New Proofs of Exact LP Reformulations for Binary Polynomial Optimization with Bounded Treewidth

Sourabh K. Choudhary[a], Santanu S. Dey[a], Nikolaos V. Sahinidis[a,b]

[a]*H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, USA*
[b]*School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, USA*

Abstract

In this work, we revisit binary polynomial optimization (BPO) problems with limited treewidth of the associated graph. We provide alternate proofs of the exactness of a reformulated linear program (LP) with $O(n2^d)$ variables, n being the number of variables and d being the treewidth of the associated graph. The first proof relies on expressing any given fractional solution as the midpoint of two points in the feasible region in order to prove integrality. The second proof is a primal-dual complementary slackness argument explicitly constructing optimal solutions to primal and dual variables.

Keywords: Dynamic programming, tree decomposition, Fixed parameter tractable

1. Introduction

A Binary Polynomial Optimization (BPO) problem, also known as pseudo-boolean optimization or $0-1$ polynomial programming, is an optimization problem of the form:

$$\begin{aligned} \min \quad & \sum_{A \subseteq N} c_A \prod_{i \in A} b_i \\ \text{s.t.} \quad & f_i(b_1, b_2, \ldots, b_n) \leq 0, \quad i \in [m], \\ & b_i \in \{0,1\}, \quad i \in [n], \end{aligned}$$

where $[m] := \{1, \ldots, m\}$, $f_i, i \in [m]$ are polynomials, and $c_A \in \mathbb{R}$ for all $A \subseteq [n]$. Even in unconstrained form, these problems are NP-hard in general. They provide a modeling framework for a variety of discrete optimization problems as many combinatorial objectives and logical constraints can be written as multilinear expressions of binary variables (see Boros and Hammer (2002) and Crama and Hammer (2011)). These problems have several applications spanning finance, communication, theoretical physics, economics, and machine learning (Elloumi et al. (2021)). A notable instance of these problems is low auto-correlation sequencing problems from the MINLP library (MINLPLib (2024)). Such problems arise in physics for determining the ground state of the Bernasconi model (Liers et al. (2010)). Another important subclass of these problems is Quadratic Unconstrained Binary Programming (QUBO), which captures a wide range of network problems like max-cut, set-partitioning, facility location, resource allocation, and clustering, among others (Kochenberger et al. (2014)). A central theme in algorithms for solving BPO is the construction of convex, often linear, relaxations. Convex relaxations can be progressively refined to yield an exact convex hull formulation, giving rise to a hierarchy of relaxations (Lovász and Schrijver (1991), Sherali and Adams (1990), Laurent (2003)). In this work, we derive a linear program equivalent to BPO and show that the representation of its feasible region is Totally Dual Integral (TDI). Access to such a linear program can help in the generation of cutting planes for BPO relaxations and potentially conducting sensitivity analysis with respect to changing objective function (Martin et al. (1990)).

1.1. A first reformulation in terms of multilinear sets

For each product of binary variables $\prod_{i\in A} b_i$, we introduce a new variable w_A including the variables $w_{\{j\}}$ for b_j. Let $\mathbf{w}$ be the resulting 2^{n-1} dimensional vector of variables, each variable for one of the nonempty subsets of $[n]$. Thus, we may rewrite BPO in this lifted space as follows:

$$\min_w \quad \sum_{A\subseteq[n]} c_A w_A \tag{1}$$

$$\text{s.t.} \quad g_i(\mathbf{w}) \geq 0, \quad i \in [m] \tag{2}$$

$$w_A = \prod_{j\in A} w_{\{j\}}, \quad A \subseteq [n] \tag{3}$$

$$w_{\{j\}} \in \{0,1\}, \quad i \in [n]. \tag{4}$$

Here g_i is the affine function obtained after replacing each product of binary variables with the corresponding introduced variable in f_i. Let $\mathcal{S}_n$ be the set of binary points described by (3)–(4) and $\tilde{\mathcal{S}}_n$ be the set of binary points described by (2)–(4). Since the objective function is linear in this lifted space, we can replace the feasible set, $\tilde{\mathcal{S}}_n$, with its convex hull $\text{conv}(\tilde{\mathcal{S}}_n)$ to obtain an equivalent formulation:

$$\begin{aligned} \min_w \quad & \sum_{A \subseteq N} c_A w_A \\ \text{s.t.} \quad & \mathbf{w} \in \text{conv}(\tilde{\mathcal{S}}_n). \end{aligned}$$

In general, such an equivalent LP has $O(2^n)$ variables and $O(2^n)$ constraints as $\text{conv}(\tilde{\mathcal{S}}_n)$ involves 2^n constraints generally. However, we note that not all monomials of binary variables necessarily appear in the objective function or the constraints. This observation can lead to a more compact representation of the convex hull, reducing the number of variables and inequalities required, as discussed next.

1.2. Treewidth and the convex hull of the objective function

As was previously known in (Laurent (2009), Bienstock and Muñoz (2018)), we next present a description of the convex hull with fewer than $O(2^n)$ variables and $O(2^n)$ constraints based on the graph-theoretic notion of treewidth. Given a set V, we use 2^V to denote the power set of V, that is the set of all subsets of V.

Definition 1 (Tree decomposition of a graph)**.** For a graph $G = (V, E)$, a tree decomposition is defined as (T, ϕ) where $T = (NS, FS)$ is a tree with node set NS and edge set FS and ϕ is a labeling function $\phi : NS \to 2^V$ such that the following three properties are satisfied:

1. For all $v \in V$, there exists $p \in NS$ such that $v \in \phi(p)$.
2. For all $(u, v) \in E$, there exists $p \in NS$ such that $u, v \subseteq \phi(p)$.
3. For each vertex $v \in V$, the set of all the nodes of T given by $\{p \in NS : v \in \phi(p)\}$ forms a subtree of T.

