On the Lifted Multicut Polytope for Trees

Jan-Hendrik Lange and Björn Andres

1 Max Planck Institute for Informatics, Saarbrücken, Germany
jan-hendrik.lange@mpi-inf.mpg.de
2 Technical University Dresden, Dresden, Germany

Abstract. We study the lifted multicut problem restricted to trees, which is NP-hard in general and solvable in polynomial time for paths. In particular, we characterize facets of the lifted multicut polytope for trees defined by the inequalities of a canonical relaxation. Moreover, we present an additional class of inequalities associated with paths that are facet-defining. Taken together, our facets yield a complete totally dual integral description of the lifted multicut polytope for paths. This description establishes a connection to the combinatorial properties of alternative formulations such as sequential set partitioning.

1 Introduction

The lifted multicut problem [12] is a graph partitioning model based on real-valued costs attributed to pairs of nodes that are in distinct components. It has been applied successfully to inference tasks in areas such as image segmentation [2,17], object tracking [22] and motion segmentation [16]. In this paper we study the lifted multicut problem and its associated polyhedral geometry under the restriction that the underlying graph is a tree. This constitutes an extreme special case, since it is natural to consider only connected graphs and trees are minimally connected graphs. The resulting tree partition problem is equivalently formulated as the minimization of a multi-linear polynomial which exhibits a sparsity pattern that is determined by the tree. Therefore, it is clearly NP-hard, as we show in Sect. 3. In Sect. 4, we study the facial structure of the lifted multicut polytope for trees. We introduce a canonical relaxation in terms of node triplets and characterize under which circumstances the basic inequalities are facet-defining. Furthermore, we present another class of facet-defining inequalities, which we call intersection inequalities. Finally, we show that the facets that we present yield a complete totally dual integral description in the case of paths (Sect. 5).

Overall, we contribute to the study of graph partition problems an analysis of the facial structure of the lifted multicut polytope for the extreme case of minimally connected graphs such as trees and paths. Our results establish connections between the lifted multicut problem on trees and pseudo-Boolean optimization as well as sequential set partitioning. Furthermore, our insights can accelerate solvers for the problem based on integer linear programming.
2 Related Work

Multicut Polytopes and Correlation Clustering. Partitions of a graph into an unconstrained number of components based on similarities or dissimilarities of neighboring nodes is referred to as weighted correlation clustering. Weighted correlation clustering has been studied for complete graphs [1] as well as for general graphs [6]. Due to the fact that any partition of a graph into connected components is characterized by the mathematical notion of a multicut, correlation clustering is closely related to the study of multicuts. The combinatorial polytopes associated with the multicuts of a graph have been studied, among others, most notably by [5, 7, 8, 10]. The more general case in which similarities or dissimilarities of non-neighboring nodes are taken into account as well has been introduced in [12]. The latter, more expressive formulation has found a number of applications in computer vision [2, 16, 17, 22].

Pseudo-Boolean Optimization. The equivalent formulation of the problem we study in the form of an unconstrained minimization of a multi-linear polynomial connects our work to the field of pseudo-Boolean optimization. The optimization of pseudo-Boolean functions plays an important role in machine learning, for instance, in MAP inference for computer vision. The general problem can be reduced to the quadratic case [3, 4], which is responsible for the $\text{NP}$-hardness of the problem. The combinatorial polytope associated with the linearization of quadratic pseudo-Boolean functions was studied extensively by [18]. Recent research also considers the linearization of more general multi-linear forms [19]. Computational approaches to pseudo-Boolean optimization based on the roof duality bound [11, 14] have been quite successful in practice [20].

Sequential Set Partitioning. A set partitioning problem where the elements are assumed to adhere to a linear order has been studied by Kernighan [15], who devises a dynamic program to solve the problem in polynomial time. The algorithm essentially solves a shortest path problem in a directed acyclic graph. The corresponding integer linear programming formulation admits a totally unimodular constraint matrix [13]. We derive a complete polyhedral description for the equivalent formulation as a lifted multicut problem on paths.

3 Tree Partition Problem

Let $T = (V, E)$ be a tree. We use the short-hand notation $uv = \{u, v\}$ for any pair of distinct vertices $u, v \in V$, which may or may not correspond to an edge. When convenient, we write $n = |E|$ and $m = \binom{n}{2} = n(n-1)/2$ for the number of edges and the total number of vertex pairs, respectively. For any pair of distinct nodes $u, v \in V$, denote by $P_{uv}$ the unique path from $u$ to $v$ in $T$. Moreover, we denote by $d(u, v)$ the distance of $u$ and $v$ in $T$, i.e. the length of $P_{uv}$.

