Decomposition of Trees and Paths via Correlation

Jan-Hendrik Lange
jlange@mpi-inf.mpg.de

Bjoern Andres
andres@mpi-inf.mpg.de

Max Planck Institute for Informatics
Saarland Informatics Campus · Saarbrücken, Germany

Abstract

We study the problem of decomposing (clustering) a tree with respect to costs attributed to pairs of nodes, so as to minimize the sum of costs for those pairs of nodes that are in the same component (cluster). For the general case and for the special case of the tree being a star, we show that the problem is NP-hard. For the special case of the tree being a path, this problem is known to be polynomial time solvable. We characterize several classes of facets of the combinatorial polytope associated with a formulation of this clustering problem in terms of lifted multicuts. In particular, our results yield a complete totally dual integral (TDI) description of the lifted multicut polytope for paths, which establishes a connection to the combinatorial properties of alternative formulations such as set partitioning.

1 Introduction

We study the problem of decomposing (clustering) a tree with respect to costs attributed to pairs of nodes, so as to minimize the sum of costs for those pairs of nodes that are in the same component (cluster). This *tree partition problem* is stated rigorously in Def. 1. One instance and its solution are depicted in Fig. 1.

On the one hand, the tree partition problem lacks a complexity of clustering problems for general graphs: Its feasible set, the set of all decompositions of a tree, is trivial to characterize. For example, the decompositions of any tree $T = (V, E)$ relate one-to-one to the binary edge labelings $y \in \{0,1\}^E$ that indicate whether neighboring nodes $\{u, v\} \in E$ are in the same component ($y_{uv} = 1$) or distinct components ($y_{uv} = 0$).

On the other hand, the tree partition problem exhibits a complexity not commonly found in clustering problems for general graphs: Its objective function has a multilinear form in the coordinates of y (in the standard unit basis) whose polynomial degree is the length of the longest path in T. This form arises from the fact that any two distinct nodes $\{u, v\} \in \binom{V}{2}$ are in the same component iff $y_e = 1$ for all edges e on the unique path P_{uv} in T from u to v, that is, iff $\prod_{e \in P_{uv}} y_e = 1$. Hence, the tree partition problem can be stated rigorously as follows:

Definition 1. For any (finite, simple, undirected) tree $T = (V, E)$ and any $c \colon \binom{V}{2} \to \mathbb{R}$, the optimization problem (1) is called the instance of the *tree partition problem* w.r.t. T and c. If T is a path, it is also called the instance of the *path partition problem* w.r.t. T and c.

$$\min_{y \in \{0,1\}^E} \sum_{\{u,v\} \in \binom{V}{2}} c_{uv} \prod_{e \in P_{uv}} y_e \qquad (1)$$

1.1 Contribution

In Section 3, we survey several representations of the tree partition problem and discuss its computational complex-
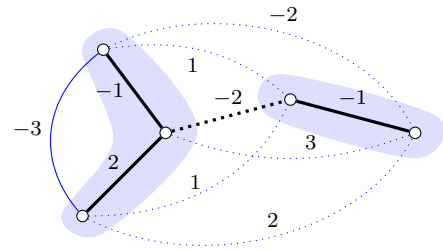

Figure 1 Depicted above is an instance of the tree partition problem with costs associated to pairs of nodes. Edges of the tree are depicted as straight bold lines (dotted or solid). Additional pairs of nodes are marked by thin blue curves (dotted or solid). The solution decomposes the tree into two components, depicted with a shaded background, leaving some pairs of nodes in the same component (solid lines and curves) and some pairs of lines in distinct components (dotted lines and curves).

ity. For the general case and for the special case of the tree being a star, we show that the problem is NP-hard. The path partition problem, on the other hand, is known to be polynomial time solvable [Kernighan, 1971].

In Section 4, we study the lifted multicut polytope [Horňáková et al., 2017] associated with the lifted multicut representation of the tree partition problem. We characterize several classes of facets for the general case, which in particular yield a complete totally dual integral description (TDI) of the lifted multicut polytope for paths. Our results relate the geometry of lifted multicuts to the combinatorial properties of the tree and path partition problem.

1.2 Notation

Throughout the paper, we consider $T = (V, E)$ to be a (finite, simple, undirected) tree with $n = |E|$ edges. Let $m = |\binom{V}{2}|$ be the number of distinct pairs of nodes in V. For any $\{u, v\} \in \binom{V}{2}$, we denote by P_{uv} the unique path in

T connecting u and v. Moreover, we denote by $\mathrm{dist}(u,v)$ the distance of u and v in T, i.e. the length of P_{uv}. To ease notation, we take the liberty of identifying any graph $G = (V_G, E_G)$ with its edge set E_G. Hence, we may write $e \in G$ instead of $e \in E_G$. For any $k \in \mathbb{N}$, we denote by $[k] := \{1, \dots, k\}$ the set of integers from 1 through k.

