Partial Optimality in Cubic Correlation Clustering

David Stein [1] **Silvia Di Gregorio** [1] **Bjoern Andres** [1]

Abstract

The higher-order correlation clustering problem is an expressive model, and recently, local search heuristics have been proposed for several applications. Certifying optimality, however, is NP-hard and practically hampered already by the complexity of the problem statement. Here, we focus on establishing partial optimality conditions for the special case of complete graphs and cubic objective functions. In addition, we define and implement algorithms for testing these conditions and examine their effect numerically, on two datasets.

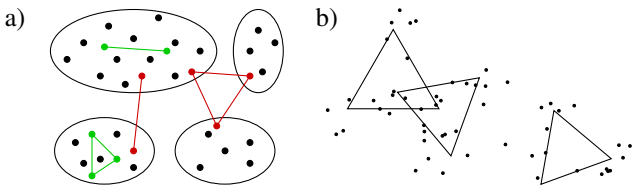

Figure 1. In order to examine the effectiveness of partial optimality conditions numerically, we implement algorithms for testing these conditions and measure the fraction of fixed variables as a function of an amount of noise, for (a) synthetic instances with four clusters and noisy costs, and (b) instances for the task of finding equilateral triangles in a noisy point cloud.

1. Introduction

We study an optimization problem whose feasible solutions are all partitions of a finite set S. Given a cost $c_p \in \mathbb{R}$ for every (unordered) pair $p \in \binom{S}{2}$ and a cost $c_t \in \mathbb{R}$ for every (unordered) triple $t \in \binom{S}{3}$, the objective is to find a partition Π of S so as to minimize the sum of the costs of those pairs and triples whose elements all belong to the same set in Π:

Definition 1.1. The instance of the *cubic set partition* problem with respect to a finite set S, the set P_S of all partitions of S, and a function $c\colon \binom{S}{3} \cup \binom{S}{2} \cup \{\emptyset\} \to \mathbb{R}$ is:

$$\min_{\Pi \in P_S} \quad \sum_{R \in \Pi} \sum_{t \in \binom{R}{3}} c_t + \sum_{R \in \Pi} \sum_{p \in \binom{R}{2}} c_p + c_\emptyset$$

The cubic set partition problem is NP-hard, as it generalizes the NP-hard clique partitioning problem for complete graphs (Grötschel & Wakabayashi, 1989), specializing to the latter in the case that $c_t = 0$ for all $t \in \binom{S}{3}$. Applications of cubic set partitioning include the task of fitting equilateral triangles to points in a plane (cf. Figure 1 and Section 6) and subspace clustering as discussed by Levinkov et al. (2022).

In this article, we ask whether we can compute a partial solution to the problem efficiently, i.e. to decide efficiently for some pairs or triples whether their elements are in the same set or distinct sets of an optimal partition. In order to find such *partial optimality*, we characterize *improving maps* and state efficiently verifiable sufficient conditions of their improvingness, a technique introduced by Shekhovtsov (2013); see also (Shekhovtsov, 2014; Shekhovtsov et al., 2015). In order to examine the effectiveness of these partial optimality conditions numerically, we implement algorithms for testing these. Moreover, we conduct experiments, cf. Figure 1. For conciseness, all proofs are deferred to the appendix.

Related Work. We choose to state the cubic set partition problem (Definition 1.1) in the form of a non-linear binary program (Proposition 2.1), a special case of the higher-order correlation clustering problem introduced by Kim et al. (2014). Combinatorial optimization problems like this involving higher-order objective functions have interesting application as accurate models of intrinsically non-linear tasks (Agarwal et al., 2005; Kappes et al., 2016; Kim et al., 2014; Levinkov et al., 2022; Ochs & Brox, 2012; Purkait et al., 2017). In particular, higher-order correlation clustering has been used for subspace clustering in (Levinkov et al., 2022, Section 5.1) by introducing negative costs for points sufficiently close to a subspace.

Being able to efficiently fix some variables to an optimal value and thus reducing the size of the problem can be valuable in practice. Consequently, much effort has been devoted to studying partial optimality for non-convex problems (Adams et al., 1998; Billionnet & Sutter, 1992; Hammer et al., 1984; Kappes et al., 2013; Kohli et al., 2008; Shekhovtsov, 2014; Shekhovtsov et al., 2015). Particularly

[1]Faculty of Computer Science, TU Dresden, Germany. Correspondence to: Bjoern Andres <bjoern.andres@tu-dresden.de>.

Proceedings of the 40th International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023. Copyright 2023 by the author(s).

impressive is an application of partial optimality conditions to Potts models for image segmentation in which more than 95% of the variables can be fixed (Shekhovtsov et al., 2015, Fig. 1). In contrast to the customary approach of considering a convex, usually linear, relaxation and establishing partial optimality conditions regarding the variables in the extended formulation, we study such conditions directly in the original variable space. Unlike the above-mentioned articles, we concentrate on taking advantage of the specific structure of the cubic correlation clustering problem. To this end, we build on the works of Alush & Goldberger (2012) and Lange et al. (2018; 2019) who establish partial optimality conditions for problems equivalent to correlation clustering with a linear objective function. Regarding their terminology, we remark that the correlation clustering problem, the clique partitioning problem, and the multicut problem are equivalent if the objective functions are linear. The correlation clustering problem keeps attracting considerable attention by the community also in the context of approximation algorithms (Veldt, 2022). The cubic set partition problem we consider here generalizes the specialization to complete graphs of both the correlation clustering problem and the multicut problem. Note that correlation clustering for arbitrary, weighted graphs does not become more specific by considering only complete graphs. Instead, any such problem with respect to an arbitrary graph can be stated as a problem with respect to a complete graph and excessive edges having cost zero.

Here, we transfer all partial optimality conditions established by Alush & Goldberger (2012) and Lange et al. (2018; 2019) for the correlation clustering problem and the multicut problem to the cubic set partition problem. In addition, we establish new results. Unlike in the work of Lange et al. (2018), the algorithm we define does not exploit the sparsity of edges with non-zero cost and, in this sense, is designed for complete graphs. Moreover, we do not contribute persistency conditions for the max cut problem.

2. Preliminaries

In order to establish partial optimality conditions for the cubic set partition problem (Definition 1.1), we state this problem in the form of the non-linear integer program introduced by Kim et al. (2014):

Proposition 2.1. *The instance of the cubic set partition problem with respect to a finite set S and a function $c\colon \binom{S}{3} \cup \binom{S}{2} \cup \{\emptyset\} \to \mathbb{R}$ has the form of the cubic integer program*

$$\min_{x:\binom{S}{2}\to\{0,1\}} \sum_{pqr\in\binom{S}{3}} c_{pqr}x_{pq}x_{pr}x_{qr} + \sum_{pq\in\binom{S}{2}} c_{pq}x_{pq} + c_\emptyset$$

subject to $\forall p \in S\ \forall q \in S \setminus \{p\}\ \forall r \in S \setminus \{p,q\}$:

$$x_{pq} + x_{qr} - x_{pr} \leq 1$$

Below, we let $\mathrm{SP3}_{S,c}$ denote this instance of the problem, ϕ_c its objective function, and X_S its feasible set, i.e. the set of all $x : \binom{S}{2} \to \{0,1\}$ that satisfy the above inequalities.

Our main technique is the construction of improving maps (Shekhovtsov, 2013), which is based on the following preliminary notions.

Definition 2.2. Let $X \neq \emptyset$, $\phi\colon X \to \mathbb{R}$ and $\sigma\colon X \to X$. If for every $x \in X$, we have $\phi(\sigma(x)) \leq \phi(x)$, then σ is called *improving* for the problem $\min_{x\in X} \phi(x)$.

Proposition 2.3. *Let $X \neq \emptyset$, $\phi\colon X \to \mathbb{R}$ and $\sigma\colon X \to X$ an improving map. Moreover, let $Q \subseteq X$. If, for every $x \in X$, $\sigma(x) \in Q$, then there is an optimal solution x^* to $\min_{x\in X} \phi(x)$ such that $x^* \in Q$.*

Corollary 2.4. *Let $S \neq \emptyset$, $X \subseteq \{0,1\}^S$, $\phi\colon X \to \mathbb{R}$ and $\sigma\colon X \to X$ an improving map. Moreover, let $s \in S$ and $\beta \in \{0,1\}$. If for every $x \in X$, $\sigma(x)_s = \beta$, then there is an optimal solution x^* to $\min_{x\in X} \phi(x)$ such that $x^*_s = \beta$.*

Our construction starts from the elementary maps of Lange et al. (2019), i.e. the map $\sigma_{\delta(R)}$ that cuts a set $R \subseteq S$ from its complement, and the map σ_R that joins all sets intersecting with a set $R \subseteq S$:

Definition 2.5. For any finite, non-empty set S and $R \subseteq S$, the *elementary cut map* $\sigma_{\delta(R)}\colon X_S \to X_S$ is such that for all $x \in X_S$ and all $pq \in \binom{S}{2}$:

$$\sigma_{\delta(R)}(x)_{pq} := \begin{cases} 0 & \text{if } |\{p,q\} \cap R| = 1 \\ x_{pq} & \text{otherwise} \end{cases} .$$

Definition 2.6. For any finite, non-empty set S and $R \subseteq S$, the *elementary join map* $\sigma_R\colon X_S \to X_S$ is such that for all $x \in X_S$ and all $pq \in \binom{S}{2}$:

$$\sigma_R(x)_{pq} := \begin{cases} 1 & \text{if } pq \in \binom{R}{2} \\ 1 & \text{if } \forall p' \in \{p,q\} \setminus R\ \exists q' \in R\colon x_{p'q'} = 1 \\ x_{pq} & \text{otherwise} \end{cases} .$$

3. Partial Optimality Conditions

In this section, we establish partial optimality conditions for the cubic set partition problem by constructing improving maps, starting from the elementary maps $\sigma_{\delta(R)}$ and σ_R defined in Section 2.

