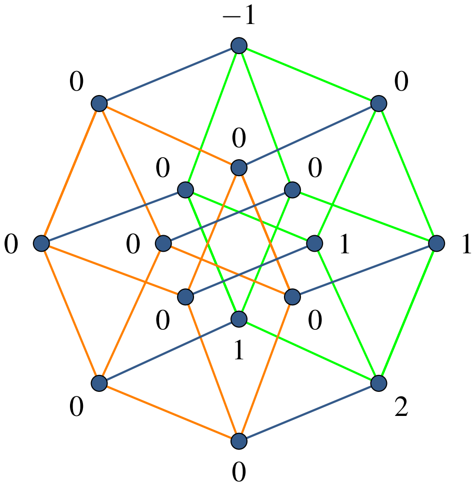
$$s_n^+ = (n-2)x_n y - \sum_{i=1}^{n-1} x_i(y - \bar{x}_n)$$

are the only *prime* 1-quadratizations of $N_n := -\prod_{i=1}^{n} x_i$.

Proof. Long and technical… □

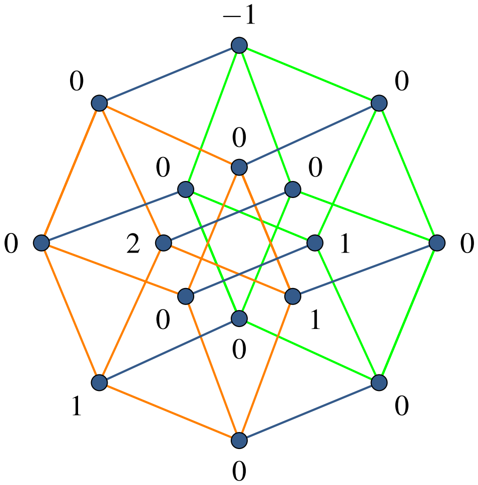
Characterizing Negative Terms

$$s_3(w; x, y, z) = 2w - xw - yw - zw$$



Characterizing Negative Terms

$$s_3^+(w; x, y, z) = x + y - xz - yz - xw - yw + zw$$



What About Positive Terms, $P_n(\mathbf{x}) := \prod_{i=1}^n x_i$?

- Can be written as $P_n(\mathbf{x}) = -x_1 \cdots x_{n-1}\bar{x}_n + P_{n-1}(\mathbf{x})$.
- So P_n can be quadratized using $n-2$ additional variables.

Ishikawa '09, '11:

$k = \left\lfloor \frac{n-1}{2} \right\rfloor$ auxiliary variables suffice:

$$P_n(\mathbf{x}) = \sum_{1 \le i < j \le n} x_i x_j + \min_{y \in \{0,1\}^k} \sum_{j=1}^{k} y_j \left(c_{j,n} \left(-\sum_{i=1}^{n} x_i + 2j \right) - 1 \right),$$

where $c_{j,n} = 1$ if $n = j$ is odd, and $c_{j,n} = 2$ otherwise.

No large coefficients, $\binom{n}{2}$ positive quadratic terms.

Every PBF has a quadratization with at most $O(n2^n)$ auxiliary variables. No characterization for P_n is known!

What About Positive Terms, $P_n(\mathbf{x}) := \prod_{i=1}^n x_i$?

- Can be written as $P_n(\mathbf{x}) = -x_1 \cdots x_{n-1}\bar{x}_n + P_{n-1}(\mathbf{x})$.
- So P_n can be quadratized using $n-2$ additional variables.

Ishikawa '09, '11:

$k = \left\lfloor \frac{n-1}{2} \right\rfloor$ auxiliary variables suffice:

$$P_n(\mathbf{x}) = \sum_{1\le i<j\le n} x_i x_j + \min_{y\in\{0,1\}^k} \sum_{j=1}^{k} y_j \left(c_{j,n}\left(-\sum_{i=1}^{n} x_i + 2j\right) - 1\right),$$

where $c_{j,n} = 1$ if $n = j$ is odd, and $c_{j,n} = 2$ otherwise.

No large coefficients, $\binom{n}{2}$ positive quadratic terms.

Every PBF has a quadratization with at most $O(n2^n)$ auxiliary variables. No characterization for P_n is known!

What About Positive Terms, $P_n(\mathbf{x}) := \prod_{i=1}^n x_i$?

- Can be written as $P_n(\mathbf{x}) = -x_1 \cdots x_{n-1}\bar{x}_n + P_{n-1}(\mathbf{x})$.
- So P_n can be quadratized using $n-2$ additional variables.

Ishikawa '09, '11:

$k = \left\lfloor \frac{n-1}{2} \right\rfloor$ auxiliary variables suffice:

$$P_n(\mathbf{x}) = \sum_{1\le i<j\le n} x_i x_j + \min_{y\in\{0,1\}^k} \sum_{j=1}^k y_j \left(c_{j,n} \left(-\sum_{i=1}^n x_i + 2j \right) - 1 \right),$$

where $c_{j,n} = 1$ if $n = j$ is odd, and $c_{j,n} = 2$ otherwise.

No large coefficients, $\binom{n}{2}$ positive quadratic terms.

Every PBF has a quadratization with at most $O(n2^n)$ auxiliary variables. No characterization for P_n is known!

What About Positive Terms, $P_n(\mathbf{x}) := \prod_{i=1}^{n} x_i$?

- Can be written as $P_n(\mathbf{x}) = -x_1 \cdots x_{n-1}\bar{x}_n + P_{n-1}(\mathbf{x})$.
- So P_n can be quadratized using $n-2$ additional variables.

Ishikawa '09, '11:

$k = \left\lfloor \frac{n-1}{2} \right\rfloor$ auxiliary variables suffice:

$$P_n(\mathbf{x}) = \sum_{1 \le i < j \le n} x_i x_j + \min_{y \in \{0,1\}^k} \sum_{j=1}^{k} y_j \left(c_{j,n} \left(-\sum_{i=1}^{n} x_i + 2j \right) - 1 \right),$$

where $c_{j,n} = 1$ if $n = j$ is odd, and $c_{j,n} = 2$ otherwise.

No large coefficients, $\binom{n}{2}$ positive quadratic terms.

Every PBF has a quadratization with at most $O(n2^n)$ auxiliary variables. No characterization for P_n is known!

Multiple Split of Terms

Assume that $\phi_i(\mathbf{w}) \in \{0,1\}$ for $i \in [q]$, $\mathbf{w} \in \{0,1\}^p$ such that

$$\min_{\mathbf{w}\in\{0,1\}^p} \sum_{i=1}^{q} \phi_i(w) \; = \; 1, \qquad \text{and}$$

$$\forall I \subset [q] \quad \exists \mathbf{w}^* \in \{0,1\}^p \quad \text{s.t.} \quad \sum_{i \in I} \phi_i(\mathbf{w}^*) = 0.$$

For instance $\phi_1 = w_1$, $\phi_2 = w_2$, and $\phi_3 = \overline{w}_1\overline{w}_2$ is such a system.

Theorem

If P_i, $i \in [q]$, are subsets of indices covering $[d]$, then

$$\prod_{j=1}^{d} x_j \; = \; \min_{\mathbf{w}\in\{0,1\}^p} \sum_{i=1}^{q} \phi_i(w) \prod_{j \in P_i} x_j.$$

With $p = \lceil \log q \rceil$ new variables we can split a degree $d = kq$ term into q terms of degree $k + p$.