2 Related Work

2.1 Pseudo-Boolean Optimization

The optimization of pseudo-Boolean functions plays an important role in machine learning, for instance, in MAP-inference for computer vision models. The general problem can be reduced to the quadratic case [Boros and Hammer, 2002, Boros and Gruber, 2012], which is responsible for the NP-hardness of the problem. The combinatorial polytope associated with the linearization of quadratic pseudo-Boolean functions was extensively studied by Padberg [1989]. Computational approaches to pseudo-Boolean optimization based on the roof duality bound [Hammer et al., 1984, Kahl and Strandmark, 2012] have been quite successful in practice [Rother et al., 2007].

2.2 Correlation Clustering and Multicut Polytopes

Decompositions of graphs into an unknown number of components based on (dis-)similarities between *neighboring* nodes is referred to as weighted correlation clustering. It has been studied for the complete graph [Bansal et al., 2004] as well as general graphs [Demaine et al., 2006]. As any decomposition of a graph is characterized by the mathematical notion of a multicut, the study of multicuts is closely related to correlation clustering. The combinatorial polytopes associated with the multicuts of a graph have been studied, among others, most notably by Grötschel and Wakabayashi [1989], Chopra and Rao [1993], Deza and Laurent [1997]. The more general case where correlations between *non-neighboring* nodes are taken into account as well has been introduced by Horňáková et al. [2017].

2.3 Sequential Set Partitioning

Set partitioning problems where the elements are assumed to adhere to a linear order have been studied by Kernighan [1971], who devises a dynamic programming algorithm that essentially solves a shortest path problem in a directed acyclic graph. The corresponding integer linear programming formulation admits a totally unimodular constraint matrix [Joseph and Bryson, 1997].

3 Representations and Complexity

3.1 Sparse Pseudo-Boolean Functions

For any $n \in \mathbb{N}$, any $f \colon \{0,1\}^n \to \mathbb{R}$ is called an n-variate *pseudo-Boolean* function (PBF). Any n-variate PBF f has

a unique multi-linear polynomial form[1]:

$$f(y) = \sum_{I \subseteq [n]} c_I \prod_{i \in I} y_i, \qquad \text{with} \qquad c \colon 2^{[n]} \to \mathbb{R}. \qquad (2)$$

We observe that the objective functions of the tree partition problem form a class of sparse PBFs whose set $\{I \subseteq [n] \colon c_I \neq 0\}$ of non-zero coefficients is constrained by a tree. This class of functions is formalized in the following definition.

Definition 2. An n-variate PBF is called *tree-sparse* w.r.t. a tree $T = (V, [n])$ iff its multi-linear polynomial form (2) is such that, for every $I \subseteq [n]$ with $c_I \neq 0$, the set I induces a path in T. The function is called *path-sparse* w.r.t. a path $P = (V, [n])$ iff it is tree-sparse w.r.t. P.

3.2 Set Partitioning

For any connected subgraph (subtree) $S \subseteq T$, introduce a binary variable $\lambda_S \in \{0,1\}$. By defining the cost $d_S = \sum_{u,v \in V_S} c_{uv}$ for each component S, problem (1) can be reformulated as the set partitioning problem

$$\min \quad d^\top \lambda \qquad (3)$$
$$\text{s.t.} \quad \sum_{S \colon u \in V_S} \lambda_S = 1, \quad \forall u \in V$$
$$\lambda_S \geq 0, \ \lambda_S \in \mathbb{Z}.$$

Here, the costs d_S account for all pairs of nodes within component S and the constraints ensure that every node is contained in exactly one component of the partition. Note that the number of variables λ_S is exponential in the number of leaves of T.

Lemma 1. *The vector $y \in \{0,1\}^E$ is a solution of problem (1) w.r.t. the tree $T = (V, E)$ and costs $c \colon \binom{V}{2} \to \mathbb{R}$ iff the vector λ defined by*

$$\lambda_S = 1 \iff \forall e \in S : y_e = 1$$

for every subtree $S \subseteq T$, is a solution of problem (3) w.r.t. the costs $d_S = \sum_{u,v \in V_S} c_{uv}$.

Proof. We set $c_{uu} = 0$ for all $u \in V$. Let S_u denote the component of the partition that contains $u \in V$. The claim follows from

$$\sum_{\{u,v\} \in \binom{V}{2}} c_{uv} \prod_{e \in P_{uv}} y_e = \sum_{u \in V} \sum_{v \in V_{S_u}} c_{uv} \prod_{e \in P_{uv}} y_e$$
$$= \sum_{u \in V} \sum_{v \in V_{S_u}} c_{uv}$$
$$= \sum_{u \in V} \sum_{\substack{\text{subtree} \\ S \subseteq T \colon u \in V_S}} \lambda_S \sum_{v \in V_S} c_{uv}$$
$$= \sum_{\substack{\text{subtree} \\ S \subseteq T}} \lambda_S \sum_{u,v \in V_S} c_{uv}. \qquad \square$$

---

[1] In line with the literature on pseudo-Boolean optimization, we call the form (2) *multi-linear* despite its constant term $c_\emptyset$.

3.3 Lifted Multicuts

In this section, we identify the tree partition problem (Def. 1) as a special case of the minimum cost lifted multicut problem [Horňáková et al., 2017] where the underlying graph is a tree. Essentially, this representation is a linearization of the tree partition problem with reversed binary encoding.