For simplicity, we introduce some notation: For any $r \in \mathbb{R}$, let $r^\pm := \max\{0, \pm r\}$. For any function $f\colon X \to Y$ and any $X' \subseteq X$, let $f|_{X'}\colon X' \to Y$ denote the restriction of f to X'. From here onwards, S will always denote a finite set. For any non-empty set S and any $R, R', R'' \subseteq S$, let

$$\delta(R, R') := \left\{ pq \in \binom{S}{2} \,\middle|\, p \in R \wedge q \in R' \right\}$$
$$\delta(R) := \delta(R, S \setminus R)$$

$$T_{RR'R''} := \left\{ pqr \in \tbinom{S}{3} \;\middle|\; p \in R \wedge q \in R' \wedge r \in R'' \right\} \ .$$

For $\mathcal{I}_S := \binom{S}{3} \cup \binom{S}{2} \cup \{\emptyset\}$ and any $c \colon \mathcal{I}_S \to \mathbb{R}$, let

$$P^{\pm} := \left\{ pq \in \tbinom{S}{2} \;\middle|\; c_{pq} \gtrless 0 \right\}$$
$$T^{\pm} := \left\{ pqr \in \tbinom{S}{3} \;\middle|\; c_{pqr} \gtrless 0 \right\}$$
$$T_{P'} := \left\{ pqr \in \tbinom{S}{3} \;\middle|\; P' \cap \tbinom{pqr}{2} \neq \emptyset \right\} \quad \forall P' \subseteq \tbinom{S}{2} \ .$$

3.1. Cut Conditions

Here, we establish partial optimality conditions that imply the existence of an optimal solution x^* to $\mathrm{SP3}_{S,c}$ such that $x^*_{ij} = 0$ for some $ij \in \binom{S}{2}$, or such that $x^* \in \{x \in X_S \mid x_{ij}x_{ik}x_{jk} = 0\}$ for some $ijk \in \binom{S}{3}$.

The following Proposition 3.1 generalizes to cubic objective functions the specialization for complete graphs of Theorem 1 of Alush & Goldberger (2012). Intuitively, it says: If there exists a subset R for which joining any pair or triple that has some elements in R and some outside of R leads to a penalty, then we can safely cut the whole set R from the rest.

Proposition 3.1. *Let $S \neq \emptyset$, and let $c \in \mathbb{R}^{\mathcal{I}_S}$. If there exists $R \subseteq S$ such that*

$$c_{pq} \geq 0 \quad \forall pq \in \delta(R) \tag{1}$$
$$c_{pqr} \geq 0 \quad \forall pqr \in T_{\delta(R)} \tag{2}$$

then there is an optimal solution x^* to $\mathrm{SP3}_{S,c}$ such that $x^*_{ij} = 0$ for all $ij \in \delta(R)$.

In the proof, we show: The map that does not change vectors $x \in X_S$ with $x^*_{ij} = 0$ for all $ij \in \delta(R)$ and applies $\sigma_{\delta(R)}$ to the other vectors is improving if (1) and (2) are satisfied.

This condition can be exploited: When satisfied for a set R, $\mathrm{SP3}_{S,c}$ decomposes into two independent subproblems. Firstly,

$$\min_{x \in X_S} \phi_c(x) = \min_{x \in X_R} \phi_{c|_{\mathcal{I}_R}}(x) + \min_{x \in X_{S \setminus R}} \phi_{c|_{\mathcal{I}_{S \setminus R}}}(x) \ .$$

Secondly, given solutions

$$x' \in \operatorname*{argmin}_{x \in X_R} \phi_{c|_{\mathcal{I}_R}}(x) \ , \quad x'' \in \operatorname*{argmin}_{x \in X_{S \setminus R}} \phi_{c|_{\mathcal{I}_{S \setminus R}}}(x) \ ,$$

an optimal solution to the problem $\min_{x \in X_S} \phi_c(x)$ is given by the $x \in X_S$ such that $x_{pq} = x'_{pq}$ if $pq \in \binom{R}{2}$, such that $x_{pq} = x''_{pq}$ if $pq \in \binom{S \setminus R}{2}$, and such that $x_{pq} = 0$ if $pq \in \delta(R)$.

The following Proposition 3.2, together with Proposition 3.4 further below, generalize to cubic objective functions the specialization for complete graphs of Theorem 1 of Lange et al. (2019). The idea behind this statement is this: If there exists a pair ij and a subset R that cuts ij such that the penalty we would have to pay if we were to join i and j is so large that it is at least the best possible reward achieved by joining R and its complement, then it is optimal to keep i and j separated.

Proposition 3.2. *Let $S \neq \emptyset$, and let $c \in \mathbb{R}^{\mathcal{I}_S}$. Moreover, let $ij \in \binom{S}{2}$. If there exists $R \subseteq S$ such that $ij \in \delta(R)$ and*

$$c^+_{ij} \geq \sum_{pqr \in T_{\delta(R)}} c^-_{pqr} + \sum_{pq \in \delta(R)} c^-_{pq} \ , \tag{3}$$

then there is an optimal solution x^* to $\mathrm{SP3}_{S,c}$ such that $x^*_{ij} = 0$.

In the proof, we apply $\sigma_{\delta(R)}$ to any $x \in X_S$ with $x_{ij} = 1$, on the one hand, and apply the identity to the feasible vectors with $x_{ij} = 0$, on the other hand. We prove that this is an improving map under the assumption.

The following Proposition 3.3 establishes a partial optimality result that implies the existence of an optimal solution x^* such that $x^*_{ij}x^*_{ik}x^*_{jk} = 0$ for some $ijk \in \binom{S}{3}$. The intuition is similar as for Proposition 3.2, except here it involves a triple instead of a pair. Note also that one cannot conclude here which of the variables x^*_{ij}, x^*_{ik} or x^*_{jk} equals zero.

Proposition 3.3. *Let $S \neq \emptyset$, and let $c \in \mathbb{R}^{\mathcal{I}_S}$. Moreover, let $ijk \in \binom{S}{3}$ and $R \subseteq S$ such that $ij, ik \in \delta(R)$. If*

$$c^+_{ijk} + c^+_{ij} + c^+_{ik} \geq \sum_{pqr \in T_{\delta(R)}} c^-_{pqr} + \sum_{pq \in \delta(R)} c^-_{pq} \ , \tag{4}$$

there is an optimal solution x^* to $\mathrm{SP3}_{S,c}$ such that $x^*_{ij}x^*_{ik}x^*_{jk} = 0$.

The map used in the proof is similar to the previous ones: If we start from a feasible vector such that $x_{ij}x_{ik}x_{jk} = 0$ already, we do not change it. Otherwise, we apply $\sigma_{\delta(R)}$. Proposition 3.3, together with Proposition 3.5, is a novel result that does not canonically extend prior work.

3.2. Join Conditions

Next, we establish partial optimality conditions that imply the existence of an optimal solution x^* to $\mathrm{SP3}_{S,c}$ such that $x^*_{ij} = 1$ for some $ij \in \binom{S}{2}$. This property can be used to simplify a given instance by joining the elements i and j.

Proposition 3.4 transfers to the cubic set partition problem a result of Lange et al. (2019). In Proposition 3.4, Condition (5) is rather restrictive. The idea behind it is this: If there exists a subset R that cuts i and j, and the total potential reward for joining i and j (meaning not only joining the pair ij but also the triples that could end up together once i and j are in the same cluster) is higher than the sum of rewards and penalties incurred by joining R and its complement, then it is optimal to put i and j together.