A multicut is a set of edges that are between the components of a partition of a graph [7]. The characteristic vector $x \in \{0, 1\}^E$ of a multicut (where $x_e = 1$...
if \( e \in E \) is cut) is lifted to the space \( \{0, 1\}^{\binom{n}{2}} \) by setting \( x_{uv} = 1 \) if, and only if \( u \) and \( v \) are in distinct components of the underlying partition [12].

The lifted multicut problem on trees is to find a partition of a tree that minimizes a sum of costs associated with pairs of nodes that are in distinct components. It is formulated as an integer linear program as follows.

**Definition 1** (Lifted Multicut Problem). The lifted multicut polytope \( \text{LMC} \) w.r.t. \( T \) is defined as the convex hull of all \( x \in \{0, 1\}^{\binom{n}{2}} \) that satisfy the path and cut inequalities:

\[
\begin{align*}
x_{uv} &\leq \sum_{e \in P_{uv}} x_e \quad \forall u, v \in V, \ d(u, v) \geq 2, \quad \text{(path)} \\
x_e &\leq x_{uv} \quad \forall u, v \in V, \ d(u, v) \geq 2, \ \forall e \in P_{uv}. \quad \text{(cut)}
\end{align*}
\]

The lifted multicut problem w.r.t. \( T \) and \( \theta \in \mathbb{R}^{\binom{n}{2}} \) is defined as

\[
\min_{x \in \text{LMC}} \sum_{uv \in \binom{V}{2}} \theta_{uv} x_{uv}.
\]

Note that any vector \( x \in \{0, 1\}^E \) is the characteristic vector of a multicut of \( T \) as any edge set of a tree defines a partition. The path and cut inequalities ensure that for any distinct pair of nodes \( u, v \in V \) it holds that \( x_{uv} = 1 \) if, and only if the path \( P_{uv} \) is cut at any of its edges.

The lifted multicut problem on \( T \) can be equivalently formulated as the minimization of a particular multi-linear polynomial over binary inputs, which we refer to as tree partition problem.

**Definition 2** (Tree Partition Problem). Let \( T = (V, E) \) be a tree and \( \bar{\theta} \in \mathbb{R}^{\binom{n}{2}} \). The optimization problem

\[
\min_{y \in \{0, 1\}^E} \sum_{uv \in \binom{V}{2}} \bar{\theta}_{uv} \prod_{e \in P_{uv}} y_e
\]

is called the instance of the tree partition problem w.r.t. \( T \) and \( \bar{\theta} \). If \( T \) is a path, then we also refer to (TPP) as the path partition problem w.r.t. \( T \) and \( \bar{\theta} \).

It is straightforward to see, by a change of variables, that the problems (TPP) and (LMP) are equivalent (up to a constant).

**Lemma 1.** The vector \( y \in \{0, 1\}^E \) is a solution of problem (TPP) w.r.t. the tree \( T = (V, E) \) and costs \( \bar{\theta} \in \mathbb{R}^{\binom{n}{2}} \) if, and only if, the unique \( x \in \text{LMC} \) such that \( x_e = 1 - y_e \) for all \( e \in E \) is a solution of problem (LMP) w.r.t. \( T \) and the cost vector \( \theta = -\bar{\theta} \).
Proof. For any distinct pair of nodes $u, v \in V$, we set

$$x_{uv} = 1 - \prod_{e \in P_{uv}} y_e$$

which implies

$$x_{uv} = 0 \iff \forall e \in P_{uv}: y_e = 1 \iff \forall e \in P_{uv}: x_e = 0.$$  \hspace{1cm} (2)

Therefore, we can reformulate problem (TPP) in terms of the variables $x_{uv}$ by transforming the objective function according to

$$\tilde{\theta}_{uv} \prod_{e \in P_{uv}} y_e = -\tilde{\theta}_{uv}(1 - \prod_{e \in P_{uv}} y_e) + \tilde{\theta}_{uv} = -\tilde{\theta}_{uv} x_{uv} + \tilde{\theta}_{uv}.$$  \hspace{1cm} (3)

This leads to the linear combinatorial optimization problem

$$\min_{x \in \text{LMC}} \sum_{uv \in \binom{V}{2}} \theta_{uv} x_{uv} + \tilde{\theta}_{uv},$$

where the definition of LMC captures the relationship (2). \hfill \Box

Complexity. The tree partition problem, and thus problem (LMP), is \textsc{np}-hard in general (Lemma 2 below). However, the path partition problem is solvable in polynomial time \cite{15}. \hfill \Box

Lemma 2. The tree partition problem is \textsc{np}-hard.

