Relaxations for Binary Polynomial Optimization via Signed Certificates

Liding Xu, Leo Liberti

Zuse Institut Berlin, Ecole Polytechnique

ICCOPT 2025

July 2025

SPONSORED BY THE

Federal Ministry of Education and Research

Introduction

Introduction: Problem Setting

- We focus on the **Binary Polynomial Optimization (BPO)** problem:

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

  where f is a polynomial in n variables.
- Closely related is the **Binary Non-negativity Problem (BNP)**:

$$\text{Is } f(x) \geq 0 \text{ for all } x \in \{0,1\}^n?$$

- The BPO is equivalent to finding the maximum λ such that $f(x) - \lambda$ is binary non-negative. This is a conic optimization problem:

$$\lambda^* = \max_{\lambda \in \mathbb{R}} \{\lambda : f - \lambda \text{ is binary non-negative}\}$$

- Conic inner approximation leads to lower bounds on λ^*. Previous constructions include Lift-Project (LP), Sherali-Adams (LP), and Sum-of-Squares (SDP), ~~SONC/SAGE (geometric programs)~~ hierarchies. [Lasserre, 2015, Parrilo and Thomas, 2020, Sherali and Tuncbilek, 1992, Sherali and Tuncbilek, 1997]

Our Contribution

- We propose a new class of sparse binary non-negativity certificates based on the polynomial's **signed support pattern**.
- We develop new LP relaxations for BPO that are sparsity-preserving.
- **Key Idea:** Decompose any polynomial f and leverage the fact that the non-negativity of certain polynomial classes can be checked efficiently.

Preliminaries

Polynomial Classification

We classify binary polynomials based on the signs of their coefficients:

- **PS (Positively Signed):** All coefficients are non-negative.
- **NS (Negatively Signed):** All coefficients are non-positive.

Key Property: NNS polynomials are submodular, and NPS polynomials are supermodular.

Polynomial Classification

We classify binary polynomials based on the signs of their coefficients:

- **PS (Positively Signed):** All coefficients are non-negative.
- **NS (Negatively Signed):** All coefficients are non-positive.
- **NNS (Nonlinearly Negatively Signed):** Coefficients of all nonlinear monomials (degree ≥ 2) are non-positive.
- **NPS (Nonlinearly Positively Signed):** Coefficients of all nonlinear monomials are non-negative.
- **NDS (Nonlinearly Differently Signed):** Neither NNS nor NPS.

Key Property: NNS polynomials are submodular, and NPS polynomials are supermodular.

Signed Support Decomposition

Any binary polynomial f can be uniquely decomposed as:

$$f(x) = \mathrm{NNS}(f)(x) + \mathrm{PS}(f)(x)$$

- NNS(f) is the **NNS component** of f:

$$\mathrm{NNS}(f)(x) := f_0 + \sum_{\alpha \in A:\ \text{degree-1 exponent vectors}} f_\alpha x^\alpha + \sum_{\alpha \in A:\ \text{high-degree exponent vectors}} \min(f_\alpha, 0) x^\alpha$$

- PS(f) is the **PS component** of f:

$$\mathrm{PS}(f)(x) := \sum_{\alpha \in A:\ \text{high-degree exponent vectors}} \max(f_\alpha, 0) x^\alpha$$

Binary Non-negativity of NNS Polynomials

Minimizing NNS Polynomials

- The problem $\min_{x\in\{0,1\}^n} f(x)$ for an NNS polynomial f can be solved efficiently, due to submodularity.
- A more efficient approach reduces the problem to a **minimum cut** problem in a specially constructed graph. [Billionnet and Minoux, 1985, Hansen, 1974, Picard and Queyranne, 1982].