Proposition 3.4. *Let $S \neq \emptyset$, and let $c \in \mathbb{R}^{\mathcal{I}_S}$. Moreover, let $ij \in \binom{S}{2}$. If there exists an $R \subseteq S$ such that $ij \in \delta(R)$ and*

$$2c^-_{ij} + \sum_{pqr \in T_{\{ij\}}} c^-_{pqr} \geq \sum_{pqr \in T_{\delta(R)}} |c_{pqr}| + \sum_{pq \in \delta(R)} |c_{pq}| \ , \quad (5)$$

there is an optimal solution x^* to SP3$_{S,c}$ *such that $x^*_{ij} = 1$.*

In the proof, we start from a feasible vector $x \in X_S$. If x satisfies our thesis already, we are done. Otherwise, we apply $\sigma_{ij} \circ \sigma_{\delta(R)}$ and show that this map is improving granted that (5) is satisfied.

While Proposition 3.4 above considers the question whether a pair ij of elements should end up together, Proposition 3.5 below considers the same question but for a triple ijk. The statement is a bit more involved but the intuition is in line with that of Proposition 3.4.

Proposition 3.5. *Let $S \neq \emptyset$, and let $c \in \mathbb{R}^{\mathcal{I}_S}$. Moreover, let $ijk \in \binom{S}{3}$ and $R \subseteq S$ such that $ij, ik \in \delta(R)$. If*

$$2c^-_{ijk} + 2c^-_{ij} + 2c^-_{ik} + c^-_{jk} - \sum_{pqr \in \binom{S}{3}} c^+_{pqr} - \sum_{pq \in \binom{S}{2}} c^+_{pq} \quad (6)$$

$$+ \min_{\substack{x \in X_{ijk} \\ x_{ij}x_{ik}x_{jk}=0}} \sum_{pq \in \binom{ijk}{2}} c_{pq}x_{pq} \geq \sum_{pqr \in T_{\delta(R)}} c^-_{pqr} + \sum_{pq \in \delta(R)} c^-_{pq} \ ,$$

there exists an optimal solution x^* to SP3$_{S,c}$ *such that $x^*_{ij}x^*_{ik}x^*_{jk} = 1$.*

The map that we show to be improving in the proof follows the actions described previously. It is constructed as $\sigma_{ijk} \circ \sigma_{\delta(R)}$.

The following Proposition 3.6 extends to the cubic set partition problem Corollary 1 of Lange et al. (2019). It considers triplets $ijk \in \binom{S}{3}$ and states three requirements under which joining the pair ik does not compromise optimality. Firstly, (7) and (8) say there is a subset $R \subseteq S$ such that the total potential reward from joining i, j and k is greater than or equal to the sum of rewards and penalties incurred by joining R and its complement. Secondly, (9) states that the cost of joining the triple ijk is at most the negative of the sum of rewards incurred when joining ijk and its complement. Under these assumptions, we can put i and k together.

Proposition 3.6. *Let $S \neq \emptyset$, and let $c \in \mathbb{R}^{\mathcal{I}_S}$. Moreover, let $ijk \in \binom{S}{3}$ and $R, R' \subseteq S$ such that $ij, ik \in \delta(R)$ and $jk, ik \in \delta(R')$. If all of the following conditions hold, there exists an optimal solution x^* to* SP3$_{S,c}$ *such that $x^*_{ik} = 1$.*

$$c^-_{ijk} + 2c^-_{ij} + 2c^-_{ik} + \sum_{pqr \in T_{\{ij,ik\}}} c^-_{pqr}$$

$$\geq \sum_{pqr \in T_{\delta(R)}} |c_{pqr}| + \sum_{pq \in \delta(R)} |c_{pq}| \quad (7)$$

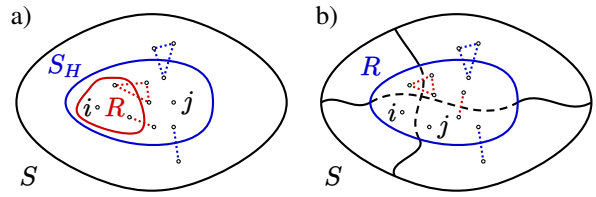

Figure 2. a) Proposition 3.7 compares the total penalty (blue, dotted tuples) of cutting S_H from its complement $S \setminus S_H$to the total cost (red, dotted tuples) of edges and triples lying within S_H and cut by R. b) Proposition 3.10 compares the total cost of separated edges and triples lying within R (red, dotted) to the total cost of edges and triples cut by R (blue, dotted), for any partition (black) of the node set separating any two nodes i and j.

$$c^-_{ijk} + 2c^-_{jk} + 2c^-_{ik} + \sum_{pqr \in T_{\{jk,ik\}}} c^-_{pqr}$$

$$\geq \sum_{pqr \in T_{\delta(R')}} |c_{pqr}| + \sum_{pq \in \delta(R')} |c_{pq}| \quad (8)$$

$$c_{ijk} + c_{ij} + c_{ik} + c_{jk}$$

$$\leq - \sum_{\substack{pqr \in T_{\delta(ijk)} \cap T^- \\ pqr \notin T_{\{ij,ik,jk\}}}} |c_{pqr}| - \sum_{pq \in \delta(ijk) \cap P^-} |c_{pq}| \quad (9)$$

In the proof, we start by fixing $x \in X_S$. Then, we distinguish four cases. Firstly, if $x_{ik} = 1$, we are good. Secondly, if $x_{ik} = x_{ij} = 0$ and $x_{jk} = 1$ and (7) is satisfied, we show $\sigma_{ik} \circ \sigma_{\delta(R)}$ to be improving. Thirdly, if $x_{ik} = x_{jk} = 0$ and $x_{ij} = 1$ and (8) is satisfied, we prove that $\sigma_{ik} \circ \sigma_{\delta(R')}$ is improving. Lastly, if $x_{ij} = x_{jk} = x_{ik} = 0$ and (9) is satisfied, we apply $\sigma_{ijk} \circ \sigma_{\delta(ijk)}$ and show that it is improving.

Next, we discuss a generalization of Theorem 2 of Lange et al. (2019) in the context of instances on complete graphs with a cubic objective function. To this end, let $S_H \subseteq S$. We define $c' \in \mathbb{R}^{\mathcal{I}_{S_H}}$ by the equations written below.

$$c'_\emptyset = \frac{1}{2} \sum_{pqr \in \binom{S_H}{3}} c_{pqr} + \sum_{pq \in \binom{S_H}{2}} c_{pq} \quad (10)$$

$$\forall pq \in \binom{S_H}{2}: \ c'_{pq} = -c_{pq} + \frac{1}{2} \sum_{\substack{r \in S_H \\ r \neq p,q}} c_{pqr} \quad (11)$$

$$\forall pqr \in \binom{S_H}{3}: \ c'_{pqr} = -2c_{pqr} \quad (12)$$

Proposition 3.7 below studies subsets $S_H \subseteq S$ such that $\mathbb{1}_{\binom{S_H}{2}}$ is a trivial solution to the problem $\max_{x \in X_{S_H}} \phi_{c'}(x)$, i.e. $\max_{x \in X_{S_H}} \phi_{c'}(x) = 0$. Here, we compare the negative parts of the costs involved in cutting S_H from its complement with the total cost of any inner cut of S_H. See also Figure 2a.

Proposition 3.7. *Let $S \neq \emptyset$ and $c \in \mathbb{R}^{\mathcal{I}_S}$. Moreover, let $S_H \subseteq S$ and $ij \in \binom{S_H}{2}$. We define $c' \in \mathbb{R}^{\mathcal{I}_{S_H}}$ as in (10),*

(11), (12). *If* $\max_{x \in X_{S_H}} \phi_{c'}(x) = 0$ *and for all* $R \subseteq S_H$ *with* $i \in R$ *and* $j \in S_H \setminus R$ *we have*

$$\sum_{pq \in \delta(S_H) \cap P^-} c_{pq} + \sum_{pqr \in T_{\delta(S_H)} \cap T^-} c_{pqr} \tag{13}$$
$$\geq \sum_{pq \in \delta(R, S_H \setminus R)} c_{pq} + \sum_{pqr \in T_{\delta(R, S_H \setminus R)} \cap \binom{S_H}{3}} c_{pqr} \ ,$$

there is an optimal solution x^* *to* $\mathrm{SP3}_{S,c}$ *such that* $x^*_{ij} = 1$.

In the proof, we apply $\sigma = \sigma_{S_H} \circ \sigma_{\delta(S_H)}$ to any $x \in X_S$ with $x_{ij} = 0$ and show that σ is improving under the assumptions of Proposition 3.7.

We are unaware of an efficient method for finding subsets $S_H \subseteq S$ and $R \subseteq S$ for which (13) are satisfied. For subsets S_H with $|S_H| \in \{2, 3\}$, two corollaries of Proposition 3.7 provide efficiently-verifiable partial optimality conditions:

Corollary 3.8. *Let* $S \neq \emptyset$, $c \in \mathbb{R}^{\mathcal{I}_S}$ *and* $ij \in \binom{S}{2}$. *If*

$$c_{ij} \leq \sum_{pq \in \delta(ij) \cap P^-} c_{pq} + \sum_{pqr \in T_{\delta(ij)} \cap T^-} c_{pqr}$$

there is an optimal solution x^* *to* $\mathrm{SP3}_{S,c}$ *such that* $x^*_{ij} = 1$.