The width of a tree decomposition is defined as $\max_{p \in NS} |\phi(p)| - 1$. The treewidth is the minimum width over all possible tree decompositions. We will use d to denote the treewidth.

For a BPO instance, one can associate a graph which we refer to as the associated graph as follows. For each variable in the problem, we create a vertex. An edge is added between two vertices if the corresponding variables appear together in a monomial in the objective function or if they appear in the same constraint. We can find a tree decomposition corresponding to the treewidth of the associated graph. We refer to the treewidth of the associated graph as the treewidth of the BPO instance.

It is well known that several NP-hard combinatorial optimization problems on graphs with limited treewidth can be solved in polynomial time. See, for example, Bodlaender and Jansen (2000) for the max-cut problem, Telle and Proskurowski (1997) for the vertex partitioning problem, and Gerber and Kobler (2003) for chromatic index and graph partitioning. Wainwright and Jordan (2004) show that obtaining the convex hulls using Sherali-Adams operations (Sherali and Adams, 1990), also known as relaxation linearization technique (RLT), simplifies when the treewidth of the objective is bounded. In particular, one needs to consider RLT up to degree equal to the treewidth, instead of n, to obtain an exact LP reformulation. A similar result presented in Section 8 of Laurent (2009) and in Bienstock and Muñoz (2018) states that a BPO instance with n variables and treedwidth d admits a linear programming reformulation with $O(n2^d)$ variables and constraints. See also Faenza et al. (2022), Del Pia and Khajavirad (2024) and Cifuentes et al. (2024) for related results. We call this result the *treewidth-convex hull theorem* and state it formally as follows.

Theorem 1 (Treewidth-convex hull theorem)**.** *Consider a binary polynomial optimization instance with treewidth d, and let k denote the number of nodes in a tree decomposition that achieves this treewidth. This binary polynomial optimization instance can be reformulated as a linear program with $O(k2^d)$ variables and $O(k2^d)$ constraints.*

This work provides two new alternate proofs of Theorem 1. Previous proofs (Laurent (2009), Bienstock and Muñoz (2018)) are based on mathematical induction on the number of nodes in the tree decomposition that yields treewidth d. The base case considers instances with a tree decomposition with a single node, where it is shown that any feasible solution to the LP can be expressed as a convex combination of integer feasible points. For the inductive step, the argument assumes that the claim holds for any tree decomposition with at most k nodes, that is, any fractional feasible solution in the LP's feasible region can be written as a convex combination of integer points. The proof then establishes the integrality for an instance with a tree decomposition with $k+1$ nodes by constructing the desired integer points from the feasible region using the integer points constructed for the monomials belonging to the two subtrees of the tree decomposition.

In this work, the first proof shows that a fractional feasible solution of the LP can be expressed as an average of two distinct feasible points, thereby proving that fractional points cannot be extreme points, and hence integrality of the LP. In the second proof, we explicitly construct an integer optimal solution for any choice of objective coefficients. Primal-dual complementary slackness is used to prove the optimality of the constructed integer feasible solution. The dual to the LP is the DP formulation for solving such problems over known tree decompositions. As a corollary, we obtain the representation of the feasible region of the LP is totally dual integral.

The central ingredient in both proofs is a new reformulation of the standard LP associated with these results. This reformulation is presented in Section 2, followed by the two proofs of Theorem 1 in Section 3.

For clarity in navigating the notation employed throughout the proof, we shall make use of the following running example:

Example 1.

$$\begin{aligned} \min \quad & b_1b_2b_3 - 5b_1b_4 + 4b_1b_5 \\ \text{s.t.} \quad & b_1 + b_2 + b_3 \geq 2 \\ & b \in \{0,1\}^5. \end{aligned}$$

Clearly, the associated graph contains a clique among the vertices representing b_1*,* b_2 *and* b_3*, whereas the pairs* $b_1 - b_4$ *and* $b_1 - b_5$ *are connected with an edge. The associated graph and the corresponding tree decomposition are shown in Figure 1. The label function of the nodes in the tree decomposition is shown in the rectangles that represent the nodes.*

2. Reformulations and the LP relaxation

In this section, we present a reformulation of BPO. Suppose T is a tree decomposition of a given BPO instance corresponding to its treewidth d, which has k nodes with indices $\{1, 2, \ldots, k\}$. For $i \in [k]$, let $\mathcal{B}_i \subseteq N$ be the set of indices in the label function of i, that is $\mathcal{B}_i = \phi(i)$. We use $w[\mathcal{B}_i] \in \mathbb{R}^{2^{|\mathcal{B}_i|}-1}$ to denote vectors of all the monomials that can be formed from the variables with indices in $\mathcal{B}_i$. For example, if $\mathcal{B}_i = \{1, 2, 3\}$ for some i, then possible values of $w[\mathcal{B}_i] = [w_{\{1\}}, w_{\{2\}}, w_{\{3\}}, w_{\{1,2\}}, w_{\{1,3\}}, w_{\{2,3\}}, w_{\{1,2,3\}}]^T$

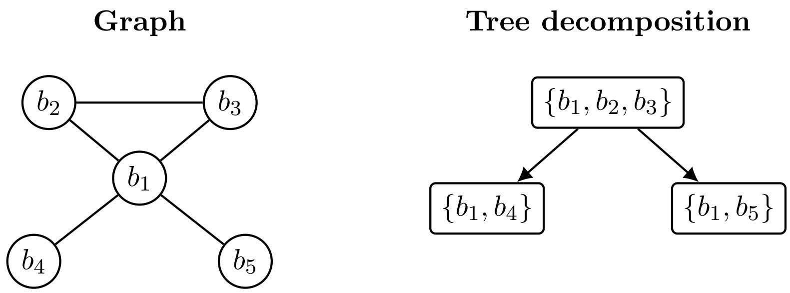

Figure 1: Associated graph and tree decomposition for the example

are: $\{0,0,0,0,0,0,0\}$, $\{1,0,0,0,0,0,0\}$, $\{0,1,0,0,0,0,0\}$, $\{0,0,1,0,0,0,0\}$, $\{1,1,0,1,0,0,0\}$, $\{1,0,1,0,1,0,0\}$, $\{0,1,1,0,0,1,0\}$, $\{1,1,1,1,1,1,1\}$.

