Analysis and Optimization of Graph Decompositions by Lifted Multicuts

Andrea Horňáková [*1] **Jan-Hendrik Lange** [*1] **Bjoern Andres** [1]

Abstract

We study the set of all decompositions (clusterings) of a graph through its characterization as a set of lifted multicuts. This leads us to practically relevant insights related to the definition of classes of decompositions by must-join and must-cut constraints and related to the comparison of clusterings by metrics. To find optimal decompositions defined by minimum cost lifted multicuts, we establish some properties of some facets of lifted multicut polytopes, define efficient separation procedures and apply these in a branch-and-cut algorithm.

1. Introduction

This article is about the set of all decompositions (clusterings) of a graph. A decomposition of a graph $G = (V, E)$ is a partition Π of the node set V such that, for every subset $U \in \Pi$ of nodes, the subgraph of G induced by U is connected. An example is depicted in Fig. 1. Decompositions of a graph arise in practice, as feasible solutions of clustering problems, and in theory, as a generalization of partitions of a set, to which they specialize for complete graphs.

We study the set of all decompositions of a graph through its characterization as a set of multicuts. A multicut of G is a subset $M \subseteq E$ of edges such that, for every (chordless) cycle $C \subseteq E$ of G, we have $|M \cap C| \neq 1$. An example is depicted in Fig. 1. For any graph G, a one-to-one relation exists between the decompositions and the multicuts of G. The multicut induced by a decomposition is the set of edges that straddle distinct components. Multicuts are useful in the study of decompositions as the characteristic function $x \in \{0, 1\}^E$ of a multicut $x^{-1}(1)$ of G makes explicit, for every pair $\{v, w\} \in E$ of neighboring nodes, whether v and w are in distinct components. To make explicit also for non-neighboring nodes, specifically, for all $\{v, w\} \in E'$ with $E \subseteq E' \subseteq \binom{V}{2}$, whether v and w are in distinct components, we define a lifting of the multicuts of G to multicuts of $G' = (V, E')$. The multicuts of G' lifted from G are still in one-to-one relation with the decompositions of G. Yet, they are a more expressive model of these decompositions than the multicuts of G. We apply lifted multicuts in three ways:

[*] Equal contribution [1] Max Planck Institute for Informatics, Saarbrücken, Germany. Correspondence to: Bjoern Andres <andres@mpi-inf.mpg.de>.

Proceedings of the 34^{th} International Conference on Machine Learning, Sydney, Australia, PMLR 70, 2017.

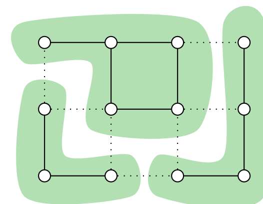

Figure 1. A decomposition of a graph is a partition of the node set into connected subsets. Above, one decomposition is depicted in green. Any decomposition is characterized by the set of those edges (depicted as dotted lines) that straddle distinct components. Such edge sets are precisely the multicuts of the graph.

Firstly, we study problems related to the definition of a class of decompositions by *must-cut* or *must-join* constraints (Section 4). Such constraints have applications where defining a decomposition totally is an ambiguous and tedious task, e.g., in the field of image segmentation. The first problem is to decide whether a set of such constraints is *consistent*, i.e., whether a decomposition of the given graph exists that satisfies the constraints. We show that this decision problem is NP-complete in general and can be solved efficienty for a subclass of constraints. The second problem is to decide whether a consistent set of must-join and must-cut constraints is *maximally specific*, i.e., whether no such constraint can be added without changing the set of decompositions that satisfy the constraints. We show that this decision problem is NP-hard in general and can be solved efficienty for a subclass of constraints. This finding is relevant for comparing the classes of decompositions definable by must-join and must-cut constraints by certain metrics, which is the next topic.

As a second application of lifted multicuts, we study the comparison of decompositions and classes of decompositions by *metrics* (Section 5). To obtain a metric on the set of all decompositions of a given graph, we define a metric on a set of lifted multicuts that characterize these decompositions.

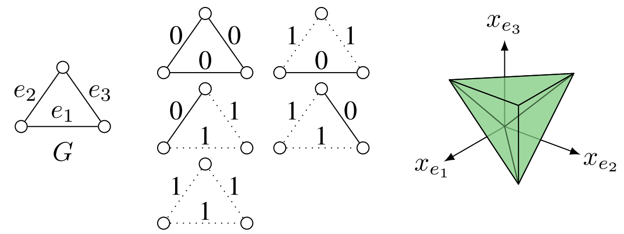

Figure 2. For any connected graph G (left), the characteristic functions of all multicuts of G (middle) span, as their convex hull in $\mathbb{R}^E$, the *multicut polytope* of G (right), a 01-polytope that is $|E|$-dimensional (Chopra & Rao, 1993).

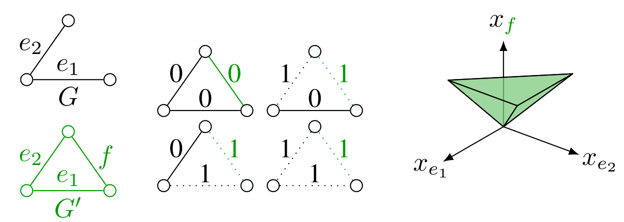

Figure 3. For any connected graph $G = (V, E)$ (top left) and any graph $G' = (V, E')$ with $E \subseteq E'$ (bottom left), those multicuts of G' that are lifted from G (middle) span, as their convex hull in $\mathbb{R}^E$, the *lifted multicut polytope* w.r.t. G and G' (right), a 01-polytope that is $|E'|$-dimensional (Thm. 7).

By lifting to different graphs, we obtain different metrics, two of which are well-known and here generalized. To extend this metric to the classes of decompositions definable by must-join and must-cut constraints, we define a metric on partial lifted multicuts that characterize these classes, connecting results of Sections 4 and 5. We show that computing this metric is NP-hard in general and efficient for a subclass of must-join and must-cut constraints. These findings have implications on the applicability of must-join and must-cut constraints as a form of supervision, more specifically, on the practicality of certain error metrics and loss functions.

As a third application of lifted multicuts, we study the optimization of graph decompositions by minimum cost lifted multicuts. The minimum cost lifted multicut problem is a generalization of the correlation clustering problem. Its applications in the field of computer vision are mentioned below. To tackle this problem, we establish some properties of some facets of lifted multicut polytopes (Fig. 2 and 3), define efficient separation procedures and apply these in a branch-and-cut algorithm.

1.1. Related Work

Initial motivation to study decompositions of a graph by multicuts came from the field of polyhedral optimization. Multicut polytopes are studied by Grötschel & Wakabayashi (1989); Deza et al. (1991; 1992); Chopra & Rao (1993; 1995) and Deza & Laurent (1997) who characterize several classes of their facets.