Multiple Split of Terms

Assume that $\phi_i(\mathbf{w}) \in \{0,1\}$ for $i \in [q]$, $\mathbf{w} \in \{0,1\}^p$ such that

$$\min_{\mathbf{w}\in\{0,1\}^p} \sum_{i=1}^{q} \phi_i(w) \;=\; 1, \qquad \text{and}$$

$$\forall I \subset [q] \quad \exists \mathbf{w}^* \in \{0,1\}^p \quad \text{s.t.} \quad \sum_{i\in I} \phi_i(\mathbf{w}^*) = 0.$$

For instance $\phi_1 = w_1$, $\phi_2 = w_2$, *and* $\phi_3 = \overline{w}_1\overline{w}_2$ *is such a system.*

Theorem

If P_i, $i \in [q]$, *are subsets of indices covering* $[d]$, *then*

$$\prod_{j=1}^{d} x_j \;=\; \min_{\mathbf{w}\in\{0,1\}^p} \sum_{i=1}^{q} \phi_i(w) \prod_{j\in P_i} x_j.$$

With $p = \lceil \log q \rceil$ new variables we can split a degree $d = kq$ term into q terms of degree $k + p$.

Multiple Split of Terms

Assume that $\phi_i(\mathbf{w}) \in \{0,1\}$ for $i \in [q]$, $\mathbf{w} \in \{0,1\}^p$ such that

$$\min_{\mathbf{w}\in\{0,1\}^p} \sum_{i=1}^{q} \phi_i(w) \;=\; 1, \qquad \text{and}$$

$$\forall I \subset [q] \quad \exists \mathbf{w}^* \in \{0,1\}^p \quad \text{s.t.} \quad \sum_{i\in I} \phi_i(\mathbf{w}^*) = 0.$$

For instance $\phi_1 = w_1$, $\phi_2 = w_2$, *and* $\phi_3 = \overline{w}_1\overline{w}_2$ *is such a system.*

Theorem

If P_i, $i \in [q]$, *are subsets of indices covering* $[d]$, *then*

$$\prod_{j=1}^{d} x_j \;=\; \min_{\mathbf{w}\in\{0,1\}^p} \sum_{i=1}^{q} \phi_i(w) \prod_{j\in P_i} x_j.$$

With $p = \lceil \log q \rceil$ new variables we can split a degree $d = kq$ term into q terms of degree $k + p$.

Multiple Split of Terms

Assume that $\phi_i(\mathbf{w}) \in \{0,1\}$ for $i \in [q]$, $\mathbf{w} \in \{0,1\}^p$ such that

$$\min_{\mathbf{w}\in\{0,1\}^p} \sum_{i=1}^{q} \phi_i(w) \; = \; 1, \qquad \text{and}$$

$$\forall I \subset [q] \quad \exists \mathbf{w}^* \in \{0,1\}^p \quad \text{s.t.} \quad \sum_{i\in I} \phi_i(\mathbf{w}^*) = 0.$$

For instance $\phi_1 = w_1$, $\phi_2 = w_2$, *and* $\phi_3 = \overline{w}_1\overline{w}_2$ *is such a system.*

Theorem

If P_i, $i \in [q]$, *are subsets of indices covering* $[d]$, *then*

$$\prod_{j=1}^{d} x_j \; = \; \min_{\mathbf{w}\in\{0,1\}^p} \sum_{i=1}^{q} \phi_i(w) \prod_{j\in P_i} x_j.$$

With $p = \lceil \log q \rceil$ new variables we can split a degree $d = kq$ term into q terms of degree $k + p$.

Splitting Off Common Parts

Let $C \subseteq [n]$, $\mathcal{H} \subseteq 2^{[n]\setminus C}$, and consider the following fragment of a pseudo-Boolean function:

$$g(x) = \sum_{H \in \mathcal{H}} \alpha_H \prod_{j \in C \cup H} x_j, \quad \text{with} \quad \alpha_H \geq 0.$$

Theorem (Set of Positive Terms)

$$g(x) = \min_{w \in \{0,1\}} \left(\sum_{H \in \mathcal{H}} \alpha_H \right) w \prod_{j \in C} x_j + \sum_{H \in \mathcal{H}} \alpha_H \overline{w} \prod_{j \in H} x_j.$$

Theorem (Set of Negative Terms)

$$-g(x) = \min_{w \in \{0,1\}} \sum_{H \in \mathcal{H}} \alpha_H w \left(1 - \prod_{j \in C} x_j - \prod_{j \in H} x_j \right).$$

Splitting Off Common Parts

Let $C \subseteq [n]$, $\mathcal{H} \subseteq 2^{[n]\setminus C}$, and consider the following fragment of a pseudo-Boolean function:

$$g(x) = \sum_{H \in \mathcal{H}} \alpha_H \prod_{j \in C \cup H} x_j, \quad \text{with} \quad \alpha_H \geq 0.$$

Theorem (Set of Positive Terms)

$$g(x) = \min_{w \in \{0,1\}} \left(\sum_{H \in \mathcal{H}} \alpha_H \right) w \prod_{j \in C} x_j + \sum_{H \in \mathcal{H}} \alpha_H \overline{w} \prod_{j \in H} x_j.$$

Theorem (Set of Negative Terms)

$$-g(x) = \min_{w \in \{0,1\}} \sum_{H \in \mathcal{H}} \alpha_H w \left(1 - \prod_{j \in C} x_j - \prod_{j \in H} x_j \right).$$

Splitting Off Common Parts

Let $C \subseteq [n]$, $\mathcal{H} \subseteq 2^{[n]\setminus C}$, and consider the following fragment of a pseudo-Boolean function:

$$g(x) = \sum_{H \in \mathcal{H}} \alpha_H \prod_{j \in C \cup H} x_j, \quad \text{with} \quad \alpha_H \geq 0.$$

Theorem (Set of Positive Terms)

$$g(x) = \min_{w \in \{0,1\}} \left(\sum_{H \in \mathcal{H}} \alpha_H \right) w \prod_{j \in C} x_j + \sum_{H \in \mathcal{H}} \alpha_H \overline{w} \prod_{j \in H} x_j.$$

Theorem (Set of Negative Terms)

$$-g(x) = \min_{w \in \{0,1\}} \sum_{H \in \mathcal{H}} \alpha_H w \left(1 - \prod_{j \in C} x_j - \prod_{j \in H} x_j \right).$$

Corollary

A PBF in n variables, with t terms of degree d has a quadratization with $\approx n + k\binom{n}{k} + \frac{td}{k}$ new variables and with at most $n-1$ positive quadratic terms, for any $k \geq 1$.

Ishikawa's method provides a quadratization with $\approx n + \frac{td}{2}$ new variables and $\max\{\binom{n}{2}, t\binom{d}{2}\}$ positive quadratic terms.

	New variables	# positive terms	# terms	% fixed
Ishikawa	224,346	421,897	1,133,811	80.4%
Our method	236,806	38,343	677,183	96.1%
Δ	$+6\%$	-90%	-40%	$+20\%$

Figure: Performance comparison of reductions, on Ishikawa's benchmarks. Relative performance of our method is shown as Δ. (Fix, Gruber, Boros, Zabih '11, '15))

Corollary

A PBF in n variables, with t terms of degree d has a quadratization with $\approx n + k\binom{n}{k} + \frac{td}{k}$ new variables and with at most $n-1$ positive quadratic terms, for any $k \geq 1$.

Ishikawa's method provides a quadratization with $\approx n + \frac{td}{2}$ new variables and $\max\{\binom{n}{2}, t\binom{d}{2}\}$ positive quadratic terms.