Let $\tilde{\mathcal{S}}_{|\mathcal{B}_i|} \subseteq \mathcal{S}_{|\mathcal{B}_i|}$ be the projection of the set described by (2), (3), and (4), onto the space of the $w[\mathcal{B}_i]$ variables. Consider the following formulation:

$$\begin{aligned} \min \quad & \sum_{A \subseteq N} c_A w_A \\ \text{s.t.} \quad & w[\mathcal{B}_i] \in \tilde{\mathcal{S}}_{|\mathcal{B}_i|} \qquad i \in [k] \\ & w[\mathcal{B}_i] \in \{0,1\}^{2^{|\mathcal{B}_i|}-1} \qquad i \in [k]. \end{aligned}$$

We claim that the above is an exact integer programming reformulation of our problem. This can be seen as a consequence of using a tree decomposition to build the reformulation. By definition of the associated graph, vertices for variables present in any particular monomial in the objective function form a clique in the associated graph. The same holds true for the vertices for variables appearing in the same constraint. According to a standard result in graph theory, all the vertices in a clique of a graph appear together in the label of at least one of the nodes of its tree decomposition. Consequently, indices of any particular monomial in the objective function appear together in the label of at least one node of T and a similar observation can be made for all the variables present in a particular constraint. Therefore, we obtain that:

- For every constraint $g_i(\mathbf{w}) \geq 0$, there exists a node in the tree decomposition, say i^*, that includes all the variables in the support of this constraint. Thus, the constraint $w[\mathcal{B}_{i^*}] \in \tilde{\mathcal{S}}_{|\mathcal{B}_{i^*}|}$ above enforces the constraint $g_i(\mathbf{w}) \geq 0$.

- The support of a monomial with a non-zero coefficient in the objective is included in at least one node of the tree decomposition. Thus, the formulation guarantees that all relevant variables are present and that no variable is left missing.

Since we continue to enforce integrality of the w variables, we can replace each $\tilde{\mathcal{S}}_{|\mathcal{B}_i|}$ with its convex hull, and therefore arrive at the following equivalent formulation:

$$\begin{aligned}
\min \quad & \sum_{A \subseteq N} c_A w_A \\
\text{s.t.} \quad & w[\mathcal{B}_i] \in \operatorname{conv}(\tilde{\mathcal{S}}_{|\mathcal{B}_i|}) \quad i \in [k] \qquad (5)\\
& w[\mathcal{B}_i] \in \{0,1\}^{2^{|\mathcal{B}_i|}-1} \quad i \in [k]. \qquad (6)
\end{aligned}$$

Next, we will express the inclusion of $w[\mathcal{B}_i]$ in sets $\operatorname{conv}(\tilde{\mathcal{S}}_{|\mathcal{B}_i|})$ by equality constraints expressing $w[\mathcal{B}_i]$ as a convex combination of points in $\tilde{\mathcal{S}}_{|\mathcal{B}_i|}$. For a block $i \in [k]$, denote the binary points in the multilinear set $\mathcal{S}_{|\mathcal{B}_i|}$ as $u^{iP} \in \{0,1\}^{2^{|\mathcal{B}_i|}-1}$, indexed with $P \subseteq \mathcal{B}_i$. Each of the points u^{iP} is given by:

$$u_Q^{iP} = \begin{cases} 1 & \emptyset \subsetneq Q \subseteq P, \\ 0 & \text{otherwise.} \end{cases}$$

Denote by $\mathcal{P}_i$ the set $\{P : u^{iP} \in \tilde{S}_{|\mathcal{B}_i|}\}$. In other words, $\mathcal{P}_i$ is the index set of feasible points in the multilinear set for the bag $\mathcal{B}_i$. Our next step is to introduce the convex-combination multiplier variables $\lambda^i \in \mathbb{R}_+^{|\tilde{\mathcal{S}}_{\mathcal{B}_i}|}$, $i \in [k]$, where λ^{iP} corresponds to u^{iP}, and rewrite (5)–(6) as follows:

$$\begin{aligned}
\min \quad & \sum_{A \subseteq N} c_A w_A \\
\text{s.t.} \quad & w[\mathcal{B}_i] = \sum_{P \in \mathcal{P}_i} \lambda^{iP} u^{iP}, \quad i \in [k] \\
& \sum_{P \in \mathcal{P}_i} \lambda^{iP} = 1, \quad i \in [k] \\
& w[\mathcal{B}_i] \in \{0,1\}^{2^{|\mathcal{B}_i|}-1} \quad i \in [k]. \\
& \lambda^i \in \mathbb{R}_+^{|\tilde{\mathcal{S}}_{\mathcal{B}_i}|} \quad i \in [k].
\end{aligned}$$

This reformulation has $O(k2^d)$ variables and constraints. In the following, we will show that it is possible to drop the integrality requirement, as the polyhedral feasible region is integral.