The *minimum cost lifted multicut problem* w.r.t. the tree $T = (V, E)$ and costs $c \colon \binom{V}{2} \to \mathbb{R}$ is the combinatorial optimization problem

$$\min_{x \in X_T} \sum_{\{u,v\} \in \binom{V}{2}} c_{uv} \, x_{uv}, \qquad (4)$$

where X_T is defined as

$$X_T = \Big\{ x \in \{0,1\}^m \mid$$
$$x_{uv} \leq \sum_{e \in P_{uv}} x_e \quad \forall u, v \in V, \text{dist}(u,v) \geq 2, \qquad (5)$$
$$x_e \leq x_{uv} \qquad \forall u, v \in V, \text{dist}(u,v) \geq 2, \, \forall e \in P_{uv} \Big\}. \qquad (6)$$

Problem (4) is the special case of the minimum cost lifted multicut problem as presented by Horňáková et al. [2017] for the specific choice $G = T$ and $G' = (V, \binom{V}{2})$, i.e. the tree T lifted to the complete graph on V (cf. appendix B). Lemma 2 states that the tree partition problem can be reformulated as a minimum cost lifted multicut problem.

Lemma 2. *The vector $y \in \{0,1\}^n$ is a solution of problem (1) w.r.t. the tree $T = (V, E)$ and costs $c \colon \binom{V}{2} \to \mathbb{R}$ iff the unique $x \in X_T$ such that $x_e = 1 - y_e$ for all $e \in E$ is a solution of problem (4) w.r.t. T and the cost function $-c$.*

Proof. For any distinct pair of nodes $u, v \in V$, we introduce a binary variable $x_{uv} \in \{0,1\}$ via

$$x_{uv} = 1 - \prod_{e \in P_{uv}} y_e \qquad (7)$$

which implies

$$\begin{aligned} x_{uv} = 0 \quad &\Longleftrightarrow \quad \forall e \in P_{uv} \colon \, y_e = 1 \\ &\Longleftrightarrow \quad \forall e \in P_{uv} \colon \, x_e = 0. \end{aligned} \qquad (8)$$

Therefore, we can reformulate problem (1) in terms of the variables x_{uv} by transforming the objective function according to

$$\begin{aligned} c_{uv} \prod_{e \in P_{uv}} y_e &= -c_{uv}\Big(1 - \prod_{e \in P_{uv}} y_e\Big) + c_{uv} \\ &= -c_{uv} \, x_{uv} + c_{uv}. \end{aligned} \qquad (9)$$

This leads to the combinatorial optimization problem

$$\min_{x \in X_T} \sum_{\{u,v\} \in \binom{V}{2}} -c_{uv} \, x_{uv} + c_{uv}, \qquad (10)$$

where $X_T \subseteq \{0,1\}^m$ captures the relationship (8). □

Note that, since the objective function of (4) is linear, we can replace X_T by its convex hull

$$\Xi_T = \text{conv} \, X_T, \qquad (11)$$

which is called the *lifted multicut polytope* w.r.t. T. This motivates the study of the structure of Ξ_T in Section 4. We refer to (5) and (6) as *path* and *cut* inequalities, respectively, and write Θ_T^0 for the naive linear relaxation of Ξ_T, i.e. the set of vectors $x \in [0,1]^m$ that satisfy (5) and (6).

3.4 Complexity

We now discuss the computational complexity of the tree partition problem and the path partition problem. In Lemma 3, we show that the tree partition problem is NP-hard, even if the tree is a star. On the contrary, the path partition problem is polynomial time solvable [Kernighan, 1971].

Lemma 3. *The tree partition problem is NP-hard. It remains NP-hard if T is a star.*

Proof. For a tree-sparse pseudo-Boolean function defined on a star with n leaves, problem (1) is equivalent to the unconstrained binary quadratic program with n variables, which is well-known to be NP-hard. □

4 Polyhedral Geometry

In this section, we establish polyhedral results for the lifted multicut polytope Ξ_T for trees. We characterize all trivial facets and offer a tighter outer relaxation of Ξ_T. In Section 4.5, we show that our results yield a complete totally dual integral (TDI) description of the lifted multicut polytope for paths. This result relates the combinatorial properties of the sequential set partitioning problem to the geometry of the minimum cost lifted multicut problem for paths. The rather technical proofs for our claims are deferred to appendix A.

4.1 Canonical Outer Relaxation

We give another simple relaxation of Ξ_T that is at least as tight as the naive linear relaxation Θ_T^0. To this end, define $\vec{u}(v)$ to be the first node on the path P_{uv} that is different from both u and v (cf. Fig. 2d) and consider the polytope

$$\Theta_T^1 = \Big\{ x \in [0,1]^m \mid$$
$$x_{uv} \leq x_{u,\vec{u}(v)} + x_{\vec{u}(v),v} \quad \forall u, v \in V, \, \text{dist}(u,v) \geq 2, \qquad (12)$$
$$x_{\vec{u}(v),v} \leq x_{uv} \qquad \forall u, v \in V, \, \text{dist}(u,v) \geq 2 \Big\}. \qquad (13)$$

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 Θ_T^1 is indeed an outer relaxation of Ξ_T which is at least as tight as Θ_T^0.