	New variables	# positive terms	# terms	% fixed
Ishikawa	224,346	421,897	1,133,811	80.4%
Our method	236,806	38,343	677,183	96.1%
Δ	$+6\%$	-90%	-40%	$+20\%$

Figure: Performance comparison of reductions, on Ishikawa's benchmarks. Relative performance of our method is shown as Δ. (Fix, Gruber, Boros, Zabih '11, '15))

Corollary

A PBF in n variables, with t terms of degree d has a quadratization with $\approx n + k\binom{n}{k} + \frac{td}{k}$ new variables and with at most $n-1$ positive quadratic terms, for any $k \geq 1$.

Ishikawa's method provides a quadratization with $\approx n + \frac{td}{2}$ new variables and $\max\{\binom{n}{2}, t\binom{d}{2}\}$ positive quadratic terms.

	New variables	# positive terms	# terms	% fixed
Ishikawa	224,346	421,897	1,133,811	80.4%
Our method	236,806	38,343	677,183	96.1%
Δ	$+6\%$	-90%	-40%	$+20\%$

Figure: Performance comparison of reductions, on Ishikawa's benchmarks. Relative performance of our method is shown as Δ. (Fix, Gruber, Boros, Zabih '11, '15))

Symmetric Pseudo-Boolean Functions

A PBF $f : \{0,1\}^n \to \mathbb{R}$ is *symmetric* if for all permutations $\sigma \in \mathfrak{S}_n$

$$f(x_1, x_2, \ldots, x_n) = f\left(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}\right).$$

Example

- Positive and Negative Monomials
- Majority:

  $\mathrm{Maj}_3(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3 - 2x_1x_2x_3$
- Parity:

  $\bigoplus_3(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 + 4x_1x_2x_3$

Symmetric Pseudo-Boolean Functions

A PBF $f : \{0,1\}^n \to \mathbb{R}$ is *symmetric* if for all permutations $\sigma \in \mathfrak{S}_n$

$$f(x_1, x_2, \ldots, x_n) = f\left(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}\right).$$

That is, $f(x_1, x_2, \ldots, x_n) = k\left(\sum_{i=1}^n x_i\right)$ for some $k : \{0, \ldots, n\} \to \mathbb{R}$.

Example

- Positive and Negative Monomials
- Majority:
  $\mathrm{Maj}_3(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3 - 2x_1x_2x_3$
- Parity:
  $\bigoplus_3(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 + 4x_1x_2x_3$

Symmetric Pseudo-Boolean Functions

A PBF $f : \{0,1\}^n \to \mathbb{R}$ is *symmetric* if for all permutations $\sigma \in \mathfrak{S}_n$

$$f(x_1, x_2, \ldots, x_n) = f\left(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}\right).$$

That is, $f(x_1, x_2, \ldots, x_n) = k\left(\sum_{i=1}^{n} x_i\right)$ for some $k : \{0, \ldots, n\} \to \mathbb{R}$.

Example

- Positive and Negative Monomials
- Majority:

  $\mathrm{Maj}_3(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3 - 2x_1x_2x_3$
- Parity:

  $\bigoplus_3(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 + 4x_1x_2x_3$

Quadratization of Symmetric PBFs

How many auxiliary variables suffice?

- Negative Monomials: 1 (Freedman and Drineas ’05).
- Positive Monomials: $\left\lfloor \frac{n-1}{2} \right\rfloor$ (Ishikawa ’09, ’11).
- $n-1$ auxiliary variables suffice (Fix ’11, ad hoc argument).

- We show that $n-2$ auxiliary variables suffice to y-linear quadratize SPBF.

Quadratization of Symmetric PBFs

How many auxiliary variables suffice?

- Negative Monomials: 1 (Freedman and Drineas '05).
- Positive Monomials: $\left\lfloor \frac{n-1}{2} \right\rfloor$ (Ishikawa '09, '11).
- $n-1$ auxiliary variables suffice (Fix '11, ad hoc argument).

- We show that $n-2$ auxiliary variables suffice to y-linear quadratize SPBF.

Quadratization of Symmetric PBFs

How many auxiliary variables suffice?

- Negative Monomials: 1 (Freedman and Drineas '05).
- Positive Monomials: $\left\lfloor \frac{n-1}{2} \right\rfloor$ (Ishikawa '09, '11).
- $n-1$ auxiliary variables suffice (Fix '11, ad hoc argument).

- We show that $n-2$ auxiliary variables suffice to y-linear quadratize SPBF.

Quadratization of Symmetric PBFs

How many auxiliary variables suffice?

- Negative Monomials: 1 (Freedman and Drineas ’05).
- Positive Monomials: $\left\lfloor \frac{n-1}{2} \right\rfloor$ (Ishikawa ’09, ’11).
- $n - 1$ auxiliary variables suffice (Fix ’11, ad hoc argument).

- We show that $n - 2$ auxiliary variables suffice to y-linear quadratize SPBF.

Quadratization of Symmetric PBFs

Representation Theorem:

For all $0 < \epsilon_i \leq 1$, $i = 0, \dots n$, every SPBF $f : \{0,1\}^n \to \mathbb{R}$ can be uniquely represented in the form

$$f(x) = \sum_{i=0}^{n} \alpha_i \left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^{-},$$

where $[a]^{-} := \min(a, 0)$ for any real value a.

- Idea: $\left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^{-}$ reflects whether $\sum_{j=1}^{n} x_j$ is larger than i.
- Lower triangular system of linear equations: $\alpha_0, \dots, \alpha_n$ can be efficiently computed.

Quadratization of Symmetric PBFs

Representation Theorem:

For all $0 < \epsilon_i \leq 1$, $i = 0, \ldots n$, every SPBF $f : \{0,1\}^n \to \mathbb{R}$ can be uniquely represented in the form

$$f(x) = \sum_{i=0}^{n} \alpha_i \left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^{-},$$

where $[a]^{-} := \min(a, 0)$ for any real value a.

- Idea: $\left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^{-}$ reflects whether $\sum_{j=1}^{n} x_j$ is larger than i.
- Lower triangular system of linear equations: $\alpha_0, \ldots, \alpha_n$ can be efficiently computed.

Quadratization of Symmetric PBFs

Representation Theorem:

For all $0 < \epsilon_i \leq 1$, $i = 0, \ldots, n$, every SPBF $f : \{0,1\}^n \to \mathbb{R}$ can be uniquely represented in the form

$$f(x) = \sum_{i=0}^{n} \alpha_i \left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^{-},$$

where $[a]^- := \min(a, 0)$ for any real value a.

- Idea: $\left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^{-}$ reflects whether $\sum_{j=1}^{n} x_j$ is larger than i.
- Lower triangular system of linear equations: $\alpha_0, \ldots, \alpha_n$ can be efficiently computed.

Quadratization of Symmetric PBFs

So, for a symmetric pseudo-Boolean function f:

$$f(x) = \sum_{i=0}^{n} \alpha_i \left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^-$$

$$\stackrel{?}{=} \min_{y \in \{0,1\}} \sum_{i=0}^{n} \alpha_i \left(i - \epsilon_i - \sum_{j=1}^{n} x_j \right) y_i.$$

Well, not quite: $-[a]^- = -\min(a,0) \neq \min_{y \in \{0,1\}}(-ay)$.

Quadratization of Symmetric PBFs

So, for a symmetric pseudo-Boolean function f:

$$f(x) = \sum_{i=0}^{n} \alpha_i \left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^{-}$$

$$\stackrel{?}{=} \min_{y \in \{0,1\}} \sum_{i=0}^{n} \alpha_i \left(i - \epsilon_i - \sum_{j=1}^{n} x_j \right) y_i.$$

Well, not quite: $-[a]^{-} = -\min(a,0) \neq \min_{y\in\{0,1\}}(-ay)$.