The binary linear program whose feasible solutions are all multicuts of a graph is known as the correlation clustering problem from the work of Bansal et al. (2004) and Demaine et al. (2006) who establish its APX-hardness and a logarithmic approximation. The stability of its solutions is analyzed by Nowozin & Jegelka (2009). Generalizations to multilinear objective functions are studied by Kim et al. (2014) and Kappes et al. (2016). The problem remains NP-hard for planar graphs (Voice et al., 2012; Bachrach et al., 2013) where it admits a PTAS (Klein et al., 2015) and relaxations that are often tight in practice (Yarkony et al., 2012).

The lifting of multicuts we define makes path connectedness explicit. For a single component, this is studied by Nowozin & Lampert (2010) who introduce the connected subgraph polytope and outer relaxations. Applications of the minimum cost lifted multicut problem and experimental comparisons to the correlation clustering problem in the field of computer vision are by Keuper et al. (2015) and Tang et al. (2017) who find feasible solutions by local search (Keuper et al., 2015; Levinkov et al., 2017), and by Beier et al. (2017) who find feasible solutions by consensus optimization (Beier et al., 2016). The complexity of several decision problems related to clustering with must-join and must-cut constraints is established by Davidson & Ravi (2007). Well-known metrics on the set of all decompositions of a graph are the metric of Rand (1971) and the variation of information (Meilă, 2007).

2. Multicuts

Definition 1 Let $G = (V, E)$ be any graph. A subgraph $G' = (V', E')$ of G is called a *component* of G iff G' is non-empty, node-induced[1], and connected[2]. A partition Π of V is called a *decomposition* of G iff, for every $U \in \Pi$, the subgraph $(U, E \cap \binom{U}{2})$ of G induced by U is connected (and hence a component of G).

For any graph G, we denote by $D_G \subset 2^{2^V}$ the set of all decompositions of G. Useful in the study of decompositions are the multicuts of a graph:

Definition 2 For any graph $G = (V, E)$, a subset $M \subseteq E$ of edges is called a *multicut* of G iff, for every cycle $C \subseteq E$ of G, we have $|C \cap M| \neq 1$.

Lemma 1 *(Chopra & Rao, 1993) It is sufficient in Def. 2 to consider only the* chordless *cycles.*

For any graph G, we denote by $M_G \subseteq 2^E$ the set of all multicuts of G. One reason why multicuts are useful in

[1] That is: $E' = E \cap \binom{V'}{2}$

[2] We do not require a component to be maximal w.r.t. the subgraph relation.

the study of decompositions is that, for every graph G, a one-to-one relation exists between the decompositions and the multicuts of G. An example is depicted in Fig. 1:

Lemma 2 *For any graph $G = (V, E)$, the map $\phi_G : D_G \to 2^E$ defined by (1) is a bijection from D_G to M_G.*

$$\forall \Pi \in D_G \; \forall \{v, w\} \in E : \\ \{v, w\} \in \phi_G(\Pi) \; \Leftrightarrow \; \forall U \in \Pi (v \notin U \vee w \notin U) \quad (1)$$

Another reason why multicuts are useful in the study of decompositions is that, for any graph $G = (V, E)$ and any decomposition Π of G, the characteristic function of the multicut induced by Π is a 01-encoding of Π of fixed length $|E|$.

Lemma 3 *(Chopra & Rao, 1993) For any graph $G = (V, E)$ and any $x \in \{0, 1\}^E$, the set $x^{-1}(1)$ of those edges that are labeled 1 is a multicut of G iff (2) holds. It is sufficient in (2) to consider only* chordless *cycles.*

$$\forall C \in \text{cycles}(G) \; \forall e \in C : \; x_e \leq \sum_{e' \in C \setminus \{e\}} x_{e'} \quad (2)$$

For any graph $G = (V, E)$, we denote by X_G the set of all $x \in \{0, 1\}^E$ that satisfy (2).

2.1. Complete Graphs

The decompositions of a complete graph $K_V := (V, \binom{V}{2})$ are precisely the partitions of the node set V (by Def. 1). The multicuts of a complete graph K_V relate one-to-one to the equivalence relations on V:

Lemma 4 *For any set V and the complete graph K_V, the map $\psi : M_{K_V} \to 2^{V \times V}$ defined by (3) is a bijection between M_{K_V} and the set of all equivalence relations on V.*

$$\forall M \in M_{K_V} \; \forall v, w \in V : \\ (v, w) \in \psi(M) \; \Leftrightarrow \; \{v, w\} \notin M \quad (3)$$

The bijection between the decompositions of a graph and the multicuts of a graph (Lemma 2) specializes, for complete graphs, to the well-known bijection between the partitions of a set and the equivalence relations on the set (by Lemma 4). In this sense, decompositions and multicuts of graphs generalize partitions of sets and equivalence relations on sets.

3. Lifted Multicuts

For any graph $G = (V, E)$, the characteristic function $x \in X_G$ of a multicut $x^{-1}(1)$ of G makes explicit, for every pair $\{v, w\} \in E$ of neighboring nodes, whether v and w are in distinct components. To make explicit also for non-neighboring nodes, specifically, for all $\{v, w\} \in E'$ with $E \subseteq E' \subseteq \binom{V}{2}$, whether v and w are in distinct components, we define a lifting of the multicuts of G to multicuts of $G' = (V, E')$:

Definition 3 For any graphs $G = (V, E)$ and $G' = (V, E')$ with $E \subseteq E'$, the composed map $\lambda_{GG'} := \phi_{G'} \circ \phi_G^{-1}$ is called the *lifting* of multicuts from G to G'.

For any graphs $G = (V, E)$ and $G' = (V, E')$ with $E \subseteq E'$, we introduce the notation $F_{GG'} := E' \setminus E$, for brevity.

Lemma 5 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$ and any $x \in \{0, 1\}^{E'}$, the set $x^{-1}(1)$ is a multicut of G' lifted from G iff*

$$\forall C \in \text{cycles}(G) \; \forall e \in C : \; x_e \leq \sum_{e' \in C \setminus \{e\}} x_{e'} \quad (4)$$

$$\forall vw \in F_{GG'} \; \forall P \in vw\text{-paths}(G) : \; x_{vw} \leq \sum_{e \in P} x_e \quad (5)$$

$$\forall vw \in F_{GG'} \; \forall C \in vw\text{-cuts}(G) : \; 1 - x_{vw} \leq \sum_{e \in C} (1 - x_e) \quad (6)$$

For any graphs $G = (V, E)$ and $G' = (V, E')$ with $E \subseteq E'$ we denote by $X_{GG'}$ the set of all $x \in \{0, 1\}^{E'}$ that satisfy (4)–(6).