Lemma 4. *It holds that $\Xi_T \subseteq \Theta_T^1 \subseteq \Theta_T^0$.*
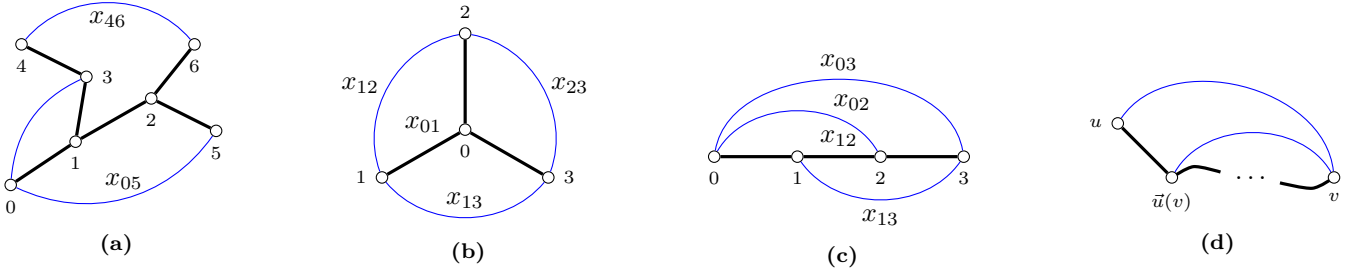
Figure 2 2a: A tree with additional lifted edges (thin blue curves) between non-neighboring nodes corresponding to long range variables. 2b: A star with three leaves lifted to the complete graph. 2c: A path of length 3 lifted to the complete graph. 2d: The node $\vec{u}(v)$ is the first node on the path P_{uv}.

4.2 Trivial Facets

It can be characterized exactly which inequalities of (12), (13) and $0 \leq x_{uv} \leq 1$ define facets of Ξ_T. These elementary results are closely related to the more general study in [Horňáková et al., 2017].

Lemma 5. *The inequality (12) defines a facet of Ξ_T if and only if $\mathrm{dist}(u, v) = 2$.*

Lemma 6. *The inequality (13) defines a facet of Ξ_T if and only if v is a leaf of T.*

Lemma 7. *For any $u, v \in V$, $u \neq v$, the inequality $x_{uv} \leq 1$ defines a facet of Ξ_T if and only if both u and v are leaves of T. Moreover, none of the inequalities $0 \leq x_{uv}$ define facets of Ξ_T.*

4.3 Non-Trivial Facets

We further present a large class of non-trivial facets of Ξ_T. Consider the set of inequalities given by

$$x_{uv} + x_{\vec{u}(v),\vec{v}(u)} \leq x_{u,\vec{v}(u)} + x_{\vec{u}(v),v} \qquad (14)$$
$$\forall u, v \in V, \ \mathrm{dist}(u, v) \geq 3.$$

We shall refer to (14) as *square* inequalities. As an example consider the graph depicted in Fig. 2c and $u = 0$, $v = 3$. Any square inequality is valid and facet-defining for Ξ_T.

Lemma 8. *The inequalities (14) are valid for Ξ_T.*

Lemma 9. *The inequalities (14) define facets of Ξ_T.*

4.4 Facets Related to the Boolean Quadric Polytope

The argument to show Lemma 3 also implies that the lifted multicut polytope for a star with n leaves is isomorphic to the *Boolean Quadric Polytope* [Padberg, 1989] of order n. Therefore, insights about the facial structure of the Boolean quadric polytope can be used to derive valid inequalities and facets for the more general polytope Ξ_T. In fact, any tree-sparse pseudo-Boolean function repeatedly admits substructures similar to the quadratic case. As this connection is beyond the scope of our paper, we only illustrate it with the following example. Consider the star graph depicted in Fig. 2b. It is known from [Padberg, 1989] that the "triangle" inequality

$$x_{01} + x_{23} \leq x_{12} + x_{13} \qquad (15)$$

is valid and facet-defining for Ξ_T. Observe that, similarly, for the tree in Fig. 2a

$$x_{02} + x_{56} \leq x_{05} + x_{06} \qquad (16)$$

holds true, but in this case x_{02}, x_{05} and x_{06} correspond to cubic instead of quadratic terms of the associated pseudo-Boolean function.

4.5 Complete TDI Description for Paths

Suppose the node set $V = \{0, \ldots, n\}$ is linearly ordered and consider the path $P = (V, E)$, where the edge set is given by $E = \big\{\{i, i+1\} \mid i \in \{0, \ldots, n-1\}\big\}$. We show that the facets arising from (12) - (14) yield a complete description of Ξ_P. For this section, we consider only paths of length $n \geq 2$, since for $n = 1$ the polytope $\Xi_P = [0, 1]$ is trivial. Let Θ_P^{Path} be the polytope defined by

$$\Theta_P^{\mathrm{Path}} = \Big\{ x \in \mathbb{R}^m \ \Big|$$
$$x_{0n} \leq 1, \qquad\qquad\qquad\qquad\qquad (17)$$
$$x_{in} \leq x_{i-1,n} \qquad \forall i \in \{1, \ldots, n-1\}, \qquad (18)$$
$$x_{0i} \leq x_{0,i+1} \qquad \forall i \in \{1, \ldots, n-1\}, \qquad (19)$$
$$x_{i-1,i+1} \leq x_{i-1,i} + x_{i,i+1} \ \forall i \in \{1, \ldots, n-1\}, \qquad (20)$$
$$x_{j,k} + x_{j+1,k-1} \leq x_{j+1,k} + x_{j,k-1}$$
$$\forall j, k \in \{0, \ldots, n\}, \ j < k-2 \Big\}. \quad (21)$$

Note that the system (17) - (21) consists precisely of those trivial and square inequalities which we have shown to define facets of Ξ_P. We first prove that Θ_P^{Path} indeed yields an outer relaxation of Ξ_P.