Example 2. *For our running example, introducing lifting variables for the products as well as convex hull multiplier variables for each node of the tree decomposition yields the following equivalent formulation:*

$$
\begin{aligned}
\min \quad & w_{123} - 5w_{14} + 4w_{15} \\
\text{s.t.} \quad & \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_{12} \\ w_{23} \\ w_{13} \\ w_{123} \end{pmatrix} = \lambda^{1,\{2,3\}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda^{1,\{1,3\}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda^{1,\{1,2\}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda^{1,\{1,2,3\}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\
& \begin{pmatrix} w_1 \\ w_4 \\ w_{14} \end{pmatrix} = \lambda^{2,\emptyset} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \lambda^{2,\{4\}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda^{2,\{1\}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda^{2,\{1,4\}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
& \begin{pmatrix} w_1 \\ w_5 \\ w_{15} \end{pmatrix} = \lambda^{3,\emptyset} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \lambda^{3,\{5\}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda^{3,\{1\}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda^{3,\{1,5\}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
& \lambda^{1,\{2,3\}} + \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}} = 1 \\
& \lambda^{2,\emptyset} + \lambda^{2,\{4\}} + \lambda^{2,\{1\}} + \lambda^{2,\{1,4\}} = 1 \\
& \lambda^{3,\emptyset} + \lambda^{3,\{5\}} + \lambda^{3,\{1\}} + \lambda^{3,\{1,5\}} = 1 \\
& \lambda^{iP} \geq 0 \qquad i \in \{1,2,3\}, \quad P \subseteq \mathcal{B}_i \\
& w_1, w_2, w_3, w_{12}, w_{13}, w_{23}, w_{123}, w_{14}, w_{15} \in \{0,1\}.
\end{aligned}
$$

□

We will pick an arbitrary node of T and assign it as the root node. This root node will be denoted as r. As is standard, for any node in the tree, the first neighbor that is visited on the path to the root node will be called the parent of the node and the node itself is the child of the parent.

Now, consider a relaxation LP where integrality constraints on $w[\mathcal{B}_i]$ are relaxed. In order to prove Theorem 1, it suffices to show that there exists an optimal solution with integral λ variables for any choice of the objective coefficients. The integrality of the remaining variables consequently follows. Therefore, it is convenient to project out the $w[\mathcal{B}_i]$ variables. This can be accomplished in the the following fashion:

- Rewriting any w_A variable in the objective in terms of the corresponding sum of λ variables. Note that this may not be unique.

- Equating various sums of λ variables that yield the same monomial among every child-parent pair (i, j) of nodes in the tree decomposition T. Due to Property (3) in the definition of tree decomposition, any common variables between two nodes are also common to every node in the path joining them. Thus, this operation effectively leads to equating monomials common in any two given nodes of the tree T.

Thus, the following formulation is obtained after projecting out the w_A variables:

$$
\begin{aligned}
\min \quad & \sum_{i\in[k]} \sum_{P\in\mathcal{P}_i} c^{iP}\lambda^{iP} \\
\text{s.t.} \quad & \sum_{T\in\mathcal{P}_i:T\supseteq P} \lambda^{iT} = \sum_{T\in\mathcal{P}_j:T\supseteq P} \lambda^{jT}, \qquad P\subseteq \mathcal{B}_i\cap\mathcal{B}_j, i \text{ is a child of } j \quad (7)\\
& \sum_{Q\in\mathcal{P}_i} \lambda^{iQ} = 1, \qquad i\in[k] \quad (8)\\
& \lambda^{iQ} \geq 0, \qquad i\in[k], Q\in\mathcal{P}_i,
\end{aligned}
$$

where (7) ensures that common monomials in adjacent nodes of the associated tree are equated in terms of the λ variables for the corresponding nodes. In particular, observe that for $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$ and $|P| > 0$, $w_P = \sum_{T\in\mathcal{P}_i:T\supseteq P} \lambda^{iT} = \sum_{T\in\mathcal{P}_j:T\supseteq P} \lambda^{jT}$. For $|P| = 0$ or $P = \emptyset$, (7) effectively equates the sum of all multipliers for nodes i and j.

Example 3. *In our running example, we relax integrality constraints. Furthermore, in the resulting LP, we project out w_1, w_2, w_3, w_{12}, w_{13}, w_{23}, w_{123}, w_{14}, and w_{15} by equating the expressions of w_1 in the vector equations. Here w_1 which is the only common variable between any two vector equations. We thus obtain the following LP relaxation:*

$$
\begin{aligned}
\min \quad & \lambda^{1,\{1,2,3\}} + \lambda^{1,\{1,4\}} + 4\lambda^{1,\{1,5\}} \\
\text{s.t.} \quad & \lambda^{2,\{1\}} + \lambda^{2,\{1,4\}} = \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}} \\
& \lambda^{2,\emptyset} + \lambda^{2,\{4\}} + \lambda^{2,\{1\}} + \lambda^{2,\{1,4\}} = \lambda^{1,\{2,3\}} + \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}} \\
& \lambda^{3,\{1\}} + \lambda^{3,\{1,5\}} = \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}} \\
& \lambda^{3,\emptyset} + \lambda^{3,\{5\}} + \lambda^{3,\{1\}} + \lambda^{3,\{1,5\}} = \lambda^{1,\{2,3\}} + \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}}
\end{aligned}
$$

$$\lambda^{1,\{2,3\}} + \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}} = 1$$
$$\lambda^{2,\emptyset} + \lambda^{2,\{4\}} + \lambda^{2,\{1\}} + \lambda^{2,\{1,4\}} = 1$$
$$\lambda^{3,\emptyset} + \lambda^{3,\{5\}} + \lambda^{3,\{1\}} + \lambda^{3,\{1,5\}} = 1$$
$$\lambda^{iP} \geq 0 \quad i \in \{1,2,3\}, P \subseteq \mathcal{B}_i$$

□

In order to prove the integrality of the above polytope, it is convenient to rewrite constraints (7) and (8) in a modified but equivalent fashion. The equivalent form has the property that any two chosen parent-child equations contain only different block variables. This equivalent form of equations is arrived at using elementary row operations. Both of our proofs rely on bringing the equations of a parent-child pair in its equivalent form, which is presented in the following claim.