4. Partial Lifted Multicuts

As a first application of lifted multicuts, we study the class of decompositions of a graph definable by must-join and must-cut constraints. For this, we consider partial functions. For any set E, a partial characteristic function of subsets of E is a function from any subset $F \subseteq E$ to $\{0, 1\}$. With some abuse of notation, we denote the set of all partial characteristic functions of subsets of E by $\{0, 1, *\}^E := \bigcup_{F \subseteq E} \{0, 1\}^F$. For any $x \in \{0, 1, *\}^E$, we denote the domain of x by $\operatorname{dom} x := x^{-1}(\{0, 1\})$.

For any connected graph $G = (V, E)$ whose decompositions we care about and any graph $G' = (V, E')$ with $E \subseteq E'$, we consider a partial function $\tilde{x} \in \{0, 1, *\}^{E'}$. For any $\{v, w\} \in \operatorname{dom} \tilde{x}$, we constrain the nodes v and w to the same component if $\tilde{x}_{vw} = 0$ and to distinct components if $\tilde{x}_{vw} = 1$.

4.1. Consistency

A natural question to ask is whether a decomposition of the graph G exists that satisfies these constraints. We show that this decision problem is NP-complete.

Definition 4 For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$, and any $\tilde{x} \in \{0, 1, *\}^{E'}$, the elements of

$$X_{GG'}[\tilde{x}] := \{x \in X_{GG'} \,|\, \forall e \in \operatorname{dom} \tilde{x} : \; x_e = \tilde{x}_e\} \quad (7)$$

are called the *completions* of $\tilde{x}$ in $X_{GG'}$. In addition, $\tilde{x}$ is called *consistent* and a *partial characterization of multicuts* *of G' lifted from G* iff

$$X_{GG'}[\tilde{x}] \neq \emptyset \ . \tag{8}$$

We denote the set of all partial characterizations of multicuts of G' lifted from G by

$$\tilde{X}_{GG'} := \left\{ \tilde{x} \in \{0,1,*\}^{E'} \,\middle|\, X_{GG'}[\tilde{x}] \neq \emptyset \right\} \ . \tag{9}$$

Theorem 1 *Deciding consistency is* NP*-complete.*

Lemma 6 *Consistency can be decided efficiently if $E \subseteq \operatorname{dom} \tilde{x}$ or*

$$\begin{aligned} &\forall vw \in \operatorname{dom} \tilde{x} \setminus E : \\ &\quad \tilde{x}_{vw} = 1 \vee \exists P \in vw\text{-path}(G)\, \forall e \in P : \ \tilde{x}_e = 0 \end{aligned} \tag{10}$$

4.2. Specificity

A less obvious question to ask for any partial characterization $\tilde{x}$ of multicuts of G' lifted from G is whether $\tilde{x}$ is maximally specific for its completions in $X_{GG'}$. In other words, is there no edge $e \in E' \setminus \operatorname{dom} \tilde{x}$ such that, for any completions x, x' of $\tilde{x}$ in $X_{GG'}$, we have $x_e = x'_e$, i.e., an edge that could be included in $\operatorname{dom} \tilde{x}$ without changing the set of completions of $\tilde{x}$ in $X_{GG'}$? We show that deciding maximal specificity is NP-hard.

Definition 5 Let $G = (V, E)$ a connected graph and $G' = (V, E')$ a graph with $E \subseteq E'$. For any $\tilde{x} \in \tilde{X}_{GG'}$, the edges

$$E'[\tilde{x}] := \{ e \in E' \,|\, \forall x, x' \in X_{GG'}[\tilde{x}] : \ x_e = x'_e \} \tag{11}$$

are called *decided*. The edges $E' \setminus E'[\tilde{x}]$ are called *undecided*. Moreover, $\tilde{x}$ is called *maximally specific* iff[3]

$$E'[\tilde{x}] \subseteq \operatorname{dom} \tilde{x} \ . \tag{12}$$

Theorem 2 *Deciding maximal specificity is* NP*-hard.*

Lemma 7 *Maximal specificity can be decided efficiently if $E' = E$ or $E \subseteq \operatorname{dom} \tilde{x}$.*

Below, we justify the term *maximal specificity* and define an operation that maps any partial characterization of lifted multicuts to one that is maximally specific.

Definition 6 For any connected graph $G = (V, E)$ and any graph $G' = (V, E')$ with $E \subseteq E'$, the relation $\leq$ on $\tilde{X}_{GG'}$ defined by (13) is called the *specificity* of partial characterizations of multicuts of G' lifted from G.

$$\begin{aligned} &\forall \tilde{x}, \tilde{x}' \in \tilde{X}_{GG'} : \\ &\quad \tilde{x} \leq \tilde{x}' \ \Leftrightarrow \ \operatorname{dom} \tilde{x} \subseteq \operatorname{dom} \tilde{x}' \wedge \forall e \in \operatorname{dom} \tilde{x} : \tilde{x}_e = \tilde{x}'_e \end{aligned} \tag{13}$$

[3]Note that (12) is equivalent to $E'[\tilde{x}] = \operatorname{dom} \tilde{x}$, as $E'[\tilde{x}] \supseteq \operatorname{dom} \tilde{x}$ holds by definition of $E'[\tilde{x}]$.

Lemma 8 *For any connected graph $G = (V, E)$ and any graph $G' = (V, E')$ with $E \subseteq E'$, specificity is a partial order on $\tilde{X}_{GG'}$.*

Note that two partial characterizations $\tilde{x}, \tilde{x}' \in \tilde{X}_{GG'}$ with the same completions $X_{GG'}[\tilde{x}] = X_{GG'}[\tilde{x}']$ need not be comparable w.r.t. $\leq$. For example, consider the graphs G, G' from Fig. 3, consider $\tilde{x} : e_1 \mapsto 0, e_2 \mapsto 0$ and $\tilde{x}' : f \mapsto 0$. Nevertheless, we have the following lemma.

Lemma 9 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$, any $\tilde{x} \in \tilde{X}_{GG'}$ and*

$$\tilde{X}_{GG'}[\tilde{x}] := \left\{ \tilde{x}' \in \tilde{X}_{GG'} \,\middle|\, X_{GG'}[\tilde{x}'] = X_{GG'}[\tilde{x}] \right\} \tag{14}$$

a maximum of $\tilde{X}_{GG'}[\tilde{x}]$ w.r.t. $\leq$ exists and is unique. Moreover, $\tilde{x}$ is maximally specific in the sense of Def. 5 iff $\tilde{x}$ is maximal w.r.t. $\leq$ in $\tilde{X}_{GG'}[\tilde{x}]$.