Lemma 10. *It holds that $\Xi_P \subseteq \Theta_P^{\mathrm{Path}} \subseteq \Theta_P^1$.*

As our main result, we prove that Θ_P^{Path} is in fact a complete description of Ξ_P. To derive this, we utilize the notion of total dual integrality. For an extensive reference on this subject we refer the reader to [Schrijver, 1986].

Definition 3. *A system of linear inequalities $Ax \leq b$ with $A \in \mathbb{Q}^{k \times m}$, $b \in \mathbb{Q}^k$ is* 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 serves as a sufficient condition on the integrality of polyhedra according to the following fact.

Fact 1 ([Edmonds and Giles, 1977]). *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* (17) - (21) *is totally dual integral.*

Corollary 1. *It holds that $\Xi_P = \Theta_P^{\mathrm{Path}}$.*

Proof. This is immediate from Lemma 10, Fact 1 and Theorem 1. $\square$

Remark 1. The constraint matrix corresponding to the system (17) - (21) is in general *not* totally unimodular. A minimal example is the path of length 4.

The set partitioning representation of the path partition problem w.r.t. P admits a quadratic number of variables and a linear number of constraints (opposed to a quadratic number of constraints in the description of Ξ_P). This representation corresponds to the dual program of the last problem in the proof of Theorem 1, cf. appendix A. Here, the integrality constraint need not be enforced, since the constraint matrix is totally unimodular.

5 Conclusion

We have characterized all trivial facets of the lifted multicut polytope for trees. Additionally, we provide a tighter relaxation compared to the standard linear relaxation by including additional classes of facets. The described facets provide a complete totally dual integral (TDI) description of the associated lifted multicut polytope for paths. This result relates the geometry of this problem to the combinatorial properties of alternative formulations such as the sequential set partitioning problem.

References

Nikhil Bansal, Avrim Blum, and Shuchi Chawla. Correlation clustering. *Machine Learning*, 56(1–3):89–113, 2004. doi: 10.1023/B:MACH.0000033116.57574.95.

Endre Boros and Aritanan Gruber. On quadratization of pseudo-boolean functions. In *International Symposium on Artificial Intelligence and Mathematics, ISAIM 2012, Fort Lauderdale, Florida, USA, January 9-11, 2012*, 2012.

Endre Boros and Peter L. Hammer. Pseudo-boolean optimization. *Discrete Applied Mathematics*, 123 (1–3):155 – 225, 2002. doi: https://doi.org/10.1016/S0166-218X(01)00341-9.

Sunil Chopra and M.R. Rao. The partition problem. *Mathematical Programming*, 59(1–3):87–115, 1993. doi: 10.1007/BF01581239.

Erik D. Demaine, Dotan Emanuel, Amos Fiat, and Nicole Immorlica. Correlation clustering in general weighted graphs. *Theoretical Computer Science*, 361(2–3):172–187, 2006. doi: 10.1016/j.tcs.2006.05.008.

Michel Marie Deza and Monique Laurent. *Geometry of Cuts and Metrics*. Springer, 1997.

Jack Edmonds and Frederick Richard Giles. A min-max relation for submodular functions on graphs. *Annals of Discrete Mathematics*, 1:185 – 204, 1977. doi: http://dx.doi.org/10.1016/S0167-5060(08)70734-9.

Martin Grötschel and Yoshiko Wakabayashi. A cutting plane algorithm for a clustering problem. *Mathematical Programming*, 45(1):59–96, 1989. doi: 10.1007/BF01589097.

Peter L. Hammer, Pierre Hansen, and Bruno Simeone. Roof duality, complementation and persistency in quadratic 0–1 optimization. *Mathematical Programming*, 28(2):121–155, 1984. doi: 10.1007/BF02612354.

Andrea Horňáková, Jan-Hendrik Lange, and Bjoern Andres. Analysis and optimization of graph decompositions by lifted multicuts. In *ICML*, 2017.

Anito Joseph and Noel Bryson. Partitioning of sequentially ordered systems using linear programming. *Computers & Operations Research*, 24(7):679 – 686, 1997. doi: http://dx.doi.org/10.1016/S0305-0548(96)00070-6.

Fredrik Kahl and Petter Strandmark. Generalized roof duality. *Discrete Applied Mathematics*, 160(16-17):2419–2434, 2012. doi: 10.1016/j.dam.2012.06.009.

Brian W. Kernighan. Optimal sequential partitions of graphs. *J. ACM*, 18(1):34–40, January 1971. ISSN 0004-5411. doi: 10.1145/321623.321627.