Quadratization of Zeros

- $E(l) = l(l-1) + 2\sum_{i=1}^{n-1} [i-l]^-$,
- $E'(l) = \frac{l(l-1)}{2} + 2\sum_{i=2\,:\,i \text{ even}}^{n} \left[i - \frac{1}{2} - l\right]^-$,
- $E''(l) = \frac{l(l+1)}{2} + 2\sum_{i=1\,:\,i \text{ odd}}^{n} \left[i - \frac{1}{2} - l\right]^-$.

Lemma:

For all $l = 0, \ldots, n$, $E(l) = E'(l) = E''(l) = 0$.

Modify first the representation of f by adding one of $E(\sum x_j)$, $E'(\sum x_j)$, or $E''(\sum x_j)$ so as to cancel negative coefficients.

Quadratization of Zeros

- $E(l) = l(l-1) + 2\sum_{i=1}^{n-1} [i-l]^-$,
- $E'(l) = \frac{l(l-1)}{2} + 2\sum_{i=2\,:\,i \text{ even}}^{n} \left[i - \frac{1}{2} - l\right]^-$,
- $E''(l) = \frac{l(l+1)}{2} + 2\sum_{i=1\,:\,i \text{ odd}}^{n} \left[i - \frac{1}{2} - l\right]^-$.

Lemma:

For all $l = 0, \ldots, n$, $E(l) = E'(l) = E''(l) = 0$.

Modify first the representation of f by adding one of $E(\sum x_j)$, $E'(\sum x_j)$, or $E''(\sum x_j)$ so as to cancel negative coefficients.

Quadratization of Zeros

- $E(l) \;\; = l(l-1) + 2\sum_{i=1}^{n-1} [i-l]^-$,
- $E'(l) \;\; = \frac{l(l-1)}{2} + 2\sum_{i=2\,:\,i \text{ even}}^{n} \left[i - \frac{1}{2} - l\right]^-$,
- $E''(l) = \frac{l(l+1)}{2} + 2\sum_{i=1\,:\,i \text{ odd}}^{n} \left[i - \frac{1}{2} - l\right]^-$.

Lemma:

For all $l = 0, \ldots, n$, $E(l) = E'(l) = E''(l) = 0$.

Modify first the representation of f by adding one of $E(\sum x_j)$, $E'(\sum x_j)$, or $E''(\sum x_j)$ so as to cancel negative coefficients.

Upper Bounds for Symmetric PBFs

Theorem: Positive monomials

The positive monomial P_n has a $\lfloor \frac{n-1}{2} \rfloor$-quadratization ($y$-linear).

Theorem: k-out-of-n and exact-k functions

The k-out-of-n function has a $\lceil \frac{n}{2} \rceil$-quadratization ($y$-linear), and the exact-$k$ function has a $\lfloor \frac{n}{2} \rfloor$-quadratization ($y$-linear).

Theorem: Parity function

The parity function has a $\lfloor \frac{n}{2} \rfloor$-quadratization ($y$-linear), and its complement has a $\lfloor \frac{n-1}{2} \rfloor$-quadratization ($y$-linear).

Theorem: General symmetric pseudo-Boolean functions

Every symmetric pseudo-Boolean function has a y-linear quadratization involving at most $(n-2)$ auxiliary variables.

Upper Bounds for Symmetric PBFs

Theorem: Positive monomials

The positive monomial P_n has a $\left\lfloor \frac{n-1}{2} \right\rfloor$-quadratization ($y$-linear).

Theorem: k-out-of-n and exact-k functions

The k-out-of-n function has a $\left\lceil \frac{n}{2} \right\rceil$-quadratization ($y$-linear), and the exact-$k$ function has a $\left\lfloor \frac{n}{2} \right\rfloor$-quadratization ($y$-linear).

Theorem: Parity function

The parity function has a $\left\lfloor \frac{n}{2} \right\rfloor$-quadratization ($y$-linear), and its complement has a $\left\lfloor \frac{n-1}{2} \right\rfloor$-quadratization ($y$-linear).

Theorem: General symmetric pseudo-Boolean functions

Every symmetric pseudo-Boolean function has a y-linear quadratization involving at most $(n-2)$ auxiliary variables.

Upper Bounds for Symmetric PBFs

Theorem: Positive monomials

The positive monomial P_n has a $\lfloor \frac{n-1}{2} \rfloor$-quadratization ($y$-linear).

Theorem: k-out-of-n and exact-k functions

The k-out-of-n function has a $\lceil \frac{n}{2} \rceil$-quadratization ($y$-linear), and the exact-$k$ function has a $\lfloor \frac{n}{2} \rfloor$-quadratization ($y$-linear).

Theorem: Parity function

The parity function has a $\lfloor \frac{n}{2} \rfloor$-quadratization ($y$-linear), and its complement has a $\lfloor \frac{n-1}{2} \rfloor$-quadratization ($y$-linear).

Theorem: General symmetric pseudo-Boolean functions

Every symmetric pseudo-Boolean function has a y-linear quadratization involving at most $(n-2)$ auxiliary variables.

Upper Bounds for Symmetric PBFs

Theorem: Positive monomials

The positive monomial P_n has a $\lfloor \frac{n-1}{2} \rfloor$-quadratization ($y$-linear).

Theorem: k-out-of-n and exact-k functions

The k-out-of-n function has a $\lceil \frac{n}{2} \rceil$-quadratization ($y$-linear), and the exact-$k$ function has a $\lfloor \frac{n}{2} \rfloor$-quadratization ($y$-linear).

Theorem: Parity function

The parity function has a $\lfloor \frac{n}{2} \rfloor$-quadratization ($y$-linear), and its complement has a $\lfloor \frac{n-1}{2} \rfloor$-quadratization ($y$-linear).

Theorem: General symmetric pseudo-Boolean functions

Every symmetric pseudo-Boolean function has a y-linear quadratization involving at most $(n-2)$ auxiliary variables.

Polyhedral Cone of Quadratizations

For

- $f : \{0,1\}^n \to \mathbb{R}$ a PBF in n variables,
- m : an upper bound on the number of auxiliary variables of (some of) the quadratizations $g : \{0,1\}^{n+m} \to \mathbb{R}$ of f,
- $\delta(n,m) = 1 + n + m + \binom{n+m}{2}$,

$$\mathscr{P}_f(n,m) := \Big\{ g(\mathbf{x},\mathbf{y}) \in \mathbb{R}^{\delta(n,m)} \ : \ g(\mathbf{x},\mathbf{y}) \text{ is quadratic and } f(\mathbf{x}) \le g(\mathbf{x},\mathbf{y}) \text{ for all } \mathbf{x} \in \{0,1\}^n, \mathbf{y} \in \{0,1\}^m \Big\},$$

is the *polyhedral cone of quadratizations of* f.