Definition 7 Let $G = (V, E)$ be a connected graph and let $G' = (V, E')$ be a graph with $E \subseteq E'$. For any $\tilde{x} \in \tilde{X}_{GG'}$, we call the unique maximum of $\tilde{X}_{GG'}[\tilde{x}]$ w.r.t. $\leq$ the *closure* of $\tilde{x}$ w.r.t. G and G' and denote it by $\operatorname{cl}_{GG'} \tilde{x}$.

We denote by $\hat{X}_{GG'}$ the set of all maximally specific partial characterizations of multicuts of G' lifted from G, i.e.:

$$\hat{X}_{GG'} := \left\{ \tilde{x} \in \tilde{X}_{GG'} \,\middle|\, \tilde{x} = \operatorname{cl}_{GG'} \tilde{x} \right\} \ . \tag{15}$$

Theorem 3 *For any $\tilde{x}, \tilde{x}' \in \tilde{X}_{GG'}$, we have $X_{GG'}[\tilde{x}] = X_{GG'}[\tilde{x}'] \Leftrightarrow \tilde{X}_{GG'}[\tilde{x}] = \tilde{X}_{GG'}[\tilde{x}'] \Leftrightarrow \operatorname{cl}_{GG'} \tilde{x} = \operatorname{cl}_{GG'} \tilde{x}'$.*

Lemma 10 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$ and any $x \in X_G$, the closure $y := \operatorname{cl}_{GG'} x$ of x w.r.t. G and G' coincides with the lifting of the multicut $x^{-1}(1)$ of G to the multicut $y^{-1}(1)$ of G', i.e.*

$$(\operatorname{cl}_{GG'} x)^{-1}(1) = \lambda_{GG'}(x^{-1}(1)) \ . \tag{16}$$

Theorem 4 *Computing closures is* NP*-hard.*

Lemma 11 *In the special case that $E' = E$ or $E \subseteq \operatorname{dom} \tilde{x}$, the closure can be computed efficiently.*

5. Metrics

5.1. Metrics on Decompositions

As a second application of lifted multicuts, we compare decompositions of a given graph by comparing lifted multicuts that characterize these decompositions. We compare these lifted multicuts by comparing their characteristic functions by Hamming metrics: For any $E \neq \emptyset$ and any $e \in E$, we define $d^1_e, d^1_E : \{0,1\}^E \times \{0,1\}^E \to \mathbb{N}^+_0$ by the forms

$$d^1_e(x, x') = \begin{cases} 0 & \text{if } x_e = x'_e \\ 1 & \text{otherwise} \end{cases} \tag{17}$$

$$d^1_E(x, x') = \sum_{e' \in E} d^1_{e'}(x, x') \ . \tag{18}$$

Theorem 5 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$, any $\mu : E' \to \mathbb{R}^+$, the set $E'' := E \cup E'$ and the graph $G'' := (V, E'')$, the function $d^\mu_{E'} : X_{GG''} \times X_{GG''} \to \mathbb{R}^+_0$ of the form (19) is a pseudo-metric on $X_{GG''}$. Iff G' is a supergraph of G, i.e., iff $E \subseteq E'$, $d^\mu_{E'}$ is a metric on $X_{GG''}$.*

$$d^\mu_{E'}(x, x') := \sum_{e \in E'} \mu_e \, d^1_e(x, x') \tag{19}$$

By the one-to-one relation between decompositions and multicuts (Lemma 2), $d^\mu_{E'}$ induces a pseudo-metric on the set D_G of all decompositions of G. Two special cases are well-known: For $E' = E$ and $\mu = 1$, we have $d^\mu_{E'} = d^1_E$, which is the Hamming metric (18) on the multicuts that characterize the decompositions, also known as the boundary metric on decompositions. For $E' = \binom{V}{2}$ and $\mu = 1$, $d^1_{E'}$ specializes to the metric of Rand (1971). Between these extremes, i.e., for $E \subseteq E' \subseteq \binom{V}{2}$, the metric $d^\mu_{E'}$ can be used to analyze more specifically how two decompositions of the same graph differ. We propose an analysis w.r.t. the distance δ_{vw} of nodes v and w in G, i.e., w.r.t. the length of a shortest vw-path in G. For this, we denote by $\delta_G := \max\{\delta_{vw} : vw \in \binom{V}{2}\}$ the diameter of G.

Definition 8 For any connected graph $G = (V, E)$ and any $n \in \mathbb{N}$, let $E[n] := \{vw \in \binom{V}{2} \mid \delta_{vw} = n\}$ the set of pairs of nodes of distance n in G. Moreover, let $\mu^n : E[n] \to \mathbb{Q}^+$ the constant function that maps any $vw \in E[n]$ to $1/|E[n]|$. For any connected graph $G = (V, E)$, we call the sequence

$$\left(d^{\mu^n}_{E[n]}\right)_{n \in \{1,\dots,\delta_G\}} \tag{20}$$

the *spectrum of pseudo-metrics* on decompositions of G. For $E' := \binom{V}{2}$ and $\mu : E' \to \mathbb{Q}^+ : vw \mapsto 1/(\delta_G |E[\delta_{vw}]|)$, we call the metric $d^\mu_{E'}$ the *δ-metric* on decompositions of G.

An example of a spectrum of pseudo-metrics is depicted in Fig. 4. For any two decompositions Π, Π' of a connected graph G and suitable lifted multicuts x, x' characterizing these decompositions, $d^{\mu^n}_{E[n]}(x, x')$ equals the fraction of pairs of nodes at distance n in G that are either cut by Π and joined by Π', or cut by Π' and joined by Π. I.e., the pseudo-metric $d^{\mu^n}_{E[n]}$ compares decompositions of G specifically w.r.t. the distance n in G. The δ-metric compares decompositions w.r.t. all distances, and each distance is weighted equally. This is in contrast to Rand's metric which is also a comparison w.r.t. all distances but each distance is weighted by the number of pairs of nodes that have this distance.