Manfred Padberg. The boolean quadric polytope: Some characteristics, facets and relatives. *Mathematical Programming*, 45(1):139–172, 1989. doi: 10.1007/BF01589101.

Carsten Rother, Vladimir Kolmogorov, Victor S. Lempitsky, and Martin Szummer. Optimizing binary mrfs via extended roof duality. In *CVPR*, 2007.

Alexander Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, Inc., New York, NY, USA, 1986. ISBN 0-471-90854-1.

A Proofs for Section 4

Proof of Lemma 4. We show first that $\Xi_T \subseteq \Theta_T^1$. For this purpose, let $x \in \Xi_T$. If x violates (12), 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. $\vec{u}(v), v$ and the path inequality w.r.t. u, v. If x violates (13), 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. $\vec{u}(v), v$. Hence, x satisfies (12)-(13) and therefore $x \in \Theta_T^1$.

Now, we show that $\Theta_T^1 \subseteq \Theta_T^0$. Let $x \in \Theta_T^1$. We need to show that x satisfies all path and cut inequalities. Let $u, v \in V$ with $\mathrm{dist}(u, v) \geq 2$. We proceed by induction on $\mathrm{dist}(u, v)$. If $\mathrm{dist}(u, v) = 2$, then the path inequality is

directly given by (12), while the two cut inequalities are given by (13) for the two possible orderings of u and v. If $\text{dist}(u,v) > 2$, then the path inequality is obtained from (12) and the induction hypothesis for the pair $\vec{u}(v), v$, since $\text{dist}(\vec{u}(v), v) = \text{dist}(u,v) - 1$. 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 (13) such that (w.l.o.g.) e is on the path from $\vec{u}(v)$ to v. Hence, $x \in \Theta_T^0$.

Proof of Lemma 5.   First, suppose $\text{dist}(u,v) = 2$. Then P_{uv} is a path of length 2 and thus chordless in the complete graph on V. Hence, facet-definingness follows directly from [Horňáková et al., 2017, Thm 4]. Now, suppose $\text{dist}(u,v) > 2$ and let $x \in \Xi_T$ be such that (12) is satisfied with equality. We show that this implies

$$x_{uv} + x_{\vec{u}(v),\vec{v}(u)} = x_{u,\vec{v}(u)} + x_{\vec{v}(u),v}. \qquad (22)$$

Then the face of Ξ_T induced by (12) has dimension at most $m-2$ and hence cannot be a facet. In order to check that (22) holds, we distinguish the following three cases. If $x_{uv} = x_{u,\vec{u}(v)} = x_{\vec{v}(u),v}$, then all terms in (22) vanish. If $x_{uv} = x_{u,\vec{u}(v)} = 1$ and $x_{\vec{v}(u),v} = 0$, then $x_{\vec{u}(v),\vec{v}(u)} = 0$ and $x_{u,\vec{v}(u)} = 1$, so (22) holds. Finally, if $x_{uv} = x_{\vec{v}(u),v} = 1$ and $x_{u,\vec{u}(v)} = 0$, then (22) holds because $x_{\vec{u}(v),\vec{v}(u)} = x_{u,\vec{v}(u)}$ by contraction of the edge $u, \vec{u}(v)$.

Proof of Lemma 6.   First, suppose v is not a leaf of T and let $x \in \Xi_T$ be such that (13) is satisfied with equality. Since v is not a leaf, there exists a neighbor $w \in V$ of v such that $P_{\vec{u}(v),v}$ is a subpath of $P_{\vec{u}(v),w}$. We show that x additionally satisfies the equality

$$x_{uw} = x_{\vec{u}(v),w} \qquad (23)$$

and thus the face of Ξ_T induced by (13) 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 $x_{uv} = x_{\vec{u}(v),v} = 0$, then $x_{uw} = x_{vw} = x_{\vec{u}(v),w}$ by contraction of the path P_{uv}, so (23) holds.

Now, suppose v is a leaf of T and let Σ be the face of Ξ_T induced by (13). We need to prove that Σ has dimension $m-1$. This can be done explicitly by showing that we can construct $m-1$ distinct indicator vectors $\chi_{w,w'}$ for $w, w' \in V$ as linear combinations of elements from the set $S = \{x \in \Xi_T \mid x_{uv} = x_{\vec{u}(v),v}\}$. In fact this construction is analogous to the one used in the proof of Theorem 7 in [Horňáková et al., 2017] where the authors derive the dimension of the lifted multicut polytope $\dim(\Xi_{GG'}) = |E'|$. The difference here is that the vector $\mathbb{1} - \chi_{uv} \notin S$, so we have to distinguish between $x \in S$ with $x_{uv} = x_{\vec{u}(v),v} = 0$ and $x \in S$ with $x_{uv} = x_{\vec{u}(v),v} = 1$ in order to show $\chi_{u,\vec{u}(v)} \in \text{lin}(S)$ and then $\chi_e \in \text{lin}(S)$ for all other $e \neq \{u,v\}, \{\vec{u}(v), v\}$.

Proof of Lemma 7.   We apply the more general characterization given in [Horňáková et al., 2017] to our special case. The nodes $u, v \in V$ are a pair of w-w'-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 [Horňáková et al., 2017, Thm 8].