Proof. If $T$ is a star (see Fig. 1a for an example), then problem (TPP) is equivalent to the unconstrained binary quadratic program with $|E|$ variables, which is well-known to be \textsc{np}-hard. \hfill \Box

4 Lifted Multicut Polytope for Trees

In this section we study the facial structure of the lifted multicut polytope LMC. We characterize all trivial facets and offer an outer relaxation of LMC that is tighter than the standard relaxation given by \cite{12}. In Sect. 5, we show that our results yield a complete totally dual integral (TDI) description of the lifted multicut polytope for paths.
We denote the standard outer relaxation of LMC by

\[ P_0 = \{ x \in [0, 1]^{\binom{n}{2}} \mid x_{uv} \leq \sum_{e \in P_{uv}} x_e \quad \forall u, v \in V, \ d(u, v) \geq 2, \]

\[ x_e \leq x_{uv} \quad \forall u, v \in V, \ d(u, v) \geq 2, \forall e \in P_{uv} \}, \]

which is obtained by dropping the integrality constraints from the definition of LMC. In particular, any \( x \in P_0 \) satisfies all path and cut inequalities. Let \( \vec{u}(v) \) be the first node on the path \( P_{uv} \) that is different from both \( u \) and \( v \) (cf. Fig. 1b) and consider the polytope

\[ P_1 = \{ x \in [0, 1]^{\binom{n}{2}} \mid x_{uv} \leq x_{u,\vec{u}(v)} + x_{\vec{u}(v),v} \quad \forall u, v \in V, \ d(u, v) \geq 2, \]

\[ x_{\vec{u}(v),v} \leq x_{uv} \quad \forall u, v \in V, \ d(u, v) \geq 2 \}. \]

This description is canonical in the sense that it only considers a quadratic number of node triplets, namely those which feature two neighboring nodes and an arbitrary third node. The following lemma states that \( P_1 \) is indeed an outer relaxation of LMC that is at least as tight as \( P_0 \).

**Lemma 3.** It holds that \( \text{LMC} \subseteq P_1 \subseteq P_0 \).

**Proof.** We show first that \( \text{LMC} \subseteq P_1 \). For this purpose, let \( x \in \text{LMC} \cap \mathbb{Z}^m \) be a vertex of LMC. If \( x_{uv} > x_{u,\vec{u}(v)} + x_{\vec{u}(v),v} \) for some \( u, v \in V \), then \( x_{uv} = 1 \) and \( x_{u,\vec{u}(v)} = x_{\vec{u}(v),v} = 0 \). This contradicts the fact that \( x \) satisfies all cut inequalities w.r.t. \( \vec{u}(v) \) and the path inequality w.r.t. \( u, v \). If \( x_{\vec{u}(v),v} > x_{uv} \) for some \( u, v \in V \), then \( x_{\vec{u}(v),v} = 1 \) and \( x_{uv} = 0 \). This contradicts the fact that \( x \) satisfies all cut inequalities w.r.t. \( u, v \) and the path inequality w.r.t. \( \vec{u}(v),v \). It follows that \( x \in P_1 \).

Now, we show that \( P_1 \subseteq P_0 \). Let \( x \in P_1 \). We need to show that \( x \) satisfies all path and cut inequalities. Let \( u, v \in V \) with \( d(u, v) \geq 2 \). We proceed by induction on \( d(u, v) \). If \( d(u, v) = 2 \), then the path and cut inequalities are directly given by the definition of \( P_1 \) (for the two possible orderings of \( u \) and \( v \)). If \( d(u, v) > 2 \), then the path inequality is obtained from \( x_{uv} \leq x_{u,\vec{u}(v)} + x_{\vec{u}(v),v} \) and the induction
hypothesis for the pair $\vec{u}(v), v$, since $d(\vec{u}(v), v) = d(u, v) - 1$. Similarly, for any edge $e$ on the path from $u$ to $v$, we obtain the cut inequality w.r.t. $e$ by using the induction hypothesis and $x_{\vec{u}(v), v} \leq x_{uv}$ such that (w.l.o.g.) $e$ is on the path from $\vec{u}(v)$ to $v$. It follows that $x \in P_0$.

4.1 Facets

In this section, we show which inequalities in the definition of $P_1$ define facets of LMC. Moreover, we present another type of inequalities associated with paths in $T$, which define facets of LMC. We note that further facets can be established by the connection of LMC to the multi-linear polytope and, as a special case, the Boolean quadric polytope [18].