Claim 1. *Consider two adjacent nodes of the tree decomposition with variable indices, say $\mathcal{B}_i$ and $\mathcal{B}_j$, where i is a child node of j. Then, constraints* (7) *and* (8) *can be equivalently expressed as:*

$$\sum_{T \in \mathcal{P}_i : T \cap (\mathcal{B}_i \cap \mathcal{B}_j) = P} \lambda^{iT} = \sum_{T \in \mathcal{P}_j : T \cap (\mathcal{B}_i \cap \mathcal{B}_j) = P} \lambda^{jT}, \qquad P \subseteq \mathcal{B}_i \cap \mathcal{B}_j \tag{9}$$
$$\sum_{P \in \mathcal{P}_j} \lambda^{jP} = 1 \tag{10}$$

Proof of claim: Observe first that the constraint (7) corresponding to $P = \emptyset$ is the following: $\sum_{P \in \mathcal{P}_i} \lambda^{iP} = \sum_{P \in \mathcal{P}_j} \lambda^{jP}$. This, together with constraint (8) corresponding to block j implies constraint (8) corresponding to block i. Therefore, due to the redundancy, we can keep just one set of constraint (8) corresponding to the parent node j as in (10), while discarding that of the child node i.

Next, we prove this claim by applying a sequence of $|\mathcal{B}_i \cap \mathcal{B}_j|$ elementary row operations to the system of equations (7) to arrive at the system of equations (9). Let R_P represent the row in constraint (7) corresponding to some $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$.

Algorithm 1 Row operations

$k \leftarrow |\mathcal{B}_i \cap \mathcal{B}_j|$
repeat
  $\mathcal{G}_k \leftarrow \{P \subseteq \mathcal{B}_i \cap \mathcal{B}_j : |P| \geq k\}$.
  $\mathcal{H}_k \leftarrow \{P \subseteq \mathcal{B}_i \cap \mathcal{B}_j : |P| = k-1\}$.
  for $Q \in \mathcal{H}_k$ **do**
    Replace row R_Q with $R_Q - \sum_{P \in \mathcal{G}_k} R_P$
  end for
  $k \leftarrow k-1$
until $k = 0$

At the end of the iteration for $k = \tilde{k}$, all the rows corresponding to $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$ such that $|P| \geq \tilde{k} - 1$ are of the form of the claim. This can be proven by reverse mathematical induction on iteration index k.

Base case: For $k = |\mathcal{B}_i \cap \mathcal{B}_j|$, the statement is true. The constraints corresponding to $|P| = |\mathcal{B}_i \cap \mathcal{B}_j|$ are already of the form of the claim. Now, observe that, for any row corresponding to $\{P : |P| = |\mathcal{B}_i \cap \mathcal{B}_j| - 1\}$) terms defined by $\{T : T \cap (\mathcal{B}_i \cap \mathcal{B}_j) = (\mathcal{B}_i \cap \mathcal{B}_j)\}$ in both LHS and RHS are canceled out by the row operations at the end of the first iteration ($k = |\mathcal{B}_i \cap \mathcal{B}_j|$). Thus, at the end of the first iteration, we get that all rows defined by $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$ such that $|P| \geq |\mathcal{B}_i \cap \mathcal{B}_j| - 1$ are of the form of equation (9).

Inductive step: Suppose at the end of iteration $k = \tilde{k}$ all the constraints defined by $\{P : |P| \geq \tilde{k} - 1\}$ are of the form of the claim. Then the row operations of iteration $k = \tilde{k} - 1$ yield a similar result for $\{P : |P| = \tilde{k} - 2\}$ by canceling out the terms defined by $\{T : |T \cap (\mathcal{B}_i \cap \mathcal{B}_j)| \geq \tilde{k} - 1\}$ for $P \in \{P : |P| = \tilde{k} - 2\}$. This completes the proof of the claim. □

Based on the claim above, we can replace constraint (7) with constraint (9) and we need the sum of multipliers equal to one (8) only for the root node, with index say r. Therefore, the claim implies the following equivalent reformulation of the problem:

$$\begin{aligned} \min \quad & \sum_{i \in [k]} \sum_{P \in \mathcal{P}_i} c^{iP} \lambda^{iP} \\ \text{s.t.} \quad & \sum_{T \in \mathcal{P}_i : T \cap (\mathcal{B}_i \cap \mathcal{B}_j) = P} \lambda^{iT} = \sum_{T \in \mathcal{P}_j : T \cap (\mathcal{B}_i \cap \mathcal{B}_j) = P} \lambda^{jT}, \\ & \qquad i \text{ is child of } j, P \subseteq \mathcal{B}_i \cap \mathcal{B}_j \end{aligned} \tag{11}$$

$$\sum_{P \in \mathcal{P}_r} \lambda^{rP} = 1 \tag{12}$$

$$\lambda^{iQ} \geq 0, \qquad i \in [k], Q \in \mathcal{P}_i.$$

We refer to the above formulation as $\mathcal{LP}$. The formulation $\mathcal{LP}$ has an important property that we record here.