The second assertion follows from [Horňáková et al., 2017, Thm 9] and the fact that we lift to the complete graph on V.

Proof of Lemma 8.   Let $x \in \Xi_T$ and suppose that either $x_{u,\vec{v}(u)} = 0$ or $x_{\vec{u}(v),v} = 0$ for some $u, v \in V$ with $\text{dist}(u,v) \geq 3$. Then, since x satisfies all cut inequalities w.r.t. $u, \vec{v}(u)$, respectively $\vec{u}(v), v$, and the path inequality w.r.t. $\vec{u}(v), \vec{v}(u)$, it must hold that $x_{\vec{u}(v),\vec{v}(u)} = 0$. Moreover, if even $x_{u,\vec{v}(u)} = 0 = x_{\vec{u}(v),v}$, then, by the same reasoning, we have $x_{uv} = 0$ as well. Hence, x satisfies (14).

Proof of Lemma 9.   Let Σ be the face of Ξ_T induced by (14). We need to prove that Σ has dimension $m-1$, which can be done explicitly by showing that we can construct $m-1$ distinct indicator vectors $\chi_{w,w'}$ for $w, w' \in V$ as linear combinations of elements from the set $S = \{x \in \Xi_T \mid x_{uv} + x_{\vec{u}(v),\vec{v}(u)} = x_{u,\vec{v}(u)} + x_{\vec{v}(u),v}\}$. Again, this construction is very technical and analogous to the proof of Theorem 7 in [Horňáková et al., 2017] about the dimension of the lifted multicut polytope.

Proof of Lemma 10.   First, we show that $\Xi_P \subseteq \Theta_P^{\text{Path}}$. Let $x \in \Xi_P$, then x satisfies (17) and (20) by definition. Suppose x violates (18), then $x_{in} = 1$ and $x_{i-1,n} = 0$. This contradicts the fact that x satisfies all cut inequalities w.r.t. $i-1, n$ and the path inequality w.r.t. i, n. So, x must satisfy (18) and, by symmetry, also (19). It follows from Lemma 8 that x satisfies (21) as well and thus $x \in \Theta_P^{\text{Path}}$.

Next, we prove that $\Theta_P^{\text{Path}} \subseteq \Theta_P^1$. To this end, let $x \in \Theta_P^{\text{Path}}$. We show that x satisfies all inequalities (13). 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 (18). Otherwise, we use (21) for $j = u$ and $k = v+1$ and the induction hypothesis on $v+1$:

$$x_{uv} + x_{u+v,v+1} \geq x_{u+1,v} + x_{u,v+1}$$
$$\geq x_{u+1,v} + x_{u+1,v+1}$$
$$\implies x_{uv} \geq x_{u+1,v}.$$

It remains to show that x satisfies all inequalities (12). Let $u, v \in V$ with $u < v-1$. We proceed by induction on $\text{dist}(u,v) = u - v$. If $\text{dist}(u,v) = 2$, then (12) is given by (20). If $\text{dist}(u,v) > 2$, then we use (21) for $j = u$ and $k = v$ as well as the induction hypothesis on $u, v-1$, which have distance $\text{dist}(u,v) - 1$:

$$x_{uv} + x_{u+1,v-1} \leq x_{u+1,v} + x_{u,v-1}$$
$$\leq x_{u+1,v} + x_{u,u+1} + x_{u+1,v-1}$$
$$\implies x_{uv} \leq x_{u,u+1} + x_{u+1,v}.$$

Hence, $x \in \Theta_P^1$, which concludes the proof.

Proof of Theorem 1   We rewrite system (17) - (21) more compactly in the following way. Introduce two artificial nodes $-\infty$ and ∞ where we associate $-\infty$ to any index less than 0 and ∞ 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 $x_{-\infty,\infty}$. Now, the system (17) - (21) is equivalent to the system

$$x_{j,k} + x_{j+1,k-1} \leq x_{j+1,k} + x_{j,k-1}$$
$$\forall j,k \in \{-\infty, 0, \ldots, n, \infty\},\ j \leq k-2 \quad (24)$$

given the additional equality constraints

$$x_{ii} = 0 \quad \forall 0 \leq i \leq n, \quad (25)$$
$$x_{-\infty,i} = 1 \quad \forall 1 \leq i \leq n-1, \quad (26)$$
$$x_{i,\infty} = 1 \quad \forall 1 \leq i \leq n-1, \quad (27)$$
$$x_{-\infty,\infty} = 1. \quad (28)$$