**Lemma 4.** The inequality

$$x_{uv} \leq x_{u, \vec{u}(v)} + x_{\vec{u}(v), v}$$

for some $u, v \in V$ defines a facet of LMC if, and only if, $d(u, v) = 2$.

**Proof.** First, suppose $d(u, v) = 2$. Then $P_{uv}$ is a path of length 2 and thus chordless in the complete graph on $V$. Hence, the facet-defining property follows directly from [12, Theorem 10]. Now, suppose $d(u, v) > 2$ and let $x \in LMC$ be such that (5) is satisfied with equality. We show that this implies

$$x_{uv} + x_{\vec{u}(v), \vec{v}(u)} = x_{u, \vec{v}(u)} + x_{\vec{u}(v), v}.$$  \hspace{1cm} (6)

Then the face of LMC induced by (5) has dimension at most $m - 2$ and hence cannot be a facet. There are two possible cases: Either $x_{uv} = x_{u, \vec{u}(v)} = x_{\vec{u}(v), v} = 1$, then $x_{uw} = x_{\vec{u}(v), w} = 1$ as well, or

**Lemma 5.** The inequality

$$x_{\vec{u}(v), v} \leq x_{uv}$$

for some $u, v \in V$ defines a facet of LMC if, and only if, $v$ is a leaf of $T$.

**Proof.** First, suppose $v$ is not a leaf of $T$ and let $x \in LMC$ be such that (7) is satisfied with equality. Since $v$ is not a leaf, there exists a neighbor $w \in V$ of $v$ such that $P_{\vec{u}(v), w}$ is a subpath of $P_{\vec{u}(v), w}$. We show that $x$ additionally satisfies the equality

$$x_{uw} = x_{\vec{u}(v), w}$$  \hspace{1cm} (8)

and thus the face of LMC induced by (7) cannot be a facet. There are two possible cases: Either $x_{uv} = x_{\vec{u}(v), v} = 1$, then $x_{uw} = x_{\vec{u}(v), w} = 1$ as well, or
Lemma 6. For any distinct $u, v \in V$, the inequality $x_{uv} \leq 1$ defines a facet of LMC if, and only if, both $u$ and $v$ are leaves of $T$. Moreover, none of the inequalities $0 \leq x_{uv}$ define facets of LMC.

Proof. We apply the more general characterization given by [12, Theorem 8] and [12, Theorem 9]. The nodes $u, v \in V$ are a pair of $ww'$-cut-vertices for some vertices $w, w' \in V$ (with at least one being different from $u$ and $v$) if, and only if, $u$ or $v$ is not a leaf of $V$. Thus, the claim follows from [12, Theorem 9]. The second assertion follows from [12, Theorem 8] and the fact that we lift to the complete graph on $V$. \hfill \Box

Intersection Inequalities. We present another large class of non-trivial facets of LMC. For any $u, v \in V$ with $d(u, v) \geq 3$ consider the inequality

$$x_{uv} + x_{\bar{u}(v), \bar{v}(u)} \leq x_{u, \bar{v}(u)} + x_{\bar{u}(v), v},$$

(10)

which we refer to as intersection inequality. As an example consider the graph depicted in Fig. 1c with $u = 0$ and $v = 3$.

Lemma 7. Any intersection inequality is valid for LMC.

Proof. Let $x \in \text{LMC} \cap \mathbb{Z}^m$ and suppose that either $x_{u, \bar{v}(u)} = 0$ or $x_{\bar{u}(v), v} = 0$ for some $u, v \in V$ with $d(u, v) \geq 3$. Then, since $x$ satisfies all cut inequalities w.r.t. $u, \bar{v}(u)$, respectively $\bar{u}(v), v$, and the path inequality w.r.t. $\bar{u}(v), \bar{v}(u)$, it must hold that $x_{\bar{u}(v), \bar{v}(u)} = 0$. Moreover, if even $x_{u, \bar{v}(u)} = 0 = x_{\bar{u}(v), v}$, then, by the same reasoning, we have $x_{uv} = 0$ as well. Hence, $x$ satisfies (10). \hfill \Box

Lemma 8. Any intersection inequality defines a facet of LMC.
Proof. Let $\Sigma$ be the face of LMC induced by (10) for some $u, v \in V$ with $d(u, v) \geq 3$. We need to prove that $\Sigma$ has dimension $m - 1$, which can be done explicitly by showing that we can construct $m - 1$ distinct indicator vectors $\mathbb{1}_{uw^i}$ for $w, w^i \in V$ as linear combinations of elements from the set $S = \{x \in \text{LMC} \cap \mathbb{Z}^m \mid x_{uv} + x_{u, \bar{v}}(u) = x_{u, \bar{v}}(u) + x_{u, \bar{v}}(u, v)\}$.