5.2. Metrics on Classes of Decompositions

We compare classes of decompositions definable by must-join and must-cut constraints by comparing partial lifted multicuts that characterize these decompositions. To compare partial lifted multicuts, we compare their partial characteristic functions by an extension of the Hamming metric: For any $E \neq \emptyset$, any $e \in E$ and any $\theta \in \mathbb{R}^+_0$, we define $d^\theta_e, d^\theta_E : \{0, 1, *\}^E \times \{0, 1, *\}^E \to \mathbb{R}^+_0$ such that for all $\tilde{x}, \tilde{x}' \in \{0, 1, *\}^E$:

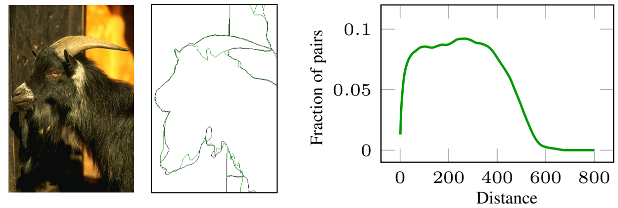

Figure 4. Depicted are two decompositions of the pixel grid graph of an image, all from (Arbeláez et al., 2011), along with the spectrum of pseudo-metrics of these decompositions.

$$d^\theta_e(\tilde{x}, \tilde{x}') = \begin{cases} 1 & \text{if } e \in \operatorname{dom} \tilde{x} \wedge e \in \operatorname{dom} \tilde{x}' \wedge \tilde{x}_e \neq \tilde{x}'_e \\ 0 & \text{if } e \in \operatorname{dom} \tilde{x} \wedge e \in \operatorname{dom} \tilde{x}' \wedge \tilde{x}_e = \tilde{x}'_e \\ 0 & \text{if } e \notin \operatorname{dom} \tilde{x} \wedge e \notin \operatorname{dom} \tilde{x}' \\ \theta & \text{otherwise} \end{cases} \tag{21}$$

$$d^\theta_E(\tilde{x}, \tilde{x}') = \sum_{e' \in E} d^\theta_{e'}(\tilde{x}, \tilde{x}') \ . \tag{22}$$

Theorem 6 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$ and any $\theta \in [\frac{1}{2}, 1]$, the function $\tilde{d}^\theta_{E'} : \tilde{X}_{GG'} \times \tilde{X}_{GG'} \to \mathbb{R}^+_0$ of the form*

$$\tilde{d}^\theta_{E'}(\tilde{x}, \tilde{x}') := d^\theta_{E'}(\operatorname{cl}_{GG'} \tilde{x}, \operatorname{cl}_{GG'} \tilde{x}') \tag{23}$$

is a pseudo-metric on $\tilde{X}_{GG'}$ and a metric on $\hat{X}_{GG'}$. Moreover, for any $\tilde{x}, \tilde{x}' \in \tilde{X}_{GG'}$:

$$\tilde{X}_{GG'}[\tilde{x}] = \tilde{X}_{GG'}[\tilde{x}'] \Leftrightarrow \tilde{d}^\theta_{E'}(\tilde{x}, \tilde{x}') = 0 \ . \tag{24}$$

By the one-to-one relation between decompositions and multicuts (Lemma 2), every partial characterization of a lifted multicut $\tilde{x} \in \tilde{X}_{GG'}$ defines a class of decompositions of the graph G, namely those defined by the lifted multicuts characterized by $X_{GG'}[\tilde{x}]$. By Theorem 6, $\tilde{d}^\theta_{E'}$ with $\theta \in [\frac{1}{2}, 1]$ well-defines a metric on these classes of decompositions and hence a means of comparing the classes of decompositions definable by must-join and must-cut constraints. Computing $\tilde{d}^\theta_{E'}(x, x')$ involves computing the closures of x and x' and is therefore NP-hard (by Theorem 4).

6. Polyhedral Optimization

As a third and final application of lifted multicuts, we turn to the optimization of graph decompositions by lifted multicuts of minimum cost.

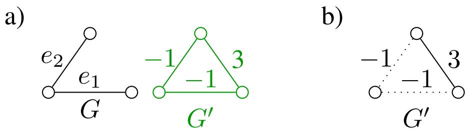

Figure 5. Depicted above in (a) is an instance of the minimum cost lifted multicut problem (Def. 9) w.r.t. graphs G, G' and costs $c = (-1, -1, 3)$. Here, the cost 3 attributed to the additional edge in G' results in the edges e_1 and e_2 not being cut in the optimum $(0, 0, 0)$ which has cost 0. Depicted in (b) is an instance of the minimum cost multicut problem w.r.t. the graph G' and the same cost function. Here, the cost 3 does not prevent the edges e_1 and e_2 from being cut in the optimum $(1, 1, 0)$ which has cost -2.

Definition 9 For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$ 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 optimization problem

$$\min \left\{ \sum_{e \in E'} c_e x_e \;\middle|\; x \in X_{GG'} \right\} \quad . \tag{25}$$

If $E' = E$, (25) specializes to the *minimum cost multicut problem* w.r.t. G' and c that is also known as *graph partition* or *correlation clustering*. If $E' \supset E$, the minimum cost lifted multicut problem w.r.t. G, G' and c differs from the minimum cost multicut problem w.r.t. G' and c. It has a smaller feasible set $X_{GG'} \subset X_{G'}$, as we have shown in Section 3 and depicted for the smallest example in Fig. 2 and 3. Unlike the minimum cost multicut problem w.r.t. G' and c, the minimum cost lifted multicut problem w.r.t. G, G' and c is such that any feasible solution $x \in X_{GG'}$ indicates by $x_{vw} = 0$ that the nodes v and w are connected in G by a path of edges labeled 0. See also Fig. 5. This property can be used to penalize by $c_{vw} > 0$ precisely those decompositions of G for which v and w are in distinct components. For nodes v and w that are not neighbors in G, such costs are sometimes called *non-local attractive*.

To solve instances of the APX-hard minimum cost lifted multicut problem by means of a branch-and-cut algorithm, we study the geometry of lifted multicut polytopes.

Definition 10 (Deza & Laurent, 1997) For any graph $G = (V, E)$, the convex hull $\Xi_G := \mathrm{conv}\, X_G$ of X_G in $\mathbb{R}^E$ is called the *multicut polytope* of G.

Definition 11 For any connected graph $G = (V, E)$ and any graph $G' = (V, E')$ with $E \subseteq E'$, $\Xi_{GG'} := \mathrm{conv}\, X_{GG'}$ is called the *lifted multicut polytope* w.r.t. G and G'.

Examples are shown in Fig. 2 and 3, respectively. In general, the lifted multicut polytope $\Xi_{GG'}$ w.r.t. graphs G and G' (Fig. 3) is a subset of the multicut polytope $\Xi_{G'}$ of the graph G' (Fig. 2). By Lemma 5, the system of cycle inequalities (2) for G' and cut inequalities (6) for G and G' is redundant as a description of $X_{GG'}$ and thus of $\Xi_{GG'}$. Below, we study the geometry of $\Xi_{GG'}$.