Observation 1. *Constraint* (11) *has the following property. Each* λ^{jQ}, $Q \in \mathcal{P}_j$ *and* $\lambda^{iQ'}$, $Q' \in \mathcal{P}_i$ *appear only once in the block of constraints corresponding to a child-parent* (i, j) *pair. Specifically,* λ^{jQ} *appears in the constraint defined by* $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$ *such that* $Q \cap (\mathcal{B}_i \cap \mathcal{B}_j) = P$*; similarly for* $\lambda^{iQ'}$.

Example 4. *To illustrate the process of getting to the formulation* $\mathcal{LP}$ *in our running example, we note that the formulation can be rearranged in the form of* $\mathcal{LP}$ *by elementary row operations. The following is the resulting formulation of the form of* $\mathcal{LP}$.

$$\begin{aligned}
\min \quad & \lambda^{1,\{1,2,3\}} + \lambda^{2,\{1,4\}} + 4\lambda^{3,\{1,5\}} \\
\text{s.t.} \quad & \lambda^{2,\{1\}} + \lambda^{2,\{1,4\}} = \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}} && (13) \\
& \lambda^{2,\emptyset} + \lambda^{2,\{4\}} = \lambda^{1,\{2,3\}} && (14) \\
& \lambda^{3,\{1\}} + \lambda^{3,\{1,5\}} = \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}} && (15) \\
& \lambda^{3,\emptyset} + \lambda^{3,\{5\}} = \lambda^{1,\{2,3\}} && (16) \\
& \lambda^{1,\{2,3\}} + \lambda^{1,\{1,3\}} + \lambda^{1,\{1,2\}} + \lambda^{1,\{1,2,3\}} = 1 && (17) \\
& \lambda^{iP} \geq 0 \qquad \forall i, P
\end{aligned}$$

□

3. Exactness of the LP and proof of Theorem 1

We now present two proofs for the exactness of $\mathcal{LP}$.

3.1. A perturbation argument

Suppose for the sake of contradiction that there is a fractional extreme point $\bar{\lambda}$ of $\mathcal{LP}$. In order to obtain a contradiction, we will create two distinct points λ' and λ'' feasible for $\mathcal{LP}$ such that $\bar{\lambda} = (\lambda' + \lambda'')/2$. Without loss of generality, $\bar{\lambda}$ is such that the root node has a fractional assignment. Otherwise, the same argument can be applied starting from a node closest to the root node that has its corresponding variables fractional. For $u \in [0,1]$, define its fractionality as $\min\{u, 1-u\}$. Let $\epsilon > 0$ be the smallest fractionality among all fractional variables in $\bar{\lambda}$. Consider two fractional convex multiplier variables for the root node, say $\bar{\lambda}^{rT_1}$ and $\bar{\lambda}^{rT_2}$. We can perturb them by increasing one, say $\bar{\lambda}^{rT_1}$, and decreasing the other, say $\bar{\lambda}^{rT_2}$, by ϵ. That is, $(\lambda^{jT_1})' \leftarrow \bar{\lambda}^{jT_1} + \epsilon$ and $(\lambda^{jT_2})' \leftarrow \bar{\lambda}^{jT_2} - \epsilon$. If $\bar{\lambda}^{jT_1}$ and $\bar{\lambda}^{jT_2}$ are present in the same constraint of the form (11), we do not need to change anything, and we have a feasible neighbor solution λ'. Otherwise, for any child i of root r, pick two fractional multipliers, say $\bar{\lambda}^{iS_1}$ and $\bar{\lambda}^{iS_2}$, one from each constraint containing $\bar{\lambda}^{jT_1}$ and $\bar{\lambda}^{jT_2}$ respectively. This is possible because of Observation 1 regarding equations (11), that is, for a system of equations defined by a parent-child pair, each variable present in the system appears in exactly one equation. Similarly, perturbing $\bar{\lambda}^{iS_1}$ by the same amount ϵ and $\bar{\lambda}^{iS_2}$ by $-\epsilon$ ensures the feasibility of constraints (11) for the pair $i - r$. Propagate the perturbations throughout the tree down to the leaves to ensure that the overall system of equations (9) is satisfied. Thus, a perturbed feasible solution λ' is generated. A similar perturbation in the opposite direction gives the second feasible point λ''. The perturbed solutions satisfy $\bar{\lambda} = (\lambda' + \lambda'')/2$ as the perturbations were in the opposite direction. Furthermore, λ' and λ'' are distinct, completing the proof.

Example 5. *In our running illustrative example, suppose that, for the sake of contradiction, there is a fractional extreme point $\bar{\lambda}$ to $\mathcal{LP}$. Say $0 < \bar{\lambda}^{1,\{2,3\}} < 1$. Then, one of $\lambda^{1,\{1,3\}}$, $\lambda^{1,\{1,2\}}$, and $\lambda^{1,\{1,2,3\}}$ is strictly fractional. Say, without loss of generality, $0 < \lambda^{1,\{1,3\}} < 1$. Then, from (13), (14), (15), and (16), we have $0 < \lambda^{2,\{1\}} + \lambda^{2,\{1,4\}} < 1$ $0 < \lambda^{2,\emptyset} + \lambda^{2,\{4\}} < 1$, $0 < \lambda^{3,\{1\}} + \lambda^{3,\{1,5\}} < 1$, and $0 < \lambda^{3,\emptyset} + \lambda^{3,\{5\}} < 1$, respectively. It is easy to see that such a solution can be perturbed to generate two corresponding solutions λ', and λ'' such that $\bar{\lambda} = (\lambda' + \lambda'')/2$. This leads to a contradiction of the point being extreme.* □

3.2. Proof 2 (Complementary slackness): Dual to $\mathcal{LP}$ and corresponding DP-based formulation

We formulate the dual to $\mathcal{LP}$, and construct an optimal solution for the dual and the primal. Consider the following notation for the dual variables:

- Let ψ be the dual corresponding to the constraint $\sum_{P \in \mathcal{P}_r} \lambda^{rP} = 1$.
- Let ρ_P^{ij} be the dual variable corresponding to the constraint for P subset of $\mathcal{B}_i \cap \mathcal{B}_j$ where node i is child of node j.