Corollary 3.9. *Let* $S \neq \emptyset$, $c \in \mathbb{R}^{\mathcal{I}_S}$ *and* $ijk \in \binom{S}{3}$. *If*

$$c_{ij} + c_{ik} \leq 0 \ , \quad c_{ij} + c_{jk} \leq 0 \ , \quad c_{ik} + c_{jk} \leq 0$$
$$c_{ij} + c_{ik} + c_{jk} \leq 0 \ , \quad c_{ij} + c_{ik} + c_{jk} + \frac{1}{2} c_{ijk} \leq 0$$
$$c_{ij} + c_{ik} + c_{ijk} \leq \sum_{pq \in \delta(ijk) \cap P^-} c_{pq} + \sum_{pqr \in T_{\delta(ijk)} \cap T^-} c_{pqr}$$
$$c_{jk} + c_{ik} + c_{ijk} \leq \sum_{pq \in \delta(ijk) \cap P^-} c_{pq} + \sum_{pqr \in T_{\delta(ijk)} \cap T^-} c_{pqr}$$

there is an optimal solution x^* *to* $\mathrm{SP3}_{S,c}$ *such that* $x^*_{ik} = 1$.

We now present the last partial optimality condition of this article. It observes that separating a whole subset R from the rest and then joining everything in it does not worsen the objective value if, for every pair $ij \in \binom{R}{2}$ and every partition of S that separates i from j, the sum of costs of separated pairs and triples within R is at most the sum of costs of joined pairs and triples cut by R. See also Figure 2b.

Proposition 3.10. *Let* $S \neq \emptyset$ *and* $c \in \mathbb{R}^{\mathcal{I}_S}$. *Moreover, let* $R \subseteq S$. *If for every* $ij \in \binom{R}{2}$ *we have*

$$\max_{\substack{x \in X_S \\ x_{ij} = 0}} \Big\{ \sum_{pqr \in \binom{R}{3}} c_{pqr}(1 - x_{pq} x_{pr} x_{qr}) + \sum_{pq \in \binom{R}{2}} c_{pq}(1 - x_{pq}) \Big\}$$
$$\leq \min_{\substack{x \in X_S \\ x_{ij} = 0}} \Big\{ \sum_{pqr \in T_{\delta(R)}} c_{pqr} x_{pq} x_{pr} x_{qr} + \sum_{pq \in \delta(R)} c_{pq} x_{pq} \Big\} \tag{14}$$

then there is an optimal solution x^* *to* $\mathrm{SP3}_{S,c}$ *such that* $\forall ij \in \binom{R}{2} \colon x^*_{ij} = 1$.

In the proof, we apply the identity to feasible vectors that satisfy our claim, and apply $\sigma = \sigma_R \circ \sigma_{\delta(R)}$ otherwise.

The above condition has been established independently of prior work on the correlation clustering problem. We are unaware of an efficient method for checking (14) for arbitrary subsets $R \subseteq S$ and costs $c \in \mathbb{R}^{\mathcal{I}_S}$. Yet, Corollary 3.11 below describes one setting in which a suitable subset can be searched for heuristically, in polynomial time. Specifically, the objective function $c \in \mathbb{R}^{\mathcal{I}_S}$ needs to be such that $c_{pq} \leq 0$ for all $pq \in \binom{R}{2}$ and such that $c_{pqr} \leq 0$ for all $pqr \in \binom{R}{3}$. An intuition for this corollary is this: For a moment, let us consider a fixed subset R and all bipartitions of R. Let us recall that the costs of all the triples and pairs inside of R are non-positive. If the worst possible cost of joining these two parts of R back together is still less than or equal to the reward obtained by joining R with the rest, then we can safely start by putting all the objects of R in the same set and decide independently whether or not to join R with other sets.

Corollary 3.11. *Let* $S \neq \emptyset$ *and* $c \in \mathbb{R}^{\mathcal{I}_S}$. *Moreover, let* $R \subseteq S$. *If*

$$c_{pq} \leq 0 \quad \forall pq \in \binom{R}{2} \tag{15}$$
$$c_{pqr} \leq 0 \quad \forall pqr \in \binom{R}{3} \tag{16}$$

and

$$\max_{\substack{R' \subset R \\ R' \neq \emptyset}} \Big\{ \sum_{pqr \in T_{\delta(R')} \cap \binom{R}{3}} c_{pqr} + \sum_{pq \in \delta(R', R \setminus R')} c_{pq} \Big\}$$
$$\leq \sum_{pqr \in T_{\delta(R)} \cap T^-} c_{pqr} + \sum_{pq \in \delta(R) \cap P^-} c_{pq} \ , \tag{17}$$

then there is an optimal solution x^* *to* $\mathrm{SP3}_{S,c}$ *such that* $x^*_{ij} = 1$, $\forall ij \in \binom{R}{2}$.

The previous corollary follows from the following two facts. Let (15) and (16) be satisfied. First of all, for any $ij \in \binom{R}{2}$, we have that the left-hand side of (14) is equal to

$$\max_{\substack{R' \subsetneq R \\ i \in R' \\ j \notin R'}} \sum_{pqr \in T_{\delta(R')} \cap \binom{R}{3}} c_{pqr} + \sum_{pq \in \delta(R', R \setminus R')} c_{pq} \ .$$

I.e., the maximizer is given by a feasible $x \in X_S$ whose restriction to $\binom{R}{2}$ corresponds to a partition $\mathcal{R}$ of R into two subsets. To see this, note for any $ij \in \binom{R}{2}$ that, instead of maximizing the left-hand side of (14) over $x \in X_S$ such that $x_{ij} = 0$, we can equivalently maximize over all $x' \in X_R$ such that $x'_{ij} = 0$. Now, let us assume there exists $ij \in \binom{R}{2}$ such that the maximizer is given by a feasible $x' \in X_R$ corresponding to a partition $\mathcal{R}'$ of R into more than two subsets. Without loss of generality, let $R_1, R_2 \in \mathcal{R}'$ such that $i \in R_1$, $j \in R_2$. Then, the vector $x'' \in X_R$

corresponding to the partition $\mathcal{R}'' = \left\{R_1, \bigcup_{R \in \mathcal{R}' \setminus \{R_1\}} R\right\}$ has objective value at least the objective value of x'. This follows from the facts that all the costs are non-positive and that $\mathcal{R}'$ is a refinement of $\mathcal{R}''$. Secondly, by using the trivial lower bound of the right-hand side of (14), we have that it is at least $\sum_{pqr \in T_{\delta(R)} \cap T^-} c_{pqr} + \sum_{pqr \in \delta(R) \cap P^-} c_{pq}$.

4. Efficient Testing of Partial Optimality

Next, we describe algorithms for testing all the partial optimality conditions introduced above. This includes exact algorithms and heuristics, and we discuss their runtimes. We start by examining Proposition 3.1. Algorithm 1 terminates in $O(|S|^3)$ time and finds a subset $R \subseteq S$ that satisfies (1)–(2). Note that (1)–(2) hold in particular for the trivial subset $R = S$. We formalize the correctness of Algorithm 1 in Proposition 4.1.

Proposition 4.1. *Let $S \neq \emptyset$, $c \in \mathbb{R}^{\mathcal{I}_S}$. Then, Algorithm 1 outputs a partition that contains a subset satisfying (1)–(2). Moreover, if there exists a non-trivial $R \subseteq S$, namely $R \neq \emptyset, S$, such that (1)–(2) are satisfied by R, then Algorithm 1 outputs a non-trivial partition.*

Algorithm 1 Region Growing

Input: $S \neq \emptyset$, $c : \mathcal{I}_S \to \mathbb{R}$
Initialize: $\mathcal{R} = \{\}$, queue $Q = S$
repeat
  $p := Q.pop$
  $R = \{p\}$
  Initialize $noChange = true$.
  repeat
    Set $noChange = true$
    if $\exists pq \in \delta(R) : c_{pq} < 0$ **then**
      $R := R \cup \{p, q\}$
      remove p, q from Q
      set $noChange = false$
    if $\exists pqr \in T_{\delta(R)} : c_{pqr} < 0$ **then**
      $R := R \cup \{p, q, r\}$
      remove p, q, r from Q
      set $noChange = false$
  until $noChange$ is $true$
  Add R to $\mathcal{R}$
until $Q = \emptyset$

Partial optimality according to Propositions 3.2–3.6 is conditional to the existence of a pair $ij \in \binom{S}{2}$ or triple $ijk \in \binom{S}{3}$, together with a subset $R \subseteq S$, and in case of Proposition 3.6 a second subset $R' \subseteq S$ independent of R, such that specific inequalities are satisfied, namely (3)–(9). For every triple, we test (9) explicitly, in quadratic time. For every pair or triple, we reduce the search for subsets R or R' that satisfy (3)–(8) *with maximum margin* to the min st-cut problem (in Section 4.1). In order to test for partial optimality efficiently, we solve the dual max st-flow problems by the implementation in the C++ library Boost (2022) of the push-relabel algorithm of Goldberg & Tarjan (1988).