We employ the same construction as used in the proof of Lemma 5. Observe that for any feasible $x \in \text{LMC} \cap \mathbb{Z}^m$ with $x \notin S$, we must have that $x_{u, \bar{v}}(v), \bar{v}(u) = 0$ and $x_{u, \bar{v}}(v) = x_{u, \bar{v}}(u, v) = 1$. This means that $x_{u, \bar{v}}(v), \bar{v}(u) \notin S$ in the construction. Thus, we simply omit $\mathbb{1}_{u, \bar{v}}(v), \bar{v}(u)$ from the construction. By the same reasoning, we conclude that the $m$ constructed vectors $\{\mathbb{1} \cup \{x_{uv} \mid uv \in \binom{V}{2} \setminus \{u, \bar{v}(u)\}\}\}$ are affine independent, so the claim follows. $\square$

5 Lifted Multicut Polytope for Paths

In this section we show that the facets established in the previous section yield a complete description of LMC when $T$ is a path. To this end, suppose that $V = \{0, \ldots, n\}$ and $E = \{\{i, i + 1\} \mid i \in \{0, \ldots, n - 1\}\}$ are linearly ordered. Therefore, $T = (V, E)$ is path. We consider only paths of length $n \geq 2$, since for $n = 1$, the polytope LMC $= [0, 1]$ is simply the unit interval. Let $P_{\text{path}}$ be the polytope of all $x \in \mathbb{R}^{\binom{V}{2}}$ that satisfy

\begin{align}
    x_{0n} & \leq 1, \quad \text{ (11)} \\
    x_{in} & \leq x_{i-1, n} \quad \forall i \in \{1, \ldots, n-1\}, \quad \text{ (12)} \\
    x_{0i} & \leq x_{0, i+1} \quad \forall i \in \{1, \ldots, n-1\}, \quad \text{ (13)} \\
    x_{i-1, i+1} & \leq x_{i-1, i} + x_{i, i+1} \quad \forall i \in \{1, \ldots, n-1\}, \quad \text{ (14)} \\
    x_{j, k} + x_{j+1, k-1} & \leq x_{j+1, k} + x_{j, k-1} \quad \forall j, k \in \{0, \ldots, n\}, \; j < k - 2. \quad \text{ (15)}
\end{align}

Note that the system of inequalities (11)–(15) consists precisely of those inequalities which we have shown to define facets of LMC in the previous section. We first prove that $P_{\text{path}}$ indeed yields an outer relaxation of LMC.

Lemma 9. It holds that LMC $\subseteq P_{\text{path}} \subseteq P_1$.

Proof. First, we show that LMC $\subseteq P_{\text{path}}$. Let $x \in \text{LMC} \cap \mathbb{Z}^m$, then $x$ satisfies (11) and (14) by definition. Suppose $x$ violates (12), then $x_{in} = 1$ and $x_{i-1, n} = 0$. This contradicts the fact that $x$ satisfies all cut inequalities w.r.t. $i, n$ and the path inequality w.r.t. $i, n$. So, $x$ must satisfy (12) and, by symmetry, also (13). It follows from Lemma 7 that $x$ satisfies (15) as well and thus $x \in P_{\text{path}}$.

Next, we prove that $P_{\text{path}} \subseteq P_1$. To this end, let $x \in P_{\text{path}}$. We show that $x$ satisfies all inequalities (7). Let $u, v \in V$ with $u < v - 1$. We need to prove that both $x_{u+1, v} \leq x_{uv}$ and $x_{u, v-1} \leq x_{uv}$ hold. For reasons of symmetry, it suffices to show only $x_{u+1, v} \leq x_{uv}$. We proceed by induction on the distance of $u$ from $n$. If $v = n$, then $x_{u+1, n} \leq x_{un}$ is given by (12). Otherwise, we use (15) for $j = u$ and $k = v + 1$ and the induction hypothesis on $v + 1$:
\[ x_{uv} + x_{u+v,v+1} \geq x_{u+1,v} + x_{u,v+1} \quad (16) \]
\[ \geq x_{u+1,v} + x_{u+1,v+1} \quad (17) \]
\[ \implies x_{uv} \geq x_{u+1,v,v+1} \quad (18) \]