Polyhedral Cone of Quadratizations

of quadratizations for pos e neg terms:

n	m	# vert (+)	# rays (+)	# vert (-)	# rays (-)
2	1	1	16	1	16
2	2	1	56	1	56
3	1	12	109	12	148
3	2	292	2109	292	3808
4	1	328	4807	353	6024
4	2	100824	1677491	74107	1484561
5	1	95954	1448766	52902	1187510

Polyhedral Cone of Quadratizations

$f(x)$	# quadr	$(2,1)$	$(2,2)$	$(3,1)$	$(3,2)$	$(4,1)$	$(4,2)$	$(5,1)$
pos	total:	1	1	8	288	2	5048	0
mono	up to $\simeq$:	1	1	2	28	1	138	0
neg	total:	1	1	8	288	10	3504	12
mono	up to $\simeq$:	1	1	2	28	2	91	2

Theorem

Every m-quadratization of a PBF $f : \{0,1\}^n \to \mathbb{R}$ is on the boundary of $\mathscr{P}_f(n,m)$. Moreover, every prime m-quadratization of $f : \{0,1\}^n \to \mathbb{R}$ is a vertex of $\mathscr{P}_f(n,m)$.

A Change of Perspective

Freedman and Drineas '05 + Ishikawa '09:

Every PBF in n variables can be quadratized with at most $O(n\,2^n)$ auxiliary variables.

Is it possible to do better? How?

Functions, not variables:

All new variables $y_1, \ldots, y_m$, when minimized, are Boolean functions of the original variables $x_1, \ldots, x_n$:

$$y^*(\mathbf{x}) \;=\; \operatorname{argmin}_{\mathbf{y}}\, g(\mathbf{x}, \mathbf{y}).$$

Then x_i, $x_i y_j^*(\mathbf{x})$ and $y_i^*(\mathbf{x}) y_j^*(\mathbf{x})$ are also Boolean functions, all viewed as vectors in $\{0,1\}^{2^n}$.

A Change of Perspective

Freedman and Drineas '05 + Ishikawa '09:

Every PBF in n variables can be quadratized with at most $O(n\,2^n)$ auxiliary variables.

Is it possible to do better? How?

Functions, not variables:

All new variables $y_1, \ldots, y_m$, when minimized, are Boolean functions of the original variables $x_1, \ldots, x_n$:

$$y^*(\mathbf{x}) \;=\; \operatorname{argmin}_{\mathbf{y}} g(\mathbf{x}, \mathbf{y}).$$

Then x_i, $x_i y_j^*(\mathbf{x})$ and $y_i^*(\mathbf{x}) y_j^*(\mathbf{x})$ are also Boolean functions, all viewed as vectors in $\{0,1\}^{2^n}$.

A Change of Perspective

Freedman and Drineas '05 + Ishikawa '09:

Every PBF in n variables can be quadratized with at most $O(n\,2^n)$ auxiliary variables.

Is it possible to do better? How?

Functions, not variables:

All new variables $y_1, \ldots, y_m$, when minimized, are Boolean functions of the original variables $x_1, \ldots, x_n$:

$$y^*(\mathbf{x}) \;=\; \mathrm{argmin}_{\mathbf{y}}\, g(\mathbf{x}, \mathbf{y}).$$

Then x_i, $x_i y_j^*(\mathbf{x})$ and $y_i^*(\mathbf{x}) y_j^*(\mathbf{x})$ are also Boolean functions, all viewed as vectors in $\{0,1\}^{2^n}$.

Non-Termwise Quadratizations

Universal Sets

A set of BFs $\mathscr{U}$ is *universal* for a set of PBFs $\mathscr{F}$, both in n variables, if

- for every $f \in \mathcal{F}$, there is a m-quadratization $g(\mathbf{x}, \mathbf{y})$ with $m \leq |\mathscr{U}|$,
- there is $\{y_1^*, y_2^*, \ldots, y_m^*\} \subseteq \mathcal{U}$ such that

$$y^*(\mathbf{x}) = \left(y_1^*(\mathbf{x}), \ldots, y_m^*(\mathbf{x})\right)$$

is a minimizer of $g(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x} \in \{0, 1\}^n$.

Non-Termwise Quadratizations

Universal Sets

- When $\mathscr{U}$ is a universal set, then all minimizers $(y_1^*(\mathbf{x}), \ldots, y_m^*(\mathbf{x}))$ can be chosen in $\mathscr{U}$, for $f \in \mathscr{F}$.
- Clearly, we can take $\mathscr{U}$ as $\mathscr{B}^n$, the set of all 2^{2^n} BFs in n-variables.
- For $\mathscr{F}$ as the set of submodular PBFs, there is no $\mathscr{U}$ comprized solely of submodular QPBFs.
- Surprisingly, rather small universal sets can be exhibited through combinatorial constructions.

Non-Termwise Quadratizations

Universal Sets

- When $\mathscr{U}$ is a universal set, then all minimizers $(y_1^*(\mathbf{x}), \ldots, y_m^*(\mathbf{x}))$ can be chosen in $\mathscr{U}$, for $f \in \mathscr{F}$.
- Clearly, we can take $\mathscr{U}$ as $\mathscr{B}^n$, the set of all 2^{2^n} BFs in n-variables.
- For $\mathscr{F}$ as the set of submodular PBFs, there is no $\mathscr{U}$ comprized solely of submodular QPBFs.
- Surprisingly, rather small universal sets can be exhibited through combinatorial constructions.

Non-Termwise Quadratizations

Universal Sets

- When $\mathscr{U}$ is a universal set, then all minimizers $(y_1^*(\mathbf{x}), \ldots, y_m^*(\mathbf{x}))$ can be chosen in $\mathscr{U}$, for $f \in \mathscr{F}$.
- Clearly, we can take $\mathscr{U}$ as $\mathscr{B}^n$, the set of all 2^{2^n} BFs in n-variables.
- For $\mathscr{F}$ as the set of submodular PBFs, there is no $\mathscr{U}$ comprized solely of submodular QPBFs.
- Surprisingly, rather small universal sets can be exhibited through combinatorial constructions.

Non-Termwise Quadratizations

Universal Sets

- When $\mathscr{U}$ is a universal set, then all minimizers $(y_1^*(\mathbf{x}), \ldots, y_m^*(\mathbf{x}))$ can be chosen in $\mathscr{U}$, for $f \in \mathscr{F}$.
- Clearly, we can take $\mathscr{U}$ as $\mathscr{B}^n$, the set of all 2^{2^n} BFs in n-variables.
- For $\mathscr{F}$ as the set of submodular PBFs, there is no $\mathscr{U}$ comprized solely of submodular QPBFs.
- Surprisingly, rather small universal sets can be exhibited through combinatorial constructions.

Universal Sets for PBFs

- Given a hypergraph $\mathscr{H}$ we call another hypergraph $\mathscr{C}$ a pairwise cover of $\mathscr{H}$, if for all $H \in \mathscr{H} \cup \mathscr{C}$, $\exists A, B \in \mathscr{C}$:
  - $A, B \subsetneq H$
  - $H = A \cup B$.
- If $\mathscr{H} = 2^V$, then assuming $V = E \cup O$, we have that

$$\mathscr{C} = 2^E \cup 2^O$$

  is a pairwise cover.
- For $\mathscr{H} = \{S \subseteq V : |S| \leq d\}$, we have

$$\mathscr{C} = \bigcup_{\substack{j \geq 1, \\ k = \left\lceil \frac{d}{2^j} \right\rceil \geq 2}}^{d} \binom{V}{k}$$

  as a pairwise cover of $\mathscr{H}$.
- Pairwise covers are related to *2-covers* (Erdős '62).