As mentioned already in Section 3.2, we are unaware of an efficient method for finding subsets that satisfy the conditions of Proposition 3.7 or 3.10. Regarding Proposition 3.7, we resort to the special case of Corollary 3.8 that we test for each pair in quadratic time, and to the special case of Corollary 3.9 that we test for each triple in quadratic time. Regarding Proposition 3.10, we employ the special case of Corollary 3.11 and search heuristically for a witness R of (17), as follows. In an outer loop, we iterate over all pairs $R = \{i, j\}$ with $c_{ij} \leq 0$. For each of these initializations of R, we add elements to R for which the costs of all pairs and triples inside R is non-positive, greedily considering ones for which the costs of newly considered pairs and triples is minimal. Upon termination of this inner loop, we take R to be a candidate. By construction of R, all coefficients on the left-hand side of (17) are non-positive. By applying Proposition 4.2 to the left-hand side of (17), this problem takes the form of a min cut problem with non-negative capacities that we solve exactly using the implementation in the C++ library Boost (2022) of the algorithm by Stoer & Wagner (1997).

4.1. Reductions to Minimum Cut Problems

Here, we discuss how, for a given pair or triple, we reduce the search for subsets $R, R' \subseteq S$ that satisfy (3)–(8) maximally to the min st-cut problem. In any of these cases, we have $S \neq \emptyset$ and $c \in \mathbb{R}^{\mathcal{I}_S}$ such that $c_{pqr} \geq 0$ for all $pqr \in \binom{S}{3}$, and $c_{pq} \geq 0$ for all $pq \in \binom{S}{2}$. Moreover, we have $i \in S$, a pair or triple $\{i\} \cup S_0 \subseteq S$ and a problem of the form

$$\min_{\substack{R \subseteq S: \\ i \in R, \\ j \notin R, \forall j \in S_0}} \sum_{pqr \in T_{\delta(R)}} c_{pqr} + \sum_{pq \in \delta(R)} c_{pq} \ . \tag{18}$$

To begin with, we move costs of triples to costs of pairs:

Proposition 4.2. *Let $S \neq \emptyset$, $c \in \mathbb{R}^{\mathcal{I}_S}$ and $R \subseteq S$. Then*

$$\sum_{pqr \in T_{\delta(R)}} c_{pqr} = \frac{1}{2} \sum_{pq \in \delta(R)} \sum_{r \in S \setminus \{p,q\}} c_{pqr} \ .$$

Consequently, (18) is equivalent to

$$\min_{\substack{R \subseteq S: \\ i \in R, \\ j \notin R, \forall j \in S_0}} \sum_{pq \in \delta(R)} c'_{pq}, \tag{19}$$

where $c'_{pq} = c_{pq} + \frac{1}{2} \sum_{r \in S \setminus \{p,q\}} c_{pqr}$ for all $pq \in \binom{S}{2}$. Next, we reduce (19) to a *quadratic unconstrained binary optimization problem*, by applying the following proposition:

Proposition 4.3. *Let $S \neq \emptyset$ and $c \in \mathbb{R}^{\binom{S}{2}}$. Moreover, let $i \in S$ and $S_0 \subseteq S \setminus \{i\}$. Furthermore, let $S' = S \setminus (S_0 \cup \{i\})$ and $c' \colon \binom{S'}{2} \cup S' \cup \{\emptyset\} \to \mathbb{R}$ such that*

$$\begin{aligned} c'_p &= \sum_{q \in S \setminus \{p\}} c_{pq} - 2c_{pi} && \forall p \in S' \\ c'_{pq} &= -2c_{pq} && \forall pq \in \tbinom{S'}{2} \\ c'_\emptyset &= \sum_{q \in S \setminus \{i\}} c_{qi} \ . \end{aligned}$$

Then:

$$\begin{aligned} &\min_{\substack{R \subseteq S: \\ i \in R, \\ j \notin R, \forall j \in S_0}} \sum_{pq \in \delta(R)} c_{pq} \\ &= \min_{y \in \{0,1\}^{S'}} \sum_{pq \in \binom{S'}{2}} c'_{pq} y_p y_q + \sum_{p \in S'} c'_p y_p + c'_\emptyset \ . \end{aligned} \tag{20}$$

For the instances of (20) that arise from testing (3)–(8), we have $\forall pq \in \binom{S'}{2} \colon c'_{pq} \leq 0$. Thus, the right-hand side of (20) is *submodular* and can be minimized in strongly polynomial time (Boros et al., 2008; Kolmogorov & Zabin, 2004). For completeness, we describe the reduction to an instance of min st-cut in detail in Appendix A.2.

5. Combining Partial Optimality Conditions

Next, we discuss how we apply partial optimality conditions iteratively and why this requires special attention. Let $S \neq \emptyset$ and $c \in \mathbb{R}^{\mathcal{I}_S}$. Furthermore, let $Q_1, Q_2 \subseteq X_S$. If there is an optimal solution $x_1^* \in X_S$ to $\text{SP3}_{S,c}$ such that $x_1^* \in Q_1$, and an optimal solution $x_2^* \in X_S$ to $\text{SP3}_{S,c}$ such that $x_2^* \in Q_2$, then there is not necessarily an optimal solution $x^* \in X_S$ to $\text{SP3}_{S,c}$ such that $x^* \in Q_1 \cap Q_2$. For example, consider $S = \{1,2,3\}$ and $c \in \mathbb{R}^{\mathcal{I}_S}$ such that $c_{123} = 5$, $c_{12} = c_{13} = c_{23} = -2$ and $c_\emptyset = 0$. Then, $\min_{x \in X_S} \phi_c(x) = -2$. Furthermore, let $Q_1 = \{x \in X_S \mid x_{12} = 1\}$ and $Q_2 = \{x \in X_S \mid x_{13} = 1\}$. It follows that $Q_1 \cap Q_2 = \{x \in X_S \mid x_{12} = 1, x_{13} = 1\} = \{(1,1,1)\}$. The set of optimal solutions is the set of all $x \in X_S$ for which there is exactly one $pq \in \binom{S}{2}$ with $x_{pq} = 1$. Thus, the feasible $x' \in X_S$ such that $x'_{12} = x'_{13} = x'_{23} = 1$ is not optimal.

5.1. Cut Conditions

For any $T_0 \subseteq \binom{S}{3}$, $P_0 \subseteq \binom{S}{2}$, we define $X_S|_{T_0,P_0} \subseteq X_S$ such that $x \in X_S|_{T_0,P_0}$ if and only if $\forall pqr \in T_0 \colon x_{pq}x_{pr}x_{qr} = 0$ and $\forall pq \in P_0 \colon x_{pq} = 0$. For any $R \subseteq S$, we have that $\sigma_{\delta(R)}$ restricted to $X_S|_{T_0,P_0}$ has image $X_S|_{T_0,P_0}$. All our cut results use either $\sigma_{\delta(R)}$ for some $R \subseteq S$ or the identity. Furthermore, the cut conditions do not change when applied to the restricted set $X_S|_{T_0,P_0}$. Therefore, we can apply all our cut conditions simultaneously. On the contrary, σ_R for some $R \subseteq S$ restricted to $X_S|_{T_0,P_0}$ does not necessarily have image $X_S|_{T_0,P_0}$, e.g. for the map σ_S we have $\sigma_S(X_S|_{T_0,P_0}) = \{\mathbb{1}_S\}$. Therefore, we cannot expect partial optimality to hold when applying conditions together that include at least one join condition.

5.2. Join Conditions

Let us assume the existence of an optimal solution x^* to $\text{SP3}_{S,c}$ such that $x^*_{ij} = 1$ for some $ij \in \binom{S}{2}$. We define $X_S|_{x_{ij}=1} = \{x \in X_S \mid x_{ij} = 1\}$. Then, we have

$$\min_{x \in X_S} \phi_c(x) = \min_{x \in X_S|_{x_{ij}=1}} \phi_c(x) \ . \tag{21}$$

Let $S' = S \setminus \{j\}$. Now, we relate feasible vectors of $X_S|_{x_{ij}=1}$ to feasible vectors of $X_{S'}$. We observe that for any $x \in X_S|_{x_{ij}=1}$ we have $\forall p \in S \setminus \{i,j\} \colon x_{pi} = x_{pj}$. We define $\varphi_{ij} \colon X_S|_{x_{ij}=1} \to X_{S'}$ as

$$\begin{aligned} \varphi_{ij}(x)_{pi} &= x_{pi} \quad \forall x \in X_S|_{x_{ij}=1} \quad \forall p \in S' \setminus \{i\} \\ \varphi_{ij}(x)_{pq} &= x_{pq} \quad \forall x \in X_S|_{x_{ij}=1} \quad \forall pq \in \tbinom{S' \setminus \{i\}}{2} \ . \end{aligned}$$

It is easy to see that φ_{ij} is bijective. Proposition 5.1 below shows that solving the right-hand side of (21) is equivalent to solving a smaller instance of the original problem.