It remains to show that \( x \) satisfies all inequalities (5). Let \( u, v \in V \) with \( u < v - 1 \). We proceed by induction on \( d(u, v) = u - v \). If \( d(u, v) = 2 \), then (5) is given by (14). If \( d(u, v) > 2 \), then we use (15) for \( j = u \) and \( k = v \) as well as the induction hypothesis on \( u, v - 1 \), which have distance \( d(u, v) - 1 \):

\[ x_{uv} + x_{u+1,v-1} \leq x_{u+1,v} + x_{u,v-1} \quad (19) \]
\[ \leq x_{u+1,v} + x_{u,u+1} + x_{u+1,v-1} \quad (20) \]
\[ \implies x_{uv} \leq x_{u,u+1} + x_{u+1,v} \quad (21) \]

Hence, \( x \in P_1 \), which concludes the proof. \( \square \)

Fig. 2. Illustration of the extension used in the proof of Theorem 1. The additional terms \( x_{ii}, x_{-\infty,i}, x_{i,\infty} \) and \( x_{-\infty,\infty} \) correspond to the thin green edges. (Color figure online)

As our main result, we prove that \( P_{\text{path}} \) is in fact a complete description of LMC and, moreover, it is \textit{totally dual integral}. For an extensive reference on the subject of total dual integrality we refer the reader to [21].

**Definition 3.** A system of linear inequalities \( Ax \leq b \) with \( A \in \mathbb{Q}^{k \times m} \), \( b \in \mathbb{Q}^k \) is called \textit{totally dual integral} (TDI) if for any \( c \in \mathbb{Z}^m \) such that the linear program \( \max \{ c^\top x \mid Ax \leq b \} \) is feasible and bounded, there exists an integral optimal dual solution.

Total dual integrality is an important concept in polyhedral geometry as it provides a sufficient condition on the integrality of polyhedra according to the following fact.

**Fact 1** ([9]). If \( Ax \leq b \) is totally dual integral and \( b \) is integral, then the polytope defined by \( Ax \leq b \) is integral.

**Theorem 1.** The system (11)–(15) is totally dual integral.

**Proof.** We rewrite system (11)–(15) more compactly in the following way. Introduce two artificial nodes \(-\infty\) and \( \infty \) where we associate \(-\infty\) to any index less than 0 and \( \infty \) to any index greater than \( n \). Moreover, we introduce variables \( x_{ii} \)
for all $0 \leq i \leq n$ as well as $x_{-\infty, i}$ and $x_{i, \infty}$ for all $1 \leq i \leq n - 1$ and finally $x_{-\infty, \infty}$ (cf. Fig. 2). Now, the system (11)–(15) is equivalent to the system

$$x_{j,k} + x_{j+1,k-1} \leq x_{j+1,k} + x_{j,k-1} \quad \forall j,k \in \{-\infty,0,\ldots,n,\infty\}, \ j \leq k - 2$$

(22)
given the additional equality constraints

$$x_{ii} = 0 \quad \forall 0 \leq i \leq n,$$
$$x_{-\infty,i} = 1 \quad \forall 1 \leq i \leq n - 1,$$
$$x_{i,\infty} = 1 \quad \forall 1 \leq i \leq n - 1,$$
$$x_{-\infty,\infty} = 1.$$

Let the system defined by (22)–(26) be represented in matrix form as $Ax \leq a, Bx = b$. Note that $P_{\text{path}}$ is non-empty and bounded. Thus, to establish total dual integrality, we need to show that for any $\theta \in \mathbb{Z}^{m+3n}$ the dual program

$$\min\{a^T y + b^T z \mid A^T y + B^T z = \theta, y \geq 0\}$$

(27)
has an integral optimal solution. Here, the $y$ variables, indexed by $j, k$, correspond to the inequalities (22) and the $z$ variables, indexed by pairs of $i, -\infty$ and $\infty$, correspond to the equations (23)–(26). Then, the equation system $A^T y + B^T z = \theta$ translates to

$$y_{i-1,i+1} + z_{i,i} = \theta_{i,i}$$
$$y_{i-1,i+2} - y_{i-1,i+1} - y_{i,i+2} = \theta_{i,i+1}$$
$$y_{i-1,\ell+1} - y_{i-1,\ell} - y_{i,\ell+1} + y_{i,\ell} = \theta_{i,\ell}$$
$$-y_{-\infty,i+1} + y_{-\infty,i} + z_{-\infty,i} = \theta_{-\infty,i}$$
$$-y_{i-1,\infty} + y_{i,\infty} + z_{i,\infty} = \theta_{i,\infty}$$
$$y_{-\infty,\infty} + z_{-\infty,\infty} = \theta_{-\infty,\infty},$$