6.1. Dimension

Theorem 7 *For any connected graph $G = (V, E)$ and any graph $G' = (V, E')$ with $E \subseteq E'$, $\dim \Xi_{GG'} = |E'|$.*

We prove Theorem 7 by constructing $|E'| + 1$ multicuts of G' lifted from G whose characteristic functions are affine independent points. The strategy is to construct, for any $e \in E'$, an $x \in X_{GG'}$ with $x_e = 0$ and "as many ones as possible". The challenge is that edges cannot be labeled independently. In particular, for $f \in F_{GG'}$, $x_f = 0$ can imply, for certain $f' \in F_{GG'} \setminus \{f\}$, that $x_{f'} = 0$, as illustrated in Fig. 6. This structure is made explicit below, in Def. 12 and 13 and Lemmata 12 and 13.

Definition 12 For any connected graph $G = (V, E)$ and any graph $G' = (V, E')$ such that $E \subseteq E'$, the sequence $(F_n)_{n \in \mathbb{N}}$ of subsets of $F_{GG'}$ defined below is called the *hierarchy* of $F_{GG'}$ with respect to G:

(a) $F_0 = \emptyset$

(b) For any $n \in \mathbb{N}$ and any $\{v, w\} = f \in F_{GG'}$: $\{v, w\} \in F_n$ iff there exists a vw-path in G such that, for any distinct nodes v' and w' in the path such that $\{v', w'\} \neq \{v, w\}$, either $\{v', w'\} \notin F_{GG'}$ or there exists a natural number $j < n$ such that $\{v', w'\} \in F_j$.

Lemma 12 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$ and any $f \in F_{GG'}$, there exists an $n \in \mathbb{N}$ such that $f \in F_n$.*

Definition 13 For any connected graph $G = (V, E)$ and any graph $G' = (V, E')$ with $E \subseteq E'$, the map $\ell : F_{GG'} \to \mathbb{N}$ such that $\forall f \in F_{GG'} \forall n \in \mathbb{N} : \ell(f) = n \Leftrightarrow f \in F_n \wedge f \notin F_{n-1}$ is called the *level function* of $F_{GG'}$.

Lemma 13 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$ and for any $f \in F_{GG'}$, there exists an $x \in X_{GG'}$, called* f-feasible, *such that*

(a) $x_f = 0$

(b) $x_{f'} = 1$ for all $f' \in F_{GG'} \setminus \{f\}$ with $\ell(f') \geq \ell(f)$.

6.2. Facets

We characterize those edges $e \in E'$ for which the inequality $x_e \leq 1$ defines a facet of the lifted multicut polytope $\Xi_{GG'}$.

Theorem 8 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$ and any $e \in E'$, the inequality $x_e \leq 1$ defines a facet of $\Xi_{GG'}$ iff there is no $\{v, w\} = f \in F_{GG'}$ such that e connects a pair of v-w-cut-vertices*[4].

[4]For any graph $G = (V, E)$ and any $v, w \in V$, a *v-w-cut-vertex* is a node $u \in V$ that lies on every vw-path of G.

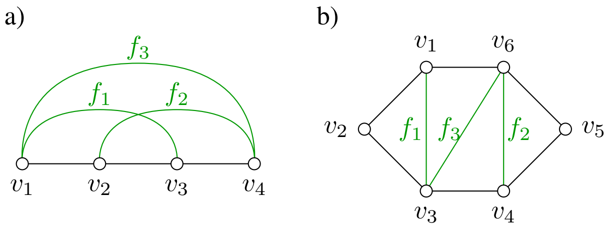

Figure 6. If two nodes $\{v, w\} = f \in F_{GG'}$ are in the same component, as indicated by $x_f = 0$, this can imply $x_{f'} = 0$ for one or more $f' \in F \setminus \{f\}$. In (a) $x_{f_3} = 0$ implies $x_{f_1} = 0$ and $x_{f_2} = 0$. In (b) $x_{f_3} = 0$ implies $x_{f_1} = 0$ or $x_{f_2} = 0$.

Next, we give conditions that contribute to identifying those edges $e \in E'$ for which the inequality $0 \leq x_e$ defines a facet of the lifted multicut polytope $\Xi_{GG'}$.

Theorem 9 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$ and any $e \in E'$, the following assertions hold: In case $e \in E$, the inequality $0 \leq x_e$ defines a facet of $\Xi_{GG'}$ iff there is no triangle in G' containing e. In case $uv = e \in F_{GG'}$, the inequality $0 \leq x_e$ defines a facet of $\Xi_{GG'}$ only if the following necessary conditions hold:*

- *(a) There is no triangle in G' containing e.*
- *(b) The distance of any pair of u-v-cut-vertices except $\{u, v\}$ is at least 3 in G'.*
- *(c) There is no triangle of nodes s, s', t in G' such that $\{s, s'\}$ is a u-v-separating node set and t is a u-v-cut-vertex.*

Next, we characterize those inequalities of (4) and (5) that are facet-defining for $\Xi_{GG'}$. Chopra & Rao (1993) have shown that an inequality of (2) defines a facet of the multicut polytope Ξ_G iff the cycle C is chordless. We establish a similar characterization of those inequalities of (4) and (5) that define a facet of the lifted multicut polytope $\Xi_{GG'}$. For clarity, we introduce some notation: For any cycle C of G and any $e \in C$, let

$$S_{GG'}(e, C) := \left\{ x \in X_{GG'} \,\middle|\, x_e = \sum_{e' \in C \setminus \{e\}} x_{e'} \right\} \tag{26}$$

$$\Sigma_{GG'}(e, C) := \operatorname{conv} S_{GG'}(e, C) \ . \tag{27}$$

For any $vw = f \in F_{GG'}$ and any vw-path P in G, let

$$S_{GG'}(f, P) := \left\{ x \in X_{GG'} \,\middle|\, x_{vw} = \sum_{e \in P} x_e \right\} \tag{28}$$

$$\Sigma_{GG'}(f, P) := \operatorname{conv} S_{GG'}(f, P) \ . \tag{29}$$

Theorem 10 *For any connected graph $G = (V, E)$ and any graph $G' = (V, E')$ with $E \subseteq E'$, the following assertions hold:*

- *(a) For any cycle C in G and any $e \in C$, the polytope $\Sigma_{GG'}(e, C)$ is a facet of $\Xi_{GG'}$ iff C is chordless in G'.*
- *(b) For any edge $vw = f \in F_{GG'}$ and any vw-path P in G, the polytope $\Sigma_{GG'}(f, P)$ is a facet of $\Xi_{GG'}$ iff $P \cup \{f\}$ is chordless in G'.*

Inequalities defined by cycles in G' that contain more than one edge from the set $F_{GG'}$ do not occur in (4) or (5). They are valid for $\Xi_{GG'}$ as they are valid for $\Xi_{G'} \supseteq \Xi_{GG'}$. They define a (non-trivial) facet of $\Xi_{GG'}$ only if the cycle is chordless (as chordal cycles are not even facet-defining for $\Xi_{G'}$). At the same time, chordlessness is not a sufficient condition for facet-definingness of non-trivial cycles. For example, in Fig. 6a, the cycle inequality $x_{f_2} \leq x_{f_3} + x_{v_1 v_2}$ is dominated by the (non-trivial) valid inequality $x_{f_2} \leq x_{f_3}$.