Proposition 5.1. *Let $S \neq \emptyset$ and $c \in \mathbb{R}^{\mathcal{I}_S}$. Moreover, let $ij \in \binom{S}{2}$ and $S' = S \setminus \{j\}$, and let $c' \in \mathbb{R}^{\mathcal{I}_{S'}}$ such that*

$$\begin{aligned} c'_{pqr} &= c_{pqr} && \forall pqr \in \tbinom{S' \setminus \{i\}}{3} \\ c'_{pqi} &= c_{pqi} + c_{pqj} && \forall pq \in \tbinom{S' \setminus \{i\}}{2} \\ c'_{pq} &= c_{pq} && \forall pq \in \tbinom{S' \setminus \{i\}}{2} \\ c'_{pi} &= c_{pi} + c_{pj} + c_{pij} && \forall p \in S' \setminus \{i\} \\ c'_\emptyset &= c_0 + c_{ij} \ . \end{aligned}$$

Furthermore, let $\varphi_{ij} \colon X_S|_{x_{ij}=1} \to X_{S'}$ be the map that relates feasible vectors of $X_S|_{x_{ij}=1}$ to feasible vectors of $X_{S'}$. Then, we have $\min_{x \in X_S|_{x_{ij}=1}} \phi_c(x) = \min_{x \in X_{S'}} \phi_{c'}(x)$. Moreover, if $x^* \in \operatorname{argmin}_{x \in X_S|_{x_{ij}=1}} \phi_c(x)$, then $\varphi_{ij}(x^*) \in \operatorname{argmin}_{x \in X_{S'}} \phi_{c'}(x)$.

5.3. Mixing Cut and Join Conditions

Here, we describe how we apply the partial optimality properties recursively. As soon as a condition leads to a smaller instance, we start the procedure again on the smaller set (in case of a join) or sets (in case of a cut). Firstly, we apply Proposition 3.1, which leaves us with independent sub-problems. Secondly, we apply our join conditions until we find a pair $ij \in \binom{S}{2}$ or triplet $ijk \in \binom{S}{3}$ to join, starting from Corollary 3.11 and then moving on to Propositions 3.4 and 3.6, Corollaries 3.8 and 3.9 and Proposition 3.5, in

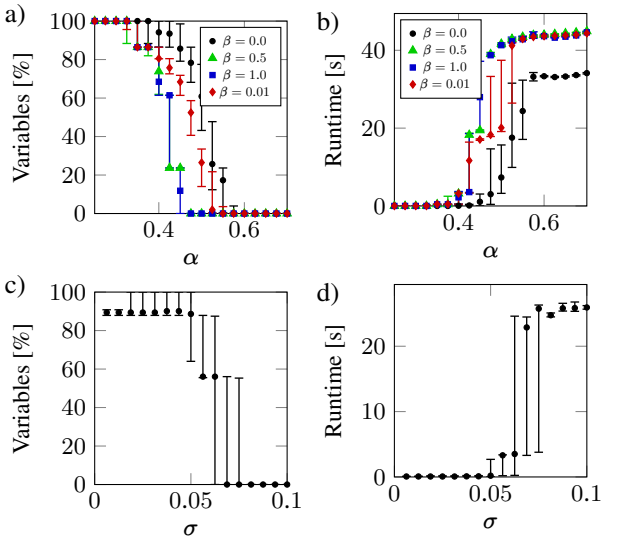

Figure 3. We report above the percentage of fixed variables and triples after applying all conditions jointly, as described in 5.3, as well as the corresponding runtimes. (a) and (b) show these for the partition dataset with respect to the parameters α and β and for 48 elements. (c) and (d) show these for the geometric dataset with respect to the parameter σ and for 45 points.

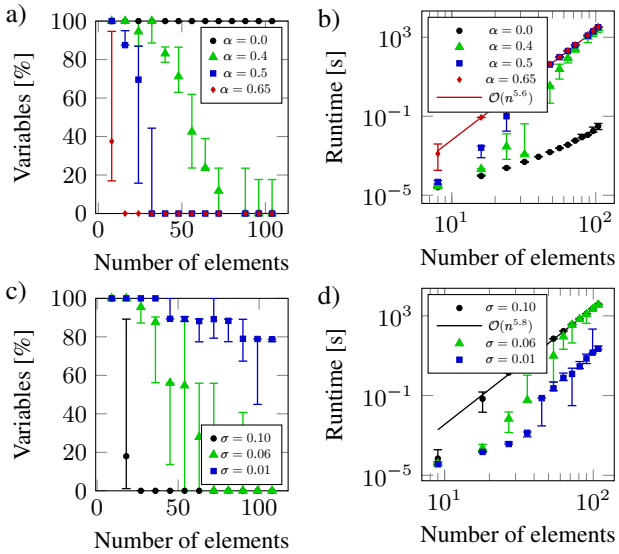

Figure 4. We report above the percentage of fixed variables and triples after applying all conditions jointly, as described in Section 5.3, as well as the corresponding runtimes. (a) and (b) show these for the partition dataset with respect to the number of elements and $\beta = 0.5$. (c) and (d) show these for the geometric dataset with respect to the number of elements.

this order. Thirdly, we apply the remaining cut conditions, which can be applied jointly, as we have seen in Section 5.1. The order in which we apply the join conditions is arbitrary and possibly sub-optimal.

6. Numerical Experiments

Next, we examine the effectiveness of the algorithms empirically, on two datasets. For both, we report the percentage of fixed variables and triples, as well as the runtime. More specifically, we report the median as well as lower and upper quartile over 30 instances. We apply all partial optimality conditions jointly, as described in Section 5.3, and we also evaluate the effectiveness of each condition separately. All algorithms are implemented in C++ and run on one core of an Intel Core i5-6600 equipped with 16 GB of RAM.

Partition Dataset. We define one dataset with respect to a partition $\mathcal{R} = \{R_1, R_2, R_3, R_4\}$ of $|S| = 8n$ elements, with $|R_1| = n$, $|R_2| = |R_3| = 2n$ and $|R_4| = 3n$, where $n \in \mathbb{N}$ is between 1 and 13. See also Figure 1a. With respect to a parameter $\alpha \in [0, 1]$, we draw the costs of pairs and triplets from two Gaussian distributions with means $-1 + \alpha$ and $1 - \alpha$, depending on whether their elements belong to the same set or distinct sets in the partition $\mathcal{R}$, and with standard deviation $\sigma = \sigma_0 + \alpha(\sigma_1 - \sigma_0)$, with $\sigma_0 = 0.1$ and $\sigma_1 = 0.4$. With respect to a parameter $\beta \in [0, 1]$, we multiply the costs of pairs by $1 - \beta$, and the costs of triples by β. The higher α is, the harder the problem becomes. The higher β is, the more important the costs of triples become.

For this dataset, the percentage of pairs and triples fixed by applying all conditions jointly, as described in Section 5.3, is shown in Figure 3a. It can be seen from this figure that the fraction of fixed variables decreases with increasing α. As α rises, the runtime increases but remains below one minute for all the instances; see Figure 3b. Varying β does not affect the overall trend. However, the fraction of fixed variables decreases as soon as triple costs are introduced. The percentage of pairs and triples fixed by applying Propositions 3.1 to 3.3 and Corollary 3.11 separately is shown in Figure 5. The other partial optimality conditions do not fix any variables of these instances. While all cut conditions settle the value of some variables, this is not the case for the join statements. In fact, only one join condition provides partial optimality in this case: Corollary 3.11. Interestingly, this is the one statement that fixes the most variables for almost all instances of this dataset. For $\beta = 0.5$ and with respect the instance size, the runtime and percentage of variables fixed by applying all conditions jointly are shown in Figure 4a and b. It can be seen that, as the size of the instance increases, the number of fixed variables decreases, while the runtime increases. For $\alpha \in \{0.4, 0.5, 0.65\}$, the runtime roughly converges to $\mathcal{O}(n^{5.6})$.