From Observation 1, for block indices i, j we can define a function $\alpha^{ij} : 2^{\mathcal{P}_j} \to 2^{\mathcal{B}_i \cap \mathcal{B}_j}$ indicating the constraint index for the variable λ^{jQ} of the parent j. That is, $\alpha^{ij}(Q) = P$ iff the variable λ^{jQ} is present in constraint (11) defined by $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$. We also define the function $\beta^{ij} : 2^{\mathcal{B}_i \cap \mathcal{B}_j} \to 2^{\mathcal{P}_i}$ giving the set of child variables in a given constraint indexed $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$. That is, for any $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$, $\beta^{ij}(P) = \{Q \in \mathcal{P}_i : \lambda^{iQ}$ is present in constraint (9) defined by P$\}$. Finally, we denote the set of all the child nodes of node i as $\mathcal{C}(i)$. By convention, if i is a leaf node, then $\mathcal{C}(i) = \emptyset$.

Hence, the dual to $\mathcal{LP}$ can be formulated as:

$$\begin{aligned} \max \quad & \psi \\ \text{s.t.} \quad & \rho_P^{ij} \leq \min_{Q \in \beta^{ij}(P)} \left(c^{iQ} + \sum_{t \in \mathcal{C}(i)} \rho^{ti}_{\alpha^{ti}(Q)} \right), \quad i \in \mathcal{C}(j), P \subseteq \mathcal{B}_i \cap \mathcal{B}_j \\ & \psi \leq \min_{Q \in \mathcal{P}_r} \left(c^{rQ} + \sum_{t \in \mathcal{C}(r)} \rho^{tr}_{\alpha^{tr}(Q)} \right). \end{aligned}$$

Next, we construct $\bar{\lambda}$ and $\bar{\rho}$ values that are feasible to $\mathcal{LP}$ and its dual, respectively.

Dual variable values. Starting from leaf-parent pairs, say (i, j), set the corresponding duals $\bar{\rho}_P^{ij} = \min_{Q \in \beta^{ij}(P)} c^{iQ}$ for all $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$ and mark the (i, j) pair as processed. Now, recursively, for all child-parent pairs (i, j) such that all pairs (t, i), $t \in \mathcal{C}(i)$ are processed, set

$$\bar{\rho}_P^{ij} = \min_{Q \in \beta^{ij}(P)} (c^{iQ} + \sum_{t \in \mathcal{C}(i)} \bar{\rho}_{\alpha^{ti}(Q)})$$

for all $P \subseteq \mathcal{B}_i \cap \mathcal{B}_j$. Mark the pair (i, j) as processed. Finally, set $\bar{\psi} = \min_{P \in \mathcal{P}_r} (c^{rP} + \sum_{t \in \mathcal{C}(r)} \bar{\rho}^{tr}_{\alpha^{tr}(P)})$.

Primal variable values. Having set the values of the dual variables, we set the primal variables as follows. For the primal variables of the root node, set the primal variable corresponding to one of the active dual constraints among $\bar{\psi} \leq (c^{rQ} + \sum_{t \in \mathcal{C}(r)} \bar{\rho}^{rt}_{\alpha^{rt}(Q)}) Q \in \mathcal{P}_r$ as 1 and the remaining as 0. That is $\bar{\lambda}^{rQ} = 1$ for a $Q \in \mathcal{P}_r$ satisfying $\bar{\psi} = c^{rQ} + \sum_{t \in \mathcal{C}(r)} \bar{\rho}^{tr}_{\alpha^{tr}(Q)}$. Now, for any node j with feasible integer primal variable values, we assign the primal variable values of its child (say i) as follows. Consider the equality constraints (11) for the parent-child pair (j, i). Exactly one of these constraints will have a right-hand side (RHS) equal to 1, depending on the variable that was assigned 1 in the parent node j. For the constraint with 1 in the RHS (say defined by $\tilde{P} \subseteq \mathcal{B}_i \cap \mathcal{B}_j$), consider the child node's primal variables λ^{iT} present in the left-hand side (i.e., $\{\ T \in \mathcal{P}_i : T \in \beta^{ij}(\tilde{P})\}$). Set the child primal variable (λ^{iT}) corresponding to an active dual constraint ($\bar{\rho}^{ij}_{\tilde{P}} \leq c^{iT} + \sum_{t \in \mathcal{C}(i)} \bar{\rho}^{ti}_{\alpha^{ti}(T)}$) as 1 and the remaining primal variables of the node (node i) as 0. By propagating this type of assignment, we assign all the primal variables. By construction, this primal and dual assignment ($\bar{\lambda}$ and $\bar{\rho}$) is feasible and satisfies complementary slackness. This implies optimality. Thus, for any choice of objective coefficient of the primal, there exists an optimal solution of the primal which is integral. Hence, the polyhedral feasible region of $\mathcal{LP}$ is integral.