Universal Sets for PBFs

- Given a hypergraph $\mathscr{H}$ we call another hypergraph $\mathscr{C}$ a pairwise cover of $\mathscr{H}$, if for all $H \in \mathscr{H} \cup \mathscr{C}$, $\exists A, B \in \mathscr{C}$:
  - $A, B \subsetneq H$
  - $H = A \cup B$.
- If $\mathscr{H} = 2^V$, then assuming $V = E \cup O$, we have that

$$\mathscr{C} = 2^E \cup 2^O$$

  is a pairwise cover.
- For $\mathscr{H} = \{S \subseteq V : |S| \leq d\}$, we have

$$\mathscr{C} = \bigcup_{\substack{j \geq 1, \\ k = \left\lceil \frac{d}{2j} \right\rceil \geq 2}}^{d} \binom{V}{k}$$

  as a pairwise cover of $\mathscr{H}$.
- Pairwise covers are related to *2-covers* (Erdős '62).

Universal Sets for PBFs

- Given a hypergraph $\mathscr{H}$ we call another hypergraph $\mathscr{C}$ a pairwise cover of $\mathscr{H}$, if for all $H \in \mathscr{H} \cup \mathscr{C}$, $\exists A, B \in \mathscr{C}$:
  - $A, B \subsetneq H$
  - $H = A \cup B$.
- If $\mathscr{H} = 2^V$, then assuming $V = E \cup O$, we have that

$$\mathscr{C} = 2^E \cup 2^O$$

  is a pairwise cover.
- For $\mathscr{H} = \{S \subseteq V : |S| \leq d\}$, we have

$$\mathscr{C} = \bigcup_{\substack{j \geq 1, \\ k = \left\lceil \frac{d}{2j} \right\rceil \geq 2}}^{d} \binom{V}{k}$$

  as a pairwise cover of $\mathscr{H}$.
- Pairwise covers are related to *2-covers* (Erdős '62).

Universal Sets for PBFs

- Given a hypergraph $\mathscr{H}$ we call another hypergraph $\mathscr{C}$ a pairwise cover of $\mathscr{H}$, if for all $H \in \mathscr{H} \cup \mathscr{C}$, $\exists A, B \in \mathscr{C}$:
  - $A, B \subsetneq H$
  - $H = A \cup B$.
- If $\mathscr{H} = 2^V$, then assuming $V = E \cup O$, we have that

$$\mathscr{C} = 2^E \cup 2^O$$

  is a pairwise cover.
- For $\mathscr{H} = \{S \subseteq V : |S| \leq d\}$, we have

$$\mathscr{C} = \bigcup_{\substack{j \geq 1, \\ k = \left\lceil \frac{d}{2^j} \right\rceil \geq 2}}^{d} \binom{V}{k}$$

  as a pairwise cover of $\mathscr{H}$.
- Pairwise covers are related to *2-covers* (Erdős '62).

Universal Sets for PBFs

- Given a hypergraph $\mathscr{H}$ we call another hypergraph $\mathscr{C}$ a pairwise cover of $\mathscr{H}$, if for all $H \in \mathscr{H} \cup \mathscr{C}$, $\exists A, B \in \mathscr{C}$:
  - $A, B \subsetneq H$
  - $H = A \cup B$.
- If $\mathscr{H} = 2^V$, then assuming $V = E \cup O$, we have that

$$\mathscr{C} = 2^E \cup 2^O$$

  is a pairwise cover.
- For $\mathscr{H} = \{S \subseteq V : |S| \leq d\}$, we have

$$\mathscr{C} = \bigcup_{\substack{j \geq 1, \\ k = \left\lceil \frac{d}{2^j} \right\rceil \geq 2}}^{d} \binom{V}{k}$$

  as a pairwise cover of $\mathscr{H}$.
- Pairwise covers are related to *2-covers* (Erdős '62).

Universal Sets for PBFs

- Given a hypergraph $\mathscr{H}$ we call another hypergraph $\mathscr{C}$ a pairwise cover of $\mathscr{H}$, if for all $H \in \mathscr{H} \cup \mathscr{C}$, $\exists A, B \in \mathscr{C}$:
  - $A, B \subsetneq H$
  - $H = A \cup B$.
- If $\mathscr{H} = 2^V$, then assuming $V = E \cup O$, we have that

$$\mathscr{C} = 2^E \cup 2^O$$

  is a pairwise cover.
- For $\mathscr{H} = \{S \subseteq V : |S| \leq d\}$, we have

$$\mathscr{C} = \bigcup_{\substack{j \geq 1, \\ k = \left\lceil \frac{d}{2j} \right\rceil \geq 2}}^{d} \binom{V}{k}$$

  as a pairwise cover of $\mathscr{H}$.
- Pairwise covers are related to *2-covers* (Erdős '62).

Upper Bounds on the Number of Auxiliary Variables

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ has a quadratization involving at most $O(2^{n/2})$ auxiliary variables.

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ of fixed degree-d has a quadratization involving at most $O(n^{d/2})$ auxiliary variables.

For a given hypergraph $\mathscr{H}$ what is the smallest pairwise cover? How easy/hard is to find a "small" pairwise cover of $\mathscr{H}$?

Upper Bounds on the Number of Auxiliary Variables

Theorem:

Every PBF $f:\{0,1\}^n \to \mathbb{R}$ has a quadratization involving at most $O\left(2^{n/2}\right)$ auxiliary variables.

Theorem:

Every PBF $f:\{0,1\}^n \to \mathbb{R}$ of fixed degree-d has a quadratization involving at most $O\left(n^{d/2}\right)$ auxiliary variables.

For a given hypergraph $\mathscr{H}$ what is the smallest pairwise cover?
How easy/hard is to find a "small" pairwise cover of $\mathscr{H}$?

Upper Bounds on the Number of Auxiliary Variables

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ has a quadratization involving at most $O(2^{n/2})$ auxiliary variables.

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ of fixed degree-d has a quadratization involving at most $O(n^{d/2})$ auxiliary variables.

For a given hypergraph $\mathscr{H}$ what is the smallest pairwise cover?
How easy/hard is to find a ”small“ pairwise cover of $\mathscr{H}$?

Upper Bounds on the Number of Auxiliary Variables

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ has a y-linear quadratization involving at most

$$O\left(\frac{2^n}{n}\log n\right)$$

auxiliary variables.

Proof:

- Different slicing of the Boolean hypercube: attractive partitions.
- Somewhat related to Turán's set systems.
- Recursive construction to bound each level of the hypercube.
- Afar from the regular construction; highly technical!

Upper Bounds on the Number of Auxiliary Variables

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ has a y-linear quadratization involving at most

$$O\left(\frac{2^n}{n}\log n\right)$$

auxiliary variables.

Proof:

- Different slicing of the Boolean hypercube: attractive partitions.
- Somewhat related to Turán's set systems.
- Recursive construction to bound each level of the hypercube.
- Afar from the regular construction; highly technical!

Upper Bounds on the Number of Auxiliary Variables

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ has a y-linear quadratization involving at most

$$O\left(\frac{2^n}{n}\log n\right)$$

auxiliary variables.

Proof:

- Different slicing of the Boolean hypercube: attractive partitions.
- Somewhat related to Turán's set systems.
- Recursive construction to bound each level of the hypercube.
- Afar from the regular construction; highly technical!

Upper Bounds on the Number of Auxiliary Variables

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ has a y-linear quadratization involving at most

$$O\left(\frac{2^n}{n}\log n\right)$$

auxiliary variables.

Proof:

- Different slicing of the Boolean hypercube: attractive partitions.
- Somewhat related to Turán's set systems.
- Recursive construction to bound each level of the hypercube.
- Afar from the regular construction; highly technical!

Upper Bounds on the Number of Auxiliary Variables

Theorem:

Every PBF $f : \{0,1\}^n \to \mathbb{R}$ has a y-linear quadratization involving at most

$$O\left(\frac{2^n}{n}\log n\right)$$

auxiliary variables.