Geometric Dataset. Next, we consider a dataset of instances that arise from the geometric problem of finding equilateral triangles in a noisy point cloud; see Figure 1b. For this, we fix three equilateral triangles in the plane. For each vertex $\vec{a}$ of a triangle, we draw a number of points from

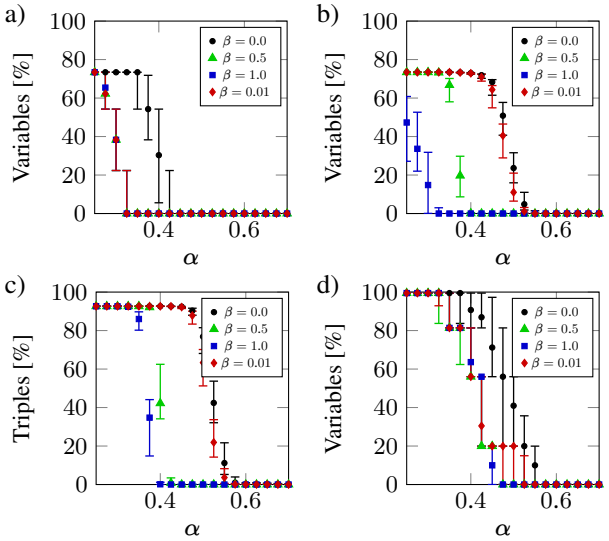

Figure 5. For the partition dataset with 48 elements, we report above the percentage of fixed pairs and triples with respect to the parameters α and β when applying Propositions 3.1 to 3.3, (a)–(c), and Corollary 3.11, (d), separateley.

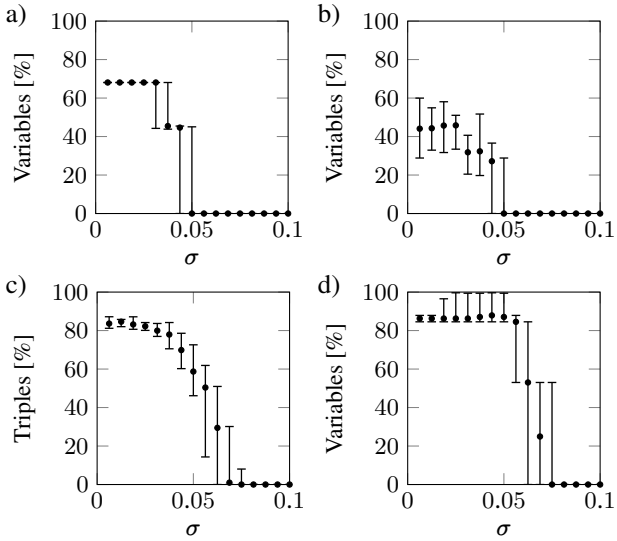

Figure 6. For the geometric dataset with 45 points, we report above the percentage of fixed variables and triples with respect to the parameter σ, when employing Propositions 3.1 to 3.3, (a)–(c), and Corollary 3.11, (d), individually.

a Gaussian distribution with mean $\vec{a}$ and covariance matrix $\sigma^2\mathbb{1}$. For any three points $\vec{a}_p, \vec{a}_q, \vec{a}_r$, let φ_p, φ_q, φ_r be the interior angles of the triangle spanned by these points, and let d^{max}_{pqr} and d^{min}_{pqr} be the maximum and minimum length of edges in this triangle. If the three points are mutually close, $d^{max}_{pqr} \leq 4\sigma$, we reward solutions in which these belong to the same set by letting $c_{pqr} = -1 + \frac{d^{max}_{pqr}}{4\sigma}$. If only two points are close, $d^{max}_{pqr} > 4\sigma$ and $d^{min}_{pqr} \leq 4\sigma$, we let $c_{pqr} = 0$. If the three points are mutually far apart, $d^{min}_{pqr} > 4\sigma$, we calculate the sum of the deviations of the inner angles from $\frac{\pi}{3}$, $\delta_{pqr} = \sum_i |\varphi_i - \frac{\pi}{3}|$. If this quantity is below $\frac{\pi}{6}$, we let $c_{pqr} = -1 + \frac{6\delta_{pqr}}{\pi}$. Otherwise, $c_{pqr} = \frac{6}{7}\frac{\delta_{pqr} - \frac{\pi}{6}}{\pi}$.

For this dataset, the percentage of pairs and triples fixed by applying all conditions jointly, as described in Section 5.3, is reported in Figure 3c. Here, the hardness of the instances is embodied by σ. The number of points is 45. As σ increases, the percentage of fixed variables decreases. The runtime increases, as can be seen from Figure 3d, and stays below one minute for all these instances. The percentage of pairs and triples fixed by applying Propositions 3.1 to 3.3 and Corollary 3.11 separately is shown in Figure 6. Also here, all the cut conditions are effective whereas the only useful join condition is Corollary 3.11. Moreover, Corollary 3.11 is the most effective condition overall. The runtime and percentage of variables fixed by applying all conditions jointly are shown with respect the size of the instance in Figure 4c and d. As the size increases, the number of fixed variables decreases while the runtime increases. For $\sigma \in \{0.06, 0.1\}$, the runtime roughly converges to $\mathcal{O}(n^{5.8})$.

7. Conclusion

We establish partial optimality conditions for the cubic set partition problem, which can be seen as the special case of cubic correlation clustering for complete graphs. In particular, we generalize all such conditions known for correlation clustering with a linear objective function to arbitrary cubic objective functions. In addition, we establish new partial optimality conditions. Furthermore, we define and implement exact algorithms and heuristics for testing all established conditions efficiently. Lastly, we quantify the effect of these algorithms on two datasets. Regarding these numerical experiments, we note that all cut conditions are effective on the tested datasets, whereas join conditions pose a bigger challenge. In fact, Proposition 3.10, in its simplified form of Corollary 3.11, is the only join property that is beneficial in our numerical experiments. Yet, in almost all cases, it is the one statement that fixes the most variables (Figures 5 and 6). We remark that Corollary 3.11 is one of the newly proposed conditions. Perspectives for future work include the exploitation of sparsity of non-zero cost coefficients, as well as applications to subspace clustering and object recognition.

Acknowledgements. Bjoern Andres and David Stein acknowledge funding by the Federal Ministry of Education and Research of Germany, from grant 01LC2006A.

References

Adams, W. P., Lassiter, J. B., and Sherali, H. D. Persistency in 0-1 polynomial programming. *Mathemat-*

ics of Operations Research, 23(2):359–389, 1998. doi: 10.1287/moor.23.2.359.

Agarwal, S., Lim, J., Zelnik-Manor, L., Perona, P., Kriegman, D., and Belongie, S. Beyond pairwise clustering. In *CVPR*, 2005. doi: 10.1109/CVPR.2005.89.

Alush, A. and Goldberger, J. Ensemble segmentation using efficient integer linear programming. *Transactions on Pattern Analysis and Machine Intelligence*, 34(10):1966–1977, 2012. doi: 10.1109/TPAMI.2011.280.

Billionnet, A. and Sutter, A. Persistency in quadratic 0–1 optimization. *Mathematical Programming*, 54(1):115–119, 1992. doi: 10.1007/BF01586044.

Boost. Boost C++ Libraries. https://www.boost.org, 2022.

Boros, E., Hammer, P. L., Sun, R., and Tavares, G. A max-flow approach to improved lower bounds for quadratic unconstrained binary optimization (qubo). *Discrete Optimization*, 5(2):501–529, 2008. doi: 10.1016/j.disopt.2007.02.001.

Goldberg, A. V. and Tarjan, R. E. A new approach to the maximum-flow problem. *Journal of the ACM*, 35(4):921–940, 1988. doi: 10.1145/48014.61051.

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

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

Kappes, J. H., Speth, M., Reinelt, G., and Schnörr, C. Towards efficient and exact map-inference for large scale discrete computer vision problems via combinatorial optimization. In *CVPR*, 2013. doi: 10.1109/CVPR.2013.229.

Kappes, J. H., Speth, M., Reinelt, G., and Schnörr, C. Higher-order segmentation via multicuts. *Computer Vision and Image Understanding*, 143:104–119, 2016. doi: 10.1016/j.cviu.2015.11.005.

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

Kohli, P., Shekhovtsov, A., Rother, C., Kolmogorov, V., and Torr, P. On partial optimality in multi-label mrfs. In *ICML*, 2008. doi: 10.1145/1390156.1390217.

Kolmogorov, V. and Zabin, R. What energy functions can be minimized via graph cuts? *Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004. doi: 10.1109/TPAMI.2004.1262177.

Lange, J., Andres, B., and Swoboda, P. Combinatorial persistency criteria for multicut and max-cut. In *CVPR*, 2019. doi: 10.1109/CVPR.2019.00625.

Lange, J.-H., Karrenbauer, A., and Andres, B. Partial optimality and fast lower bounds for weighted correlation clustering. In *ICML*, 2018. URL https://proceedings.mlr.press/v80/lange18a.html.

Levinkov, E., Kardoost, A., Andres, B., and Keuper, M. Higher-order multicuts for geometric model fitting and motion segmentation. *Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2022. doi: 10.1109/TPAMI.2022.3148795.

Ochs, P. and Brox, T. Higher order motion models and spectral clustering. In *CVPR*, 2012. doi: 10.1109/CVPR.2012.6247728.

Purkait, P., Chin, T.-J., Sadri, A., and Suter, D. Clustering with hypergraphs: The case for large hyperedges. *Transactions on Pattern Analysis and Machine Intelligence*, 39(9):1697–1711, 2017. doi: 10.1109/TPAMI.2016.2614980.