Example 6. *In our running example, introducing duals $\rho^{12}_{\{1\}}$, $\rho^{12}_{\emptyset}$, $\rho^{13}_{\{1\}}$, $\rho^{13}_{\emptyset}$, and Ψ for the constraints* (13)*,* (14)*,* (15)*,* (16)*, and* (17)*, respectively, we obtain the dual as:*

$$\begin{aligned}
\max \quad & \Psi \\
\text{s.t.} \quad & \rho^{12}_{\{1\}} \leq \min(0, 1) \\
& \rho^{12}_{\emptyset} \leq \min(0, 0) \\
& \rho^{13}_{\{1\}} \leq \min(0, 4) \\
& \rho^{13}_{\emptyset} \leq \min(0, 0) \\
& \Psi \leq \min(\rho^{12}_{\{1\}} + \rho^{13}_{\{1\}},\ \rho^{12}_{\emptyset} + \rho^{13}_{\emptyset},\ \rho^{12}_{\{1\}} + \rho^{13}_{\{1\}},\ 1 + \rho^{12}_{\{1\}} + \rho^{13}_{\{1\}})
\end{aligned}$$

For the above primal-dual pair, with the construction given in the proof, we obtain $\bar{\rho}^{12}_{\{1\}} = 0$, $\bar{\rho}^{12}_{\emptyset} = 0$, $\bar{\rho}^{13}_{\{1\}} = 0$, $\bar{\rho}^{13}_{\emptyset} = 0$, and $\bar{\Psi} = 0$ for the dual whereas $\bar{\lambda}^{1,\{2,3\}} = 1$, $\bar{\lambda}^{1,\{1,3\}} = 0$, $\bar{\lambda}^{1,\{1,2\}} = 0$, $\bar{\lambda}^{1,\{1,2,3\}} = 0$, $\bar{\lambda}^{2,\emptyset} = 1$, $\bar{\lambda}^{2,\{4\}} = 0$, $\bar{\lambda}^{2,\{1\}} = 0$, $\bar{\lambda}^{2,\{1,4\}} = 0$ and $\bar{\lambda}^{3,\emptyset} = 1$, $\bar{\lambda}^{3,\{5\}} = 0$, $\bar{\lambda}^{3,\{1\}} = 0$, and $\bar{\lambda}^{3,\{1,5\}} = 0$ for the primal. The solution satisfies primal-dual complementary slackness and hence is optimal. In addition, the primal solutions are integral.

Since we can produce an optimal dual solution that is integral for every integral objective function of the primal, we obtain the following result.

Corollary 2. The representation of the feasible region of $\mathcal{LP}$ is totally dual integral.

References

Bienstock, D., Muñoz, G., 2018. LP formulations for polynomial optimization problems. SIAM Journal on Optimization 28, 1121–1150.

Bodlaender, H.L., Jansen, K., 2000. On the complexity of the maximum cut problem. Nordic Journal of Computing 7, 14–31.

Boros, E., Hammer, P.L., 2002. Pseudo-boolean optimization. Discrete Applied Mathematics 123, 155–225.

Cifuentes, D., Dey, S.S., Xu, J., 2024. Lagrangian dual with zero duality gap that admits decomposition. arXiv preprint arXiv:2411.12085 .

Crama, Y., Hammer, P.L., 2011. Boolean functions: Theory, algorithms, and applications. Cambridge University Press.

Del Pia, A., Khajavirad, A., 2024. Beyond hypergraph acyclicity: limits of tractability for pseudo-boolean optimization. arXiv preprint arXiv:2410.23045 .

Elloumi, S., Lambert, A., Lazare, A., 2021. Solving unconstrained 0-1 polynomial programs through quadratic convex reformulation. Journal of Global Optimization 80, 231–248.

Faenza, Y., Muñoz, G., Pokutta, S., 2022. New limits of treewidth-based tractability in optimization. Mathematical Programming 191, 559–594.

Gerber, M.U., Kobler, D., 2003. Algorithms for vertex-partitioning problems on graphs with fixed clique-width. Theoretical Computer Science 299, 719–734.

Kochenberger, G., Hao, J.K., Glover, F., Lewis, M., Lü, Z., Wang, H., Wang, Y., 2014. The unconstrained binary quadratic programming problem: A survey. Journal of Combinatorial Optimization 28, 58–81.

Laurent, M., 2003. A comparison of the Sherali–Adams, Lovász–Schrijver, and Lasserre relaxations for 0–1 programming. Mathematics of Operations Research 28, 470–496.

Laurent, M., 2009. Sums of squares, moment matrices and optimization over polynomials. *In M. Putinar and S. Sullivant (eds.)*, Emerging applications of algebraic geometry, The IMA Volumes in Mathematics and its Applications, vol 149 *Springer, New York, NY* , 157–270.

Liers, F., Marinari, E., Pagacz, U., Ricci-Tersenghi, F., Schmitz, V., 2010. A non-disordered glassy model with a tunable interaction range. Journal of Statistical Mechanics: Theory and Experiment 2010, L05003.

Lovász, L., Schrijver, A., 1991. Cones of matrices and set-functions and 0–1 optimization. SIAM Journal on Optimization 1, 166–190.

Martin, R.K., Rardin, R.L., Campbell, B.A., 1990. Polyhedral characterization of discrete dynamic programming. Operations Research 38, 127–138.

MINLPLib, 2024. MINLPLib—A library of mixed-integer and continuous nonlinear programming instances. `https://minlplib.org/instances.html`. Accessed: 2024-06-29.

Sherali, H.D., Adams, W.P., 1990. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. SIAM Journal on Discrete Mathematics 3, 411–430.

Telle, J.A., Proskurowski, A., 1997. Algorithms for vertex partitioning problems on partial k-trees. SIAM Journal on Discrete Mathematics 10, 529–550.

Wainwright, M.J., Jordan, M.I., 2004. Treewidth-based conditions for exactness of the Sherali-Adams and Lasserre relaxations. Technical Report. Technical Report 671, University of California, Berkeley.