(28)–(33)

where (28)–(30) hold for all $0 \leq i < i + 1 < \ell \leq n$ and (31), (32) hold for all $1 \leq i \leq n - 1$. Observe that (28) includes only $y$ variables with indices of distance 2, (29) couples $y$ variables of distance 3 with those of distance 2 and (30) couples the remaining $y$ variables of any distance $d > 3$ with those of distance $d - 1$ and $d - 2$. Hence, any choice of values for the free variables $z_{ij}$ completely determines all $y$ variables. This means we can eliminate $y$ and reformulate the dual program entirely in terms of the $z$ variables, as follows. It holds that

$$0 \leq y_{i-1,i+1} = \theta_{i,i} - z_{i,i} \quad \forall 0 \leq i \leq n,$$
$$0 \leq y_{i-1,i+2} = \theta_{i,i+1} + \theta_{i,i} + \theta_{i+1,i+1} - z_{i,i} - z_{i+1,i+1} \quad \forall 0 \leq i < i + 1 \leq n$$

and thus, by (30),

$$0 \leq y_{i-1,\ell+1} = \sum_{i \leq j \leq k \leq \ell} \theta_{j,k} - \sum_{k=i}^{\ell} z_{k,k}$$

(34)
for all $0 \leq i \leq \ell \leq n$. Substituting the $y$ variables in (31)–(33) yields the following equivalent formulation of the dual program (27):

$$\begin{align*}
\min & \quad z_{-\infty,\infty} + \sum_{i=1}^{n-1} z_{-\infty,i} + z_{i,\infty} \\
\text{s.t.} & \quad \sum_{k=i}^{\ell} z_{k,k} \leq \sum_{i \leq j \leq k \leq \ell} \theta_{j,k} \quad \forall 0 \leq i \leq \ell \leq n \\
& \quad z_{-\infty,i} + z_{i,i} = \theta_{-\infty,i} + \sum_{0 \leq j \leq i} \theta_{j,i} \quad \forall 1 \leq i \leq n - 1 \\
& \quad z_{i,\infty} + z_{i,i} = \theta_{i,\infty} + \sum_{i \leq k \leq n} \theta_{i,k} \quad \forall 1 \leq i \leq n - 1 \\
& \quad z_{-\infty,\infty} - \sum_{k=0}^{n} z_{k,k} = \theta_{-\infty,\infty} - \sum_{0 \leq j \leq k \leq n} \theta_{j,k}.
\end{align*}$$

The variables $z_{-\infty,i}$, $z_{i,\infty}$ and $z_{-\infty,\infty}$ occur only in a single equation each. Furthermore, the matrix corresponding to the inequality constraints satisfies the consecutive-ones property w.r.t. its rows. Therefore, the constraint matrix of the whole system is totally unimodular, which concludes the proof. □

**Remark.** The constraint matrix corresponding to the system (11)–(15) is in general not totally unimodular. A minimal example is the path of length 4.

**Corollary 1.** It holds that $\text{LMC} = P_{\text{path}}$.

**Proof.** This is immediate from Lemma 9, Fact 1 and Theorem 1. □

**Remark.** The path partition problem admits a more efficient representation as a set partitioning problem as follows. For each $0 \leq i \leq \ell \leq n$, let

$$d_{i,\ell} = \sum_{i \leq j \leq k \leq \ell} \theta_{j,k},$$

then taking the dual of problem (35) and simplifying yields the problem

$$\begin{align*}
\min & \quad d^T \lambda \\
\text{s.t.} & \quad \sum_{0 \leq k \leq \ell \leq n} \lambda_{i,\ell} = 1, \quad \forall 0 \leq k \leq n \\
& \quad \lambda \geq 0.
\end{align*}$$

Each variable $\lambda_{i,\ell}$ corresponds to the component containing nodes $i$ to $\ell$. Problem (41) is precisely the sequential set partitioning formulation of the path partition problem as used by [13]. It admits a quadratic number of variables and a linear number of constraints (opposed to a quadratic number of constraints in the description of LMC). The integrality constraint need not be enforced, since the constraint matrix is totally unimodular.
6 Conclusion

We studied the lifted multicut polytope for the special case of trees lifted to the complete graph. We characterized a number of its facets and provided a tighter relaxation compared to the standard linear relaxation. Our analysis establishes a connection between the lifted multicut problem on trees and pseudo-Boolean optimization. Moreover, the described facets yield a complete totally dual integral description of the lifted multicut polytope for paths. This main results relates the geometry of the path partition problem to the combinatorial properties of the sequential set partitioning problem. Moreover, our insights can accelerate solvers for the tree and path partition problem based on integer linear programming.