Reduction to Min-Cut

For an NNS polynomial $f(x) = f_0 + \sum_{\alpha \in A} f_\alpha x^\alpha + \sum_{j \in \mathcal{N}} f_j x_j$ (with $f_\alpha \leq 0$ for $\alpha \in A$), we have:

$$\min_{x \in \{0,1\}^n} f(x) = f^a + \min_{x \in \{0,1\}^n} f^b(x)$$

where $f^a = f_0 + \sum_{\alpha \in A} f_\alpha$ is a constant and

$$f^b(x) = \sum_{\alpha \in A} -f_\alpha (1 - x^\alpha) + \sum_{j \in \mathcal{N}} f_j x_j$$

$$f^c(\mathbf{x}) := \sum_{\alpha \in A:\text{high-degree}} -f_\alpha (1 - \mathbf{x}^\alpha) + \sum_{j \in \mathcal{N}} \max(f_j, 0) x_j. \tag{1}$$

The graph will be a bipartite graph of nodes for nonlinear monomials A and linear monomials $\mathcal{N}$.

Min cut and max flow

Example: $f^c(\mathbf{x}) = (1 - x_2x_3) + 2(1 - x_1x_3x_4) + 5(1 - x_3x_5) + x_2 + x_3$.

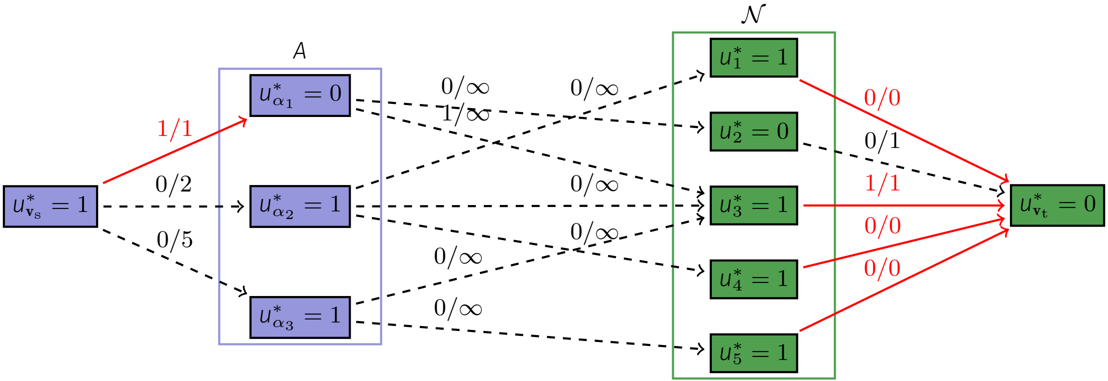

Figure: The network G^c with source $\mathbf{s}$ and terminate $\mathbf{v_t}$ is derived from f^c. Nodes are labeled by their values in the MIN CUT, the MAX FLOWS and capacities of edges are labeled above, and the cut crosses solid red edges.

LP Formulation for NNS Non-negativity

- The min-cut problem has a dual max-flow problem, which can be formulated as a linear program (LP).
- We show that this duality yields an extended LP formulation for the cone $\mathrm{NNS}^+(\mathbf{s})$.
- Technical reductions: The condition $\min_{x \in \{0,1\}^n} f(x) \geq 0$ is equivalent to:

$$
\begin{aligned}
f_{\mathbf{0}} + \sum_{\alpha \in A} f_\alpha + \sum_{j \in \mathcal{N}} \rho_{j\,\mathbf{v}_\mathrm{t}} &\geq 0 & \\
\rho_{\mathbf{v}_\mathrm{s}\,\alpha} &\leq -f_\alpha & \forall \alpha \in A \\
\rho_{\mathbf{v}_\mathrm{s}\,\alpha} &= \sum_{j \in \mathrm{supp}(\alpha)} \rho_{\alpha j} & \forall \alpha \in A \\
\rho_{\mathbf{v}_\mathrm{s}\,j} + \sum_{\alpha \in A_j} \rho_{\alpha j} &= \rho_{j\,\mathbf{v}_\mathrm{t}} & \forall j \in \mathcal{N} \\
\rho_j &\leq f_j & \forall j \in \mathcal{N}
\end{aligned}
$$

Here, ρ are flow variables.