Proof:

- Different slicing of the Boolean hypercube: attractive partitions.
- Somewhat related to Turán's set systems.
- Recursive construction to bound each level of the hypercube.
- Afar from the regular construction; highly technical!

Lower Bounds on the Number of Auxiliary Variables

Recall that:

All new variables, when minimized, are Boolean functions of the original variables $x_1, \ldots, x_n$

$$\mathbf{h}(\mathbf{x}) \;=\; \mathrm{argmin}_{\mathbf{y}}\, g(\mathbf{x}, \mathbf{y}).$$

Then x_i, $x_i h_j(\mathbf{x})$ and $h_j(\mathbf{x}) h_i(\mathbf{x})$ are also Boolean functions, all viewed as vectors in $\{0, 1\}^{2^n}$.

Lower Bounds on the Number of Auxiliary Variables

- A quadratization $g \in \mathbb{R}^{2^n}$ (viewed as a vector of values) in $n + m$ variables is a linear combination of

$$N = 1 + n + m + \binom{n+m}{2}$$

vectors, out of 2^{2^n} vectors in $\mathbb{R}^{2^n}$.

- Finitely many subspaces of dimension $< 2^n$ cannot cover $\mathbb{R}^{2^n}$. Thus, we must have $N \geq 2^n$.

Theorem:

Almost all PBFs in n variables require at least $\Omega(2^{n/2})$ auxiliary variables in any quadratization.

Lower Bounds on the Number of Auxiliary Variables

- A quadratization $g \in \mathbb{R}^{2^n}$ (viewed as a vector of values) in $n + m$ variables is a linear combination of

$$N = 1 + n + m + \binom{n+m}{2}$$

vectors, out of 2^{2^n} vectors in $\mathbb{R}^{2^n}$.

- Finitely many subspaces of dimension $< 2^n$ cannot cover $\mathbb{R}^{2^n}$. Thus, we must have $N \geq 2^n$.

Theorem:

Almost all PBFs in n variables require at least $\Omega(2^{n/2})$ auxiliary variables in any quadratization.

Lower Bounds on the Number of Auxiliary Variables

- A quadratization $g \in \mathbb{R}^{2^n}$ (viewed as a vector of values) in $n+m$ variables is a linear combination of

$$N = 1 + n + m + \binom{n+m}{2}$$

vectors, out of 2^{2^n} vectors in $\mathbb{R}^{2^n}$.

- Finitely many subspaces of dimension $< 2^n$ cannot cover $\mathbb{R}^{2^n}$. Thus, we must have $N \geq 2^n$.

Theorem:

Almost all PBFs in n variables require at least $\Omega(2^{n/2})$ auxiliary variables in any quadratization.

Lower Bounds on the Number of Auxiliary Variables

Theorem:

Almost all PBFs of fixed degree-d, in n variables require at least $\Omega(n^{d/2})$ auxiliary variables in any quadratization.

Theorem:

Almost all PBFs in n variables require at least

$$\Omega\left(\frac{2^n}{n}\right)$$

auxiliary variables in any y-linear quadratization.

Lower Bounds on the Number of Auxiliary Variables

Theorem:

Almost all PBFs of fixed degree-d, in n variables require at least $\Omega(n^{d/2})$ auxiliary variables in any quadratization.

Theorem:

Almost all PBFs in n variables require at least

$$\Omega\left(\frac{2^n}{n}\right)$$

auxiliary variables in any y-linear quadratization.

Lower and Upper Bounds in a Nutshell

For $f : \{0,1\}^n \to \mathbb{R}$ a PBF in n variables:

$f(x)$	$g(x,y)$	lower bound	upper bound
any	regular	$\Omega(2^{n/2})$	$O(2^{n/2})$
degree-d	regular	$\Omega(n^{d/2})$	$O(n^{d/2})$
any	y-linear	$\Omega\left(\frac{2^n}{n}\right)$	$O\left(\frac{2^n}{n}\log n\right)$

Lower Bounds for Symmetric PBFs

Theorem

There exist SPBFs on n variables for which any quadratization must involve at least $\Omega(\sqrt{n})$ auxiliary variables.

Theorem

There exist SPBFs on n variables for which any y-linear quadratization must involve at least $\Omega(n/\log n)$ auxiliary variables.

Lower Bounds for Symmetric PBFs

Theorem

There exist SPBFs on n variables for which any quadratization must involve at least $\Omega(\sqrt{n})$ auxiliary variables.

Theorem

There exist SPBFs on n variables for which any y-linear quadratization must involve at least $\Omega(n/\log n)$ auxiliary variables.

Lower Bounds for Symmetric PBFs

Proof idea:

- inspired by Siu, Roychowdhury, Kailath '95 work on representation of Boolean functions by *threshold circuits*;
- relates quadratizations of arbitrary functions to the quadratization of symmetric functions on a larger number of variables;
- given $f(\mathbf{x})$ on n variables, let $F(\mathbf{z})$ on $2^n - 1$ variables be such that $F(\mathbf{z}) = f(\mathbf{x})$ if $\mathbf{x}$ is the binary representation of the Hamming weight of $\mathbf{z}$;
- if F has an m-quadratization, then so does f;
- use lower bounds $\Omega(2^{n/2})$ and $\Omega(2^n/n)$ on the number of auxiliary variables for f. □

Lower Bounds for Symmetric PBFs

Proof idea:

- inspired by Siu, Roychowdhury, Kailath '95 work on representation of Boolean functions by *threshold circuits*;
- relates quadratizations of arbitrary functions to the quadratization of symmetric functions on a larger number of variables;
- given $f(\mathbf{x})$ on n variables, let $F(\mathbf{z})$ on $2^n - 1$ variables be such that $F(\mathbf{z}) = f(\mathbf{x})$ if $\mathbf{x}$ is the binary representation of the Hamming weight of $\mathbf{z}$;
- if F has an m-quadratization, then so does f;
- use lower bounds $\Omega(2^{n/2})$ and $\Omega(2^n/n)$ on the number of auxiliary variables for f. □

Lower Bounds for Symmetric PBFs

Proof idea:

- inspired by Siu, Roychowdhury, Kailath '95 work on representation of Boolean functions by *threshold circuits*;
- relates quadratizations of arbitrary functions to the quadratization of symmetric functions on a larger number of variables;
- given $f(\mathbf{x})$ on n variables, let $F(\mathbf{z})$ on $2^n - 1$ variables be such that $F(\mathbf{z}) = f(\mathbf{x})$ if $\mathbf{x}$ is the binary representation of the Hamming weight of $\mathbf{z}$;
- if F has an m-quadratization, then so does f;
- use lower bounds $\Omega(2^{n/2})$ and $\Omega(2^n/n)$ on the number of auxiliary variables for f. □

Lower Bounds for Symmetric PBFs

Proof idea:

- inspired by Siu, Roychowdhury, Kailath '95 work on representation of Boolean functions by *threshold circuits*;
- relates quadratizations of arbitrary functions to the quadratization of symmetric functions on a larger number of variables;
- given $f(\mathbf{x})$ on n variables, let $F(\mathbf{z})$ on $2^n - 1$ variables be such that $F(\mathbf{z}) = f(\mathbf{x})$ if $\mathbf{x}$ is the binary representation of the Hamming weight of $\mathbf{z}$;
- if F has an m-quadratization, then so does f;
- use lower bounds $\Omega(2^{n/2})$ and $\Omega(2^n/n)$ on the number of auxiliary variables for f. □

Lower Bounds for Symmetric PBFs

Proof idea:

- inspired by Siu, Roychowdhury, Kailath '95 work on representation of Boolean functions by *threshold circuits*;
- relates quadratizations of arbitrary functions to the quadratization of symmetric functions on a larger number of variables;
- given $f(\mathbf{x})$ on n variables, let $F(\mathbf{z})$ on $2^n - 1$ variables be such that $F(\mathbf{z}) = f(\mathbf{x})$ if $\mathbf{x}$ is the binary representation of the Hamming weight of $\mathbf{z}$;
- if F has an m-quadratization, then so does f;
- use lower bounds $\Omega(2^{n/2})$ and $\Omega(2^n/n)$ on the number of auxiliary variables for f. □

Lower Bounds for Symmetric PBFs

Theorem (parity functions)

Every y-linear quadratization of the parity function on n variables must involve at least $\Omega(\sqrt{n})$ auxiliary variables.