References

1. Bansal, N., Blum, A., Chawla, S.: Correlation clustering. Mach. Learn. 56(1–3), 89–113 (2004). https://doi.org/10.1023/B:MACH.0000033116.57574.95
2. Beier, T., et al.: Multicut brings automated neurite segmentation closer to human performance. Nat. Methods 14(2), 101–102 (2017). https://doi.org/10.1038/nmeth.4151
3. Boros, E., Gruber, A.: On quadratization of pseudo-Boolean functions. In: International Symposium on Artificial Intelligence and Mathematics, ISAIM 2012, Fort Lauderdale, Florida, USA, 9–11 January 2012 (2012)
4. Boros, E., Hammer, P.L.: Pseudo-Boolean optimization. Discrete Appl. Math. 123(1–3), 155–225 (2002). https://doi.org/10.1016/S0166-218X(01)00341-9
5. Chopra, S., Rao, M.: The partition problem. Math. Program. 59(1–3), 87–115 (1993). https://doi.org/10.1007/BF01581239
6. Demaine, E.D., Emanuel, D., Fiat, A., Immorlica, N.: Correlation clustering in general weighted graphs. Theoret. Comput. Sci. 361(2–3), 172–187 (2006). https://doi.org/10.1016/j.tcs.2006.05.008
7. Deza, M., Grötschel, M., Laurent, M.: Complete descriptions of small multicut polytopes. In: Applied Geometry and Discrete Mathematics, Proceedings of a DIMACS Workshop, pp. 221–252 (1990)
8. Deza, M.M., Laurent, M.: Geometry of Cuts and Metrics. AC, vol. 15. Springer, Heidelberg (1997). https://doi.org/10.1007/978-3-642-04295-9
9. Edmonds, J., Giles, F.R.: A min-max relation for submodular functions on graphs. Ann. Discrete Math. 1, 185–204 (1977). https://doi.org/10.1016/S0167-5060(08)70734-9
10. Grötschel, M., Wakabayashi, Y.: A cutting plane algorithm for a clustering problem. Math. Program. 45(1), 59–96 (1989). https://doi.org/10.1007/BF01589907
11. Hammer, P.L., Hansen, P., Simeone, B.: Roof duality, complementation and persistency in quadratic 0–1 optimization. Math. Program. 28(2), 121–155 (1984). https://doi.org/10.1007/BF02612354
12. Horňáková, A., Lange, J.H., Andres, B.: Analysis and optimization of graph decompositions by lifted multicut. In: ICML (2017)
13. Joseph, A., Bryson, N.: Partitioning of sequentially ordered systems using linear programming. Comput. Oper. Res. 24(7), 679–686 (1997). https://doi.org/10.1016/S0305-0548(96)00070-6
14. Kahl, F., Strandmark, P.: Generalized roof duality. Discrete Appl. Math. 160(16–17), 2419–2434 (2012). https://doi.org/10.1016/j.dam.2012.06.009
15. Kernighan, B.W.: Optimal sequential partitions of graphs. J. ACM 18(1), 34–40 (1971). https://doi.org/10.1145/321623.321627
16. Keuper, M.: Higher-order minimum cost lifted multicuts for motion segmentation. In: ICCV, pp. 4252–4260 (2017). https://doi.org/10.1109/ICCV.2017.455
17. Keuper, M., Levinkov, E., Bonneel, N., Lavoué, G., Brox, T., Andres, B.: Efficient decomposition of image and mesh graphs by lifted multicuts. In: ICCV (2015). https://doi.org/10.1109/ICCV.2015.204
18. Padberg, M.: The Boolean quadric polytope: some characteristics, facets and relatives. Math. Program. 45(1), 139–172 (1989). https://doi.org/10.1007/BF01589101
19. Pia, A.D., Khajavirad, A.: A polyhedral study of binary polynomial programs. Math. Oper. Res. 42(2), 389–410 (2017). https://doi.org/10.1287/moor.2016.0804
20. Rother, C., Kolmogorov, V., Lempitsky, V.S., Szummer, M.: Optimizing binary MRFs via extended roof duality. In: CVPR (2007)
21. Schrijver, A.: Theory of Linear and Integer Programming. John Wiley & Sons Inc., New York (1986)
22. Tang, S., Andriluka, M., Andres, B., Schiele, B.: Multiple people tracking by lifted multicuts and person re-identification. In: CVPR (2017). https://doi.org/10.1109/CVPR.2017.394