Concave Extensions of PS Polynomials

Handling the PS Component

- The PS component $\mathrm{PS}(f)$ is a supermodular function.
- We use **piecewise linear concave extensions** to find a set of linear functions that overestimate $\mathrm{PS}(f)$.
- Let $\mathcal{M}(\mathbf{s}^{\mathrm{p}})$ be a set of "overestimation matrices". For each $M \in \mathcal{M}(\mathbf{s}^{\mathrm{p}})$, $M\mathrm{PS}(f)$ is a linear polynomial and $(M\mathrm{PS}(f))(x) \geq \mathrm{PS}(f)(x)$ for $x \in \{0,1\}^n$.
- The extension is **exact** if $\min_{M \in \mathcal{M}(\mathbf{s}^{\mathrm{p}})}(M\mathrm{PS}(f))(x) = \mathrm{PS}(f)(x)$.

Types of Concave Extensions

- **Standard Extension:** Based on monomial linearization. For each monomial $\mathbf{x}^\alpha$, pick one variable x_j where $j \in \mathrm{supp}(\alpha)$.

$$(M_\sigma f)(x) = \sum_{\alpha \in \mathrm{supp}(\mathbf{s})} f_\alpha x_{\sigma(\alpha)}$$

  Number of matrices: $\prod_{\alpha \in \mathrm{supp}(\mathbf{s})} |\alpha| \leq d^m$.

- **Lovász Extension:** Based on permutations of variables. For each permutation π of $\{1, ..., n\}$:

$$(M_\pi f)(x) = \sum_{j=1}^{n} \left(f\left(\sum_{i=1}^{j} \mathbf{e}_{\pi(i)} \right) - f\left(\sum_{i=1}^{j-1} \mathbf{e}_{\pi(i)} \right) \right) x_{\pi(j)}$$

  Number of matrices: $n!$. Can be filtered down to 2^n.

Loose extension is better than tight extension!

BPO Reformulation

Certifying Non-negativity for NDS Polynomials

Lemma

A polynomial $f = \mathrm{NNS}(f) + \mathrm{PS}(f)$ *is binary non-negative if and only if for every overestimation matrix* M *for* $\mathrm{PS}(f)$*, the NNS polynomial*

$$\mathrm{NNS}(f) + M\mathrm{PS}(f)$$

is binary non-negative.

- This reduces the BNP for a general polynomial f to a set of BNPs for NNS polynomials.
- Each of these NNS-BNPs can be checked efficiently.
- $\mathrm{NNS}(f) + M\mathrm{PS}(f) \geq 0$ is a polyhedral cone!

Signed Support Decomposition of a Pattern s

Definition 5

Given a signed support pattern $\mathbf{s} \in \{-1, 0, 1\}^{\{0,1\}^n}$ (with linear terms in the NNS part), we decompose it into two disjoint parts:

$$\mathbf{s} = \mathbf{s}^{\mathsf{nn}} + \mathbf{s}^{\mathsf{p}}$$

Conditions on the Decomposition:

- $\mathbf{s}^{\mathsf{nn}}, \mathbf{s}^{\mathsf{p}} \in \{-1, 0, 1\}^{\{0,1\}^n}$
- $s^{\mathsf{nn}}_{\{0,1\}^n_{2:n}} \leq \mathbf{0}$
  (The non-linear part of $\mathbf{s}^{\mathsf{nn}}$ is purely negative)
- $s^{\mathsf{p}}_{\{0,1\}^n_{2:n}} \geq \mathbf{0}$
  (The non-linear part of $\mathbf{s}^{\mathsf{p}}$ is purely positive)
- $s^{\mathsf{p}}_{\{0,1\}^n_{0:1}} = \mathbf{0}$
  (The PS part has no linear or constant terms)
- $\mathrm{supp}(\mathbf{s}^{\mathsf{nn}}) \cap \mathrm{supp}(\mathbf{s}^{\mathsf{p}}) = \varnothing$
  (The supports are disjoint)