Proof idea: y-linear quadratization implies no quadratic term $y_i y_j$, i.e.,

$$g(\mathbf{x}, \mathbf{y}) = q(\mathbf{x}) + \sum_{i=1}^{m} y_i(\ell_i(x) - b_i),$$

where $q(\mathbf{x})$ is quadratic, and $\ell_1(\mathbf{x}), \ldots, \ell_m(\mathbf{x})$ are linear functions of $\mathbf{x}$.

- link with threshold circuit representations of parity and slicing of hypercubes by hyperplanes (Impagliazzo, Paturi, Saks '97);
- regions obtained by fixing y to all possible 2^m values;
- if m is too small, some region contains the parity function on 3 variables.

Lower Bounds for Symmetric PBFs

Theorem (parity functions)

Every y-linear quadratization of the parity function on n variables must involve at least $\Omega(\sqrt{n})$ auxiliary variables.

Proof idea: y-linear quadratization implies no quadratic term $y_i y_j$, i.e.,

$$g(\mathbf{x}, \mathbf{y}) = q(\mathbf{x}) + \sum_{i=1}^{m} y_i(\ell_i(x) - b_i),$$

where $q(\mathbf{x})$ is quadratic, and $\ell_1(\mathbf{x}), \ldots, \ell_m(\mathbf{x})$ are linear functions of $\mathbf{x}$.

- link with threshold circuit representations of parity and slicing of hypercubes by hyperplanes (Impagliazzo, Paturi, Saks '97);
- regions obtained by fixing y to all possible 2^m values;
- if m is too small, some region contains the parity function on 3 variables.

Lower Bounds for Symmetric PBFs

Theorem (parity functions)

Every y-linear quadratization of the parity function on n variables must involve at least $\Omega(\sqrt{n})$ auxiliary variables.

Proof idea: y-linear quadratization implies no quadratic term $y_i y_j$, i.e.,

$$g(\mathbf{x}, \mathbf{y}) = q(\mathbf{x}) + \sum_{i=1}^{m} y_i(\ell_i(x) - b_i),$$

where $q(\mathbf{x})$ is quadratic, and $\ell_1(\mathbf{x}), \ldots, \ell_m(\mathbf{x})$ are linear functions of $\mathbf{x}$.

- link with threshold circuit representations of parity and slicing of hypercubes by hyperplanes (Impagliazzo, Paturi, Saks '97);
- regions obtained by fixing y to all possible 2^m values;
- if m is too small, some region contains the parity function on 3 variables.

Lower Bounds for Symmetric PBFs

Theorem (parity functions)

Every y-linear quadratization of the parity function on n variables must involve at least $\Omega(\sqrt{n})$ auxiliary variables.

Proof idea: y-linear quadratization implies no quadratic term $y_i y_j$, i.e.,

$$g(\mathbf{x}, \mathbf{y}) = q(\mathbf{x}) + \sum_{i=1}^{m} y_i(\ell_i(x) - b_i),$$

where $q(\mathbf{x})$ is quadratic, and $\ell_1(\mathbf{x}), \ldots, \ell_m(\mathbf{x})$ are linear functions of $\mathbf{x}$.

- link with threshold circuit representations of parity and slicing of hypercubes by hyperplanes (Impagliazzo, Paturi, Saks '97);
- regions obtained by fixing y to all possible 2^m values;
- if m is too small, some region contains the parity function on 3 variables.

Conclusions

- Tight lower and upper bounds on the number of auxiliary variables required for arbitrary and for fixed-degree functions.
- Generic constructions for quadratization of symmetric functions.
- Structure and properties of quadratizations are still poorly understood.
- Many intriguing questions and conjectures, much computational and theoretical work to be done.

Some references

M. Anthony, E. Boros, Y. Crama, and **A. Gruber**, Quadratization of Symmetric Pseudo-Boolean Functions, *Discrete Applied Mathematics*, vol. 203, 2016, p. 1–12.

M. Anthony, E. Boros, Y. Crama, and **A. Gruber**, Quadratic Reformulations of Nonlinear Binary Optimization Problems, to appear in *Mathematical Programming*, Series A.

E. Boros and **A. Gruber**, On Quadratization of Pseudo-Boolean Functions, *Int. Symposium of Artificial Intelligence and Mathematics*, 2012.

C. Buchheim and G. Rinaldi, Efficient Reduction of Polynomial Zero-one Optimization to the Quadratic Case, *SIAM Journal on Optimization* 18 (2007) 1398-1413.

A. Fix, **A. Gruber**, E. Boros and R. Zabih, A Graph Cut Algorithm for Higher-order Markov Random Fields, Proceedings of the 2011 *IEEE International Conference on Computer Vision (ICCV)*, pages 1020-1027.

A. Fix, **A. Gruber**, E. Boros and R. Zabih, A Hypergraph-based Reduction for Higher-order Binary Markov Random Fields, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.

Some references

D. Freedman and P. Drineas, Energy Minimization Via Graph Cuts: Settling What is Possible, in: *IEEE Conference on Computer Vision and Pattern Recognition* (2) (2005) pp. 939-946.

A. Gruber, Algorithmic and Complexity Results for Boolean and Pseudo-Boolean Functions, *Ph.D. Thesis*, Rutgers University, New Brunswick, NJ, USA, 2014.

H. Ishikawa, Transformation of General Binary MRF Minimization to the First-order Case, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(6) (2011) 1234-1249.

V. Kolmogorov and C. Rother, Minimizing Non-submodular Functions with Graph Cuts - A Review, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (2007) 1274-1279.

V. Kolmogorov and R. Zabih, What Energy Functions can be Minimized via Graph Cuts?, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(2) (2004) 147-159.

Some references

S. Ramalingam, Ch. Russell, L. Ladický and Ph.H.S. Torr, Efficient Minimization of Higher Order Submodular Functions using Monotonic Boolean Functions, arXiv:1109.2304v1, 2011.

I.G. Rosenberg, Reduction of Bivalent Maximization to the Quadratic Case, *Cahiers du Centre d'Etudes de Recherche Opérationnelle* 17 (1975), 71–74.

C. Rother, P. Kohli, W. Feng and J. Jia, Minimizing Sparse Higher-order Energy Functions of Discrete Variables, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1382-1389.

C. Rother, V. Kolmogorov, V. Lempitsky and M. Szummer, Optimizing Binary MRFs via Extended Roof Duality, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

S. Živný, D.A. Cohen and P.G. Jeavons, The Expressive Power of Binary Submodular Functions, *Discrete Applied Mathematics* 157 (2009) 3347-3358.