Next, we consider the cut inequalities (6). Examples of cuts that are not facet-defining for $\Xi_{GG'}$ are shown in Fig. 4 in the appendix. To constrain the class of cuts that are facet-defining, we introduce additional notation: For any connected graph $G = (V, E)$, any distinct nodes $v, w \in V$ and any $C \in vw\text{-cuts}(G)$, we denote by

$$G(v, C) = (V(v, C), E(v, C)) \tag{30}$$

$$G(w, C) = (V(w, C), E(w, C)) \tag{31}$$

the largest components of the graph $(V, E \setminus C)$ that contain v and w, respectively. By definition of a vw-cut[5], we have

$$V(v, C) \cap V(w, C) = \emptyset \tag{32}$$

$$\wedge \quad V(v, C) \cup V(w, C) = V \ . \tag{33}$$

We denote by $F_{GG'}(vw, C)$ the set of those edges in $F_{GG'}$, except vw, that cross the vw-cut C of G, i.e.

$$F_{GG'}(vw, C) := \{f \in F_{GG'} \setminus \{vw\} \mid f \not\subseteq V(v, C) \wedge f \not\subseteq V(w, C)\} \ . \tag{34}$$

We denote by $G'(vw, C) := (V, F_{GG'}(vw, C) \cup C)$ the subgraph of G' that comprises all edges from $F_{GG'}(vw, C)$ and C. Finally, we define

$$S_{GG'}(vw, C) := \left\{ x \in X_{GG'} \,\middle|\, 1 - x_{vw} = \sum_{e \in C} (1 - x_e) \right\} \tag{35}$$

$$\Sigma_{GG'}(vw, C) := \operatorname{conv} S_{GG'}(vw, C) \ . \tag{36}$$

Definition 14 For any connected graph $G = (V, E)$, any distinct $v, w \in V$ and any $C \in vw\text{-cuts}(G)$, a component (V^*, E^*) of G is called *properly (vw, C)-connected* iff

$$v \in V^* \ \wedge\ w \in V^* \ \wedge\ |E^* \cap C| = 1 \ . \tag{37}$$

[5]For any graph $G = (V, E)$ and any distinct nodes $v, w \in V$, a vw-cut of G is a minimal (with respect to inclusion) set $C \subseteq E$ such that v and w are not connected in $(V, E \setminus C)$.

It is called *improperly* (vw, C)*-connected* iff

$$V^* \subseteq V(v, C) \ \vee \ V^* \subseteq V(w, C) \ . \tag{38}$$

It is called (vw, C)*-connected* iff it is properly or improperly (vw, C)-connected.

For any (vw, C)-connected component (V^*, E^*) of G, we denote by $F_{V^*} := \{v'w' = f' \in F_{GG'}(vw, C) \mid v' \in V^* \wedge w' \in V^*\}$ the set of those edges $v'w' = f' \in F_{GG'}(vw, C)$ such that (V^*, E^*) is also $(v'w', C)$-connected.

Theorem 11 *For any connected graph $G = (V, E)$, any graph $G' = (V, E')$ with $E \subseteq E'$, any $vw = f \in F_{GG'}$ and any $C \in vw$-cuts(G), $\Sigma_{GG'}(vw, C)$ is a facet of $\Xi_{GG'}$ only if the following necessary conditions hold:*

C1 For any $e \in C$, there exists a (vw, C)-connected component (V^*, E^*) of G such that $e \in E^*$.

C2 For any $\emptyset \neq F \subseteq F_{GG'}(vw, C)$, there exists an edge $e \in C$ and (vw, C)-connected components (V^*, E^*) and (V^{**}, E^{**}) of G such that $e \in E^*$ and $e \in E^{**}$ and $|F \cap F_{V^*}| \neq |F \cap F_{V^{**}}|$.

C3 For any $f' \in F_{GG'}(vw, C)$, any $\emptyset \neq F \subseteq F_{GG'}(vw, C) \setminus \{f'\}$ and any $k \in \mathbb{N}$, there exist (vw, C)-connected components (V^*, E^*) and (V^{**}, E^{**}) with $f' \in F_{V^*}$ and $f' \notin F_{V^{**}}$ such that

$$|F \cap F_{V^*}| \neq k \text{ or } |F \cap F_{V^{**}}| \neq 0 \ . \tag{39}$$

C4 For any $v' \in V(v, C)$, any $w' \in V(w, C)$ and any $v'w'$-path $P = (V_P, E_P)$ in $G'(vw, C)$, there exists a properly (vw, C)-connected component (V^*, E^*) of G such that

$$\begin{aligned} & (v' \notin V^* \vee \exists w'' \in V_P \cap V(w, C) : w'' \notin V^*) \\ \wedge \quad & (w' \notin V^* \vee \exists v'' \in V_P \cap V(v, C) : v'' \notin V^*) \ . \end{aligned} \tag{40}$$

C5 For any cycle $Y = (V_Y, E_Y)$ in $G'(vw, C)$, there exists a properly (vw, C)-connected component (V^*, E^*) of G such that

$$\begin{aligned} & (\exists v' \in V_Y \cap V(v, C) : v' \notin V^*) \\ \wedge \quad & (\exists w' \in V_Y \cap V(w, C) : w' \notin V^*) \ . \end{aligned} \tag{41}$$

6.3. Algorithms

To study the relevance of geometric properties established above, we compare two separation procedures, α and β, for lifted multicut polytopes. We implement these for the branch-and-cut algorithm in the software Gurobi. Our code is available at https://github.com/bjoern-andres/graph. The procedure α is canonical and serves as a reference. It separates infeasible points by any of the inequalities (4)–(6). Violated inequalities of (4) and (5) are found by searching for shortest chordless paths. Violated inequalities of (5) are found by searching for minimum vw-cuts. The procedure β is less canonical: It separates infeasible points by some cycle inequalities w.r.t. G' (cf. Theorem 10) and by cut inequalities (6). Violated cycle inequalities of G' are found by first seaching for paths and cycles as before but then replacing sub-paths by chords in G'. Violated cut-inequalities are found as before but added to the problem only conditionally: For each violated inequality of (6) that we find and the corresponding $\{v, w\} \in F_{GG'}$ and $C \in vw$-cuts(G), we search for a vw-path P in G' such that one of the cycle inequalities for the cycle formed by P and $\{v, w\}$ is violated. If it exists, only the cycle inequality is added. Otherwise, the cut-inequality is added. The advantage of β over α can be seen in Fig. 7 for an instance of the min cost lifted multicut problem by Keuper et al. (2015) with $|V| = 126$, $|E| = 229$ and $|E'| = 1860$.