Derived Complexity Parameters:

- For each part $i \in \{\mathsf{nn}, \mathsf{p}\}$:
  - $m_i := |\,\mathrm{supp}(\mathbf{s}^i)|$
    (Number of monomials)
  - $d_i := \max_{\alpha \in \mathrm{supp}(\mathbf{s}^i)} |\alpha|$
    (Maximum degree)
  - $n_i := |\mathcal{N}(\mathbf{s}^i)|$
    (Number of variables)
- For the combined pattern $\mathbf{s}$:
  - $m := m_{\mathsf{nn}} + m_{\mathsf{p}}$

The Cone of Non-Negative Polynomials

We define the cone of binary non-negative NDS polynomials with a given signed support pattern $\mathbf{s}$:

$$\mathsf{NDS}^+(\mathbf{s}) := \{f \in \mathbb{R}(x) : \mathsf{NNS}(f) \in \mathsf{SSC}(\mathbf{s}^{\mathsf{nn}}), \mathsf{PS}(f) \in \mathsf{SSC}(\mathbf{s}^{\mathsf{p}}), \\ \forall M \in \mathcal{M}(\mathbf{s}^{\mathsf{p}}), \mathsf{NNS}(f) + M\mathsf{PS}(f) \in \mathsf{NNS}^+(\mathbf{s}^{\mathsf{nn}})\}$$

Theorem

$\mathsf{NDS}^+(\mathbf{s})$ *is a convex polyhedral cone with an extended LP formulation of size polynomial in* m, d *and linear in* $\Gamma(\mathbf{s}^{\mathsf{p}})$ *(the number of overestimation matrices).*

Signed Reformulation of BPO

The original BPO problem is equivalent to the following conic optimization problem:

$$\lambda^\star = \max_{\lambda \in \mathbb{R}} \{\lambda : f - \lambda \in \mathsf{NDS}^+(\mathbf{s})\}$$

- This is an LP with a potentially large number of constraints, depending on $\Gamma(\mathbf{s}^{\mathsf{p}})$.

Hierarchies of Relaxations

Refined Signed Support Decomposition

- To manage the complexity from $\Gamma(\mathbf{s}^{\mathrm{p}})$, we don't handle the whole PS part at once.
- We create a **refined signed support decomposition** of f:

$$f = g + \sum_{k=1}^{\ell} f^k$$

where g is a PS polynomial and each f^k is a signed certificate from a simpler cone $\mathrm{NDS}^+(\theta^k)$.

- This defines an inner approximation of the full cone:

$$\mathrm{SoSC}(\Theta(\mathbf{s})) \subseteq \mathrm{NDS}^+(\mathbf{s})$$

Hierarchical Partition

- We use a **hierarchical partition tree** to systematically create nested families of these inner approximations.
- At level *i* of the hierarchy, we partition the ”difficult” part of the problem (either monomials or variables of the PS part) into smaller, manageable chunks.
- This gives a sequence of cones:

$$\mathsf{SoSC}(\Theta^1(\mathbf{s})) \subseteq \mathsf{SoSC}(\Theta^2(\mathbf{s})) \subseteq \cdots \subseteq \mathsf{SoSC}(\Theta^{\bar{h}}(\mathbf{s})) \approx \mathsf{NDS}^+(\mathbf{s})$$

- This results in a hierarchy of LP relaxations with improving bounds.