Let the system defined by (24) - (28) be represented in matrix form as $Ax \leq a$, $Bx = b$. Note that Θ_P^{Path} is non-empty and bounded. Thus, to establish TDI, we need to show that for any $c \in \mathbb{Z}^{m+3n}$ the dual program

$$\min\{a^\top y + b^\top z \mid A^\top y + B^\top z = c, y \geq 0\} \quad (29)$$

has an integral optimal solution. Here, the y variables, indexed by j,k, correspond to the inequalities (24) and the z variables, indexed by pairs of $i, -\infty$ and ∞, correspond to the equations (25) - (28). Then, the equation system $A^\top y + B^\top z = c$ translates to

$$y_{i-1,i+1} + z_{i,i} = c_{i,i} \quad (30)$$
$$y_{i-1,i+2} - y_{i-1,i+1} - y_{i,i+2} = c_{i,i+1} \quad (31)$$
$$y_{i-1,\ell+1} - y_{i-1,\ell} - y_{i,\ell+1} + y_{i,\ell} = c_{i,\ell} \quad (32)$$
$$-y_{-\infty,i+1} + y_{-\infty,i} + z_{-\infty,i} = c_{-\infty,i} \quad (33)$$
$$-y_{i-1,\infty} + y_{i,\infty} + z_{i,\infty} = c_{i,\infty} \quad (34)$$
$$y_{-\infty,\infty} + z_{-\infty,\infty} = c_{-\infty,\infty}, \quad (35)$$

where (30) - (32) hold for all $0 \leq i < i+1 < \ell \leq n$ and (33), (34) hold for all $1 \leq i \leq n-1$. Observe that (30) includes only y variables with indices of distance 2, (31) couples y variables of distance 3 with those of distance 2 and (32) 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_{ii} 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

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

and thus, by (32),

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

for all $0 \leq i \leq \ell \leq n$. Substituting the y variables in (33) - (35) yields the following equivalent formulation of (29):

$$\min \quad z_{-\infty,\infty} + \sum_{i=0}^{n} z_{-\infty,i} + z_{i,\infty} \quad (36)$$

$$\text{s.t.} \quad \sum_{k=i}^{\ell} z_{k,k} \leq \sum_{i \leq j \leq k \leq \ell} c_{j,k} \quad \forall 0 \leq i \leq \ell \leq n$$

$$z_{-\infty,i} + z_{i,i} = c_{-\infty,i} + \sum_{0 \leq j \leq i} c_{j,i} \quad \forall 1 \leq i \leq n-1$$

$$z_{i,\infty} + z_{i,i} = c_{i,\infty} + \sum_{i \leq k \leq n} c_{i,k} \quad \forall 1 \leq i \leq n-1$$

$$z_{-\infty,\infty} - \sum_{k=1}^{n} z_{k,k} = c_{-\infty,\infty} - \sum_{0 \leq j \leq k \leq n} c_{j,k}.$$

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. Therefore, the constraint matrix of the whole system is totally unimodular, which concludes the proof.

Path Partition as Sequential Set Partitioning For each $0 \leq i \leq \ell \leq n$, let

$$d_{i,\ell} = \sum_{0 \leq i \leq j \leq k \leq n} c_{j,k},$$

then taking the dual of problem (36) and simplifying yields the problem

$$\min \quad d^\top \lambda \quad (37)$$
$$\text{s.t.} \quad \sum_{0 \leq i \leq k \leq \ell \leq n} \lambda_{i,\ell} = 1, \quad \forall 0 \leq k \leq n$$
$$\lambda \geq 0.$$

Each variable $\lambda_{i,\ell}$ corresponds to the component containing nodes i to ℓ. Problem (37) is precisely the sequential set partitioning formulation of the path partition problem as used by Joseph and Bryson [1997].

B Lifted Multicuts of General Graphs

For any connected graph $G = (V, E)$, any supergraph $G' = (V, E')$ with $E' = E \cup F$ and any $c : E' \to \mathbb{R}$ the instance of the *minimum cost lifted multicut problem* w.r.t. G, G' and c is the binary linear program

$$\min_{x \in X_{GG'}} \sum_{e \in E'} c_e\, x_e \quad (38)$$

with

$$X_{GG'} = \Big\{ x \in \{0,1\}^{E'} \mid$$

$$\forall C \in \text{cycles}(G)\, \forall \hat{e} \in C : x_{\hat{e}} \leq \sum_{e \in C \setminus \{\hat{e}\}} x_e, \quad (39)$$

$$\forall uv \in F\, \forall P \in uv\text{-paths}(G) : x_{uv} \leq \sum_{e \in P} x_e, \quad (40)$$

$$\forall uv \in F\, \forall C \in uv\text{-cuts}(G) : 1 - x_{uv} \leq \sum_{e \in C} 1 - x_e \Big\}.$$
$$(41)$$

The inequalities (39), (40) and (41) are referred to as *cycle*, *path* and *cut* inequalities, respectively. The convex hull

$$\Xi_{GG'} = \text{conv}\, X_{GG'} \qquad (42)$$

of $X_{GG'}$ in $\mathbb{R}^{E'}$ is called the *lifted multicut polytope* w.r.t. G and G'. It was shown by Horňáková et al. [2017] that $\Xi_{GG'}$ has dimension $|E'|$.

For any $x \in X_{GG'}$, the set $M := \{e \in E \mid x_e = 1\}$ of those edges of the graph G that are labeled 1 is a so-called *multicut* of G. Hence, there exists a decomposition of G such that M is precisely the set of those edges that span across distinct components. In addition, the set $M' := \{e \in E' \mid x_e = 1\}$ of those edges $\{u, v\} = e \in E'$ of the graph G' that are labeled 1 is a multicut of G' lifted from the multicut M of G. A multicut M' of G' lifted from G makes explicit for all pairs $\{u, v\} \in E'$ of nodes (not only those neighboring in G) whether u and v are in the same component in the decomposition of G.