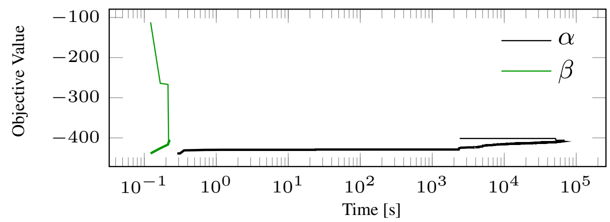

Figure 7. Compared above are the separation procedures α (black) and β (green) in a branch-and-cut search for solutions of an instance of the minimum cost lifted multicut problem by Keuper et al. (2015). Upper and lower bounds are depicted as thin and thick lines, respectively.

7. Conclusion

By studying the set of all decompositions (clusterings) of a graph through its characterization as a set of lifted multicuts, we have gained three insights: 1. Toward the definition of classes of decompositions by must-join and must-cut constraints, we have seen that consistency and maximal specificity are NP-hard to decide. 2. Toward the comparison of decompositions by metrics, we have defined a generalization of Rand's metric and the boundary metric that enables more detailed analyses of how two decompositions of the same graph differ. This metric extends to classes of decompositions definable by must-join and must-cut constraints for which it is NP-hard to compute. 3. Toward the optimization of graph decompositions by minimum cost lifted multicuts, we have established some properties of some facets of lifted multicut polytopes. These properties have led us to efficient separation procedures and a branch-and-cut algorithm for the minimum cost lifted multicut problem.

References

Arbeláez, Pablo, Maire, Michael, Fowlkes, Charless C., and Malik, Jitendra. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916, 2011.

Bachrach, Yoram, Kohli, Pushmeet, Kolmogorov, Vladimir, and Zadimoghaddam, Morteza. Optimal coalition structure generation in cooperative graph games. In *AAAI*, 2013.

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

Beier, Thorsten, Andres, Bjoern, Köthe, Ullrich, and Hamprecht, Fred A. An efficient fusion move algorithm for the minimum cost lifted multicut problem. In *ECCV*, 2016. doi: 10.1007/978-3-319-46475-6_44.

Beier, Thorsten, Pape, Constantin, Rahaman, Nasim, Prange, Timo, Berg, Stuart, Bock, Davi D., Cardona, Albert, Knott, Graham W., Plaza, Stephen M., Scheffer, Louis K., Koethe, Ullrich, Kreshuk, Anna, and Hamprecht, Fred A. Multicut brings automated neurite segmentation closer to human performance. *Nature Methods*, 14(2):101–102, 2017. doi: 10.1038/nmeth.4151.

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

Chopra, Sunil and Rao, M.R. Facets of the k-partition polytope. *Discrete Applied Mathematics*, 61(1):27–48, 1995. doi: 10.1016/0166-218X(93)E0175-X.

Davidson, Ian and Ravi, S. S. Intractability and clustering with constraints. In *ICML*, 2007. doi: 10.1145/1273496.1273522.

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

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

Deza, Michel Marie, Grötschel, Martin, and Laurent, Monique. Complete descriptions of small multicut polytopes. In *Applied Geometry and Discrete Mathematics – The Victor Klee Festschrift, volume 4 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pp. 221–252, 1991.

Deza, Michel Marie, Grötschel, Martin, and Laurent, Monique. Clique-web facets for multicut polytopes. *Mathematics of Operations Research*, 17(4):981–1000, 1992. doi: 10.1287/moor.17.4.981.

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

Kappes, Jörg Hendrik, Speth, Markus, Reinelt, Gerhard, and Schnörr, Christoph. Higher-order segmentation via multicuts. *Computer Vision and Image Understanding*, 143(C):104–119, 2016. doi: 10.1016/j.cviu.2015.11.005.

Keuper, Margret, Levinkov, Evgeny, Bonneel, Nicolas, Lavoué, Guillaume, Brox, Thomas, and Andres, Bjoern. Efficient decomposition of image and mesh graphs by lifted multicuts. In *ICCV*, 2015. doi: 10.1109/ICCV.2015.204.

Kim, Sungwoong, Yoo, Chang D., Nowozin, Sebastian, and Kohli, Pushmeet. Image segmentation using higher-order correlation clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(9):1761–1774, 2014. doi: 10.1109/TPAMI.2014.2303095.

Klein, Philip N., Mathieu, Claire, and Zhou, Hang. Correlation clustering and two-edge-connected augmentation for planar graphs. In *Symposium on Theoretical Aspects of Computer Science*, 2015. doi: 10.4230/LIPIcs.STACS.2015.554.

Levinkov, Evgeny, Uhrig, Jonas, Tang, Siyu, Omran, Mohamed, Insafutdinov, Eldar, Kirillov, Alexander, Rother, Carsten, Brox, Thomas, Schiele, Bernt, and Andres, Bjoern. Joint graph decomposition and node labeling by local search. In *CVPR*, 2017. (Forthcoming).

Meilă, Marina. Comparing clusteringsan information based distance. *Journal of Multivariate Analysis*, 98(5):873–895, 2007. doi: 10.1016/j.jmva.2006.11.013.

Nowozin, Sebastian and Jegelka, Stefanie. Solution stability in linear programming relaxations: Graph partitioning and unsupervised learning. In *ICML*, 2009.

Nowozin, Sebastian and Lampert, Christoph H. Global interactions in random field models: A potential function ensuring connectedness. *SIAM Journal on Imaging Sciences*, 3(4):1048–1074, 2010. doi: 10.1137/090752614.

Rand, William M. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971. doi: 10.1080/01621459.1971.10482356.

Tang, Siyu, Andriluka, Mykhaylo, Andres, Bjoern, and Schiele, Bernt. Multiple people tracking by lifted multicut and person re-identification. In *CVPR*, 2017. (Forthcoming).

Voice, Thomas, Polukarov, Maria, and Jennings, Nicholas R. Coalition structure generation over graphs. *Journal of Artificial Intelligence Research*, 45(1):165–196, 2012.

Yarkony, Julian, Ihler, Alexander, and Fowlkes, Charless C. Fast planar correlation clustering for image segmentation. In *ECCV*, 2012.