Two Hierarchies of Relaxations

- **Standard Signed Relaxations:**
  - Partitions the set of monomials of the PS part.
  - Converges in at most $\bar{h} \leq \lceil \log m_{\mathrm{p}} \rceil$ steps.
  - Complexity of level i: $\mathcal{O}(m_{\mathrm{nn}} d_{\mathrm{nn}} m_{\mathrm{p}} d_{\mathrm{p}}^{2^i})$.
- **Lovász Signed Relaxations:**
  - Partitions the set of variables of the PS part.
  - Converges in at most $\bar{h} \leq \lceil \log n_{\mathrm{p}} \rceil$ steps.
  - Complexity of level i: $\mathcal{O}(m_{\mathrm{nn}} d_{\mathrm{nn}} m_{\mathrm{p}} 2^{2^i})$.

Computational Results

Experimental Setup

- We tested our relaxations on MAX CUT problem instances from the Biq Mac library.
- We compared our **Standard Signed Relaxations** (levels 1, 2, 3) with the first level of the **Sherali-Adams** and **Lasserre** hierarchies.
- Metrics: solution time and relative duality gap.

Results

2Setting	pm1s_ni		w01_100		t2gn_seed		t3gn_seed		All	
	gap	time	gap	time	gap	time	gap	time	gap	time
SheraliAdams level 1	0.509	**0.0**	0.47	**0.0**	0.173	**0.0**	0.277	**0.0**	0.373	**0.0**
Lasserre level 1	**0.127**	4.1	**0.115**	7.4	0.183	340.8	0.189	443.5	**0.144**	27.9
Standard signed 1	0.275	7.6	0.252	12.2	0.104	9.3	0.167	24.6	0.21	11.0
Standard signed 2	0.253	14.7	0.24	26.4	0.095	55.5	0.161	125.0	0.196	32.1
Standard signed 3	0.239	29.3	0.229	49.2	**0.088**	128.7	**0.156**	304.9	0.186	67.2

Table: Summary of performance metrics.

Summary of Results

- The Sherali-Adams relaxation is fast but gives weak bounds.
- The Lasserre relaxation gives strong bounds but can be slow, especially for larger instances.
- Our Standard Signed Relaxations offer a good trade-off:
    - They are competitive with the Lasserre relaxation in terms of bound quality.
    - They show better scalability on larger problem instances (but depends on signs).
    - Higher levels of our hierarchy consistently improve the bounds.

Conclusion

Conclusion

- We introduced a new method for constructing LP relaxations for BPO based on the signed support pattern of the polynomial.
- Our method leverages the efficient minimization of NNS polynomials and concave extensions of PS polynomials.
- We proposed two hierarchies of relaxations (Standard and Lovász) that are sparsity-preserving and converge to the true optimum.
- Tailored LP solvers?

Billionnet, A. and Minoux, M. (1985).
Maximizing a supermodular pseudoboolean function: A polynomial algorithm for supermodular cubic functions.
Discrete Applied Mathematics, 12(1):1–11.

Hansen, P. (1974).
Programmes mathématiques en variables 0-1.
PhD thesis, Université libre de Bruxelles, Faculté des sciences appliquées.

Lasserre, J. B. (2015).
An introduction to polynomial and semi-algebraic optimization.
Cambridge Texts in Applied Mathematics. Cambridge: Cambridge University Press.

Parrilo, P. A. and Thomas, R. R., editors (2020).
Sum of squares: theory and applications. AMS short course, Baltimore, MD, USA, January 14–15, 2019, volume 77 of *Proceedings of Symposia in Applied Mathematics.*
Providence, RI: AMS.

Picard, J.-C. and Queyranne, M. (1982).
A network flow solution to some nonlinear 0-1 programming problems, with applications to graph theory.
Networks, 12(2):141–159.

Sherali, H. D. and Tuncbilek, C. H. (1992).

A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique.
Journal of Global Optimization, 2(1):101–112.

Sherali, H. D. and Tuncbilek, C. H. (1997).
New reformulation linearization/convexification relaxations for univariate and multivariate polynomial programming problems.
Operations Research Letters, 21(1):1–9.