Shekhovtsov, A. *Exact and Partial Energy Minimization in Computer Vision*. PhD thesis, Center for Machine Perception, Czech Technical University, Prague, 2013.

Shekhovtsov, A. Maximum persistency in energy minimization. In *CVPR*, 2014. doi: 10.1109/CVPR.2014.152.

Shekhovtsov, A., Swoboda, P., and Savchynskyy, B. Maximum persistency via iterative relaxed inference with graphical models. In *CVPR*, 2015. doi: 10.1109/CVPR.2015.7298650.

Stoer, M. and Wagner, F. A simple min-cut algorithm. *Journal of the ACM*, 44(4):585–591, 1997. doi: 10.1145/263867.263872.

Veldt, N. Correlation clustering via strong triadic closure labeling: Fast approximation algorithms and practical lower bounds. In *ICML*, 2022. URL https://proceedings.mlr.press/v162/veldt22a.html.

A. Appendices

A.1. Proofs

Proof of Proposition 2.1. For each partition Π of the set S and every distinct $p, q \in S$, let $x_{pq} = 1$ if and only if p and q are in the same set of Π. This establishes a one-to-one relation between the set P_S of all partitions of S and the set X_S (Grötschel & Wakabayashi, 1989). Under this bijection, the objective functions of Definition 1.1 and Proposition 2.1 are equivalent. □

Proof of Proposition 2.3. Let x^* be an optimal solution to $\min_{x \in X} \phi(x)$ such that $x^* \notin Q$. Then $\sigma(x^*)$ is also an optimal solution to $\min_{x \in X} \phi(x)$ and $\sigma(x^*) \in Q$. □

Proof of Proposition 3.1. We define $\sigma \colon X_S \to X_S$ such that for all $x \in X_S$ we have that

$$\sigma(x) := \begin{cases} x & \text{if } x_{ij} = 0 \ \forall ij \in \delta(R) \\ \sigma_{\delta(R)}(x) & \text{otherwise} \end{cases}.$$

For any $x \in X_S$, let $x' = \sigma(x)$. First, the map σ is such that $x'_{ij} = 0$ for all $ij \in \delta(R)$. Second, for any $x \in X_S$ such that there exists $ij \in \delta(R)$ such that $x_{ij} = 1$, we have that

$$\phi_c(x') - \phi_c(x) = - \sum_{pqr \in T_{\delta(R)}} c_{pqr} x_{pq} x_{pr} x_{qr} - \sum_{pq \in \delta(R)} c_{pq} x_{pq} \leq - \sum_{pqr \in T_{\delta(R)} \cap T^-} c_{pqr} - \sum_{pq \in \delta(R) \cap P^-} c_{pq} = 0.$$

The last equality is due to the fact that those sums vanish by assumptions (1) and (2). Applying Corollary 2.4 concludes the proof. □

Proof of Proposition 3.2. Let $\sigma \colon X_S \to X_S$ be constructed as

$$\sigma(x) := \begin{cases} x & \text{if } x_{ij} = 0 \\ \sigma_{\delta(R)}(x) & \text{otherwise} \end{cases}.$$

For any $x \in X_S$, let $x' = \sigma(x)$. First of all, the map σ is such that $x'_{ij} = 0$ for all $x \in X_S$. Next, for any $x \in X_S$ such that $x_{ij} = 1$ we have that

$$\begin{aligned} \phi_c(x') - \phi_c(x) &= -c_{ij} - \sum_{pqr \in T_{\delta(R)}} c_{pqr} x_{pq} x_{pr} x_{qr} - \sum_{\substack{pq \in \delta(R) \\ pq \neq ij}} c_{pq} x_{pq} \\ &\leq -c_{ij} + \sum_{pqr \in T_{\delta(R)}} c^-_{pqr} + \sum_{\substack{pq \in \delta(R) \\ pq \neq ij}} c^-_{pq} = -c^+_{ij} + \sum_{pqr \in T_{\delta(R)}} c^-_{pqr} + \sum_{pq \in \delta(R)} c^-_{pq} \leq 0. \end{aligned}$$

The last inequality follows from assumption (3). We conclude the proof by applying Corollary 2.4. □

Proof of Proposition 3.3. We define $\sigma \colon X_S \to X_S$ as

$$\sigma(x) := \begin{cases} x & \text{if } x_{ij} x_{ik} x_{jk} = 0 \\ \sigma_{\delta(R)}(x) & \text{otherwise} \end{cases}.$$

For any $x \in X_S$, we denote $\sigma(x)$ by x'. We observe that $x'_{ij} x'_{ik} x'_{jk} = 0$ for all $x \in X_S$. Second, for any $x \in X_S$ such that $x_{ij} x_{ik} x_{jk} = 1$ it holds that

$$\begin{aligned} \phi_c(x') - \phi_c(x) &= - \sum_{pqr \in T_{\delta(R)}} c_{pqr} x_{pq} x_{pr} x_{qr} - \sum_{pq \in \delta(R)} c_{pq} x_{pq} \leq -c_{ijk} - c_{ij} - c_{ik} + \sum_{\substack{pqr \in T_{\delta(R)} \\ pqr \neq ijk}} c^-_{pqr} + \sum_{\substack{pq \in \delta(R) \\ pq \notin \{ij, ik\}}} c^-_{pq} \\ &= -c^+_{ijk} - c^+_{ij} - c^+_{ik} + \sum_{pqr \in T_{\delta(R)}} c^-_{pqr} + \sum_{pq \in \delta(R)} c^-_{pq} \leq 0. \end{aligned}$$

The last inequality holds because of assumption (4). Applying Proposition 2.3 with $Q = \{x \in X_S \mid x_{ij} x_{ik} x_{jk} = 0\}$ concludes the proof. □

Proof of Proposition 3.4. Let $\sigma\colon X_S \to X_S$ be such that for all $x \in X_S$ it holds that

$$\sigma(x) := \begin{cases} x & \text{if } x_{ij} = 1 \\ \left(\sigma_{ij} \circ \sigma_{\delta(R)}\right)(x) & \text{otherwise} \end{cases}.$$

For any $x \in X_S$, let $x' = \sigma(x)$. Firstly, the map σ is such that $x'_{ij} = 1$ for all $x \in X_S$. We now show that σ is improving. In particular, let $x \in X_S$ such that $x_{ij} = 0$. We observe that $x'_{pq} = x_{pq}$ for all $pq \notin \delta(R)$. Therefore,

$$\begin{aligned}
\phi_c(x') - \phi(x) &= \sum_{pqr \in T_{\delta(R)}} c_{pqr}\left(x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr}\right) + \sum_{pq \in \delta(R)} c_{pq}\left(x'_{pq} - x_{pq}\right) \\
&= \sum_{pqr \in T_{\delta(R)} \setminus T_{\{ij\}}} c_{pqr}\left(x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr}\right) + \sum_{pqr \in T_{\{ij\}}} c_{pqr}x'_{pq}x'_{pr}x'_{qr} + c_{ij} \\
&+ \sum_{\substack{pq \in \delta(R) \\ pq \neq ij}} c_{pq}\left(x'_{pq} - x_{pq}\right) \\
&\leq \sum_{pqr \in T_{\delta(R)} \setminus T_{\{ij\}}} |c_{pqr}| + \sum_{pqr \in T_{\{ij\}}} c^+_{pqr} + c_{ij} + \sum_{\substack{pq \in \delta(R) \\ pq \neq ij}} |c_{pq}| \\
&= \sum_{pqr \in T_{\delta(R)}} |c_{pqr}| - \sum_{pqr \in T_{\{ij\}}} c^-_{pqr} - 2c^-_{ij} + \sum_{pq \in \delta(R)} |c_{pq}| \leq 0.
\end{aligned}$$

We remark that the last inequality is due to assumption (5). □

Proof of Proposition 3.5. We construct $\sigma\colon X_S \to X_S$ as follows

$$\sigma(x) := \begin{cases} x & \text{if } x_{ij}x_{ik}x_{jk} = 1 \\ \left(\sigma_{ijk} \circ \sigma_{\delta(R)}\right)(x) & \text{otherwise} \end{cases}.$$

For any $x \in X_S$, let $x' = \sigma(x)$. First, the map σ is such that $x'_{ij}x'_{ik}x'_{jk} = 1$ for all $x \in X_S$. Note that, for any $x \in X_S$ such that $x_{ij}x_{ik}x_{jk} = 0$, we have that $x'_{pq} \geq x_{pq}$ for all $pq \notin \delta(R)$. It follows that

$$\begin{aligned}
\phi_c(x') - \phi(x) &= \sum_{\substack{pqr \in T_{\delta(R)} \\ pqr \neq ijk}} c_{pqr}(x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr}) + c_{ijk} + \sum_{pqr \notin T_{\delta(R)}} c_{pqr}(x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr}) \\
&+ \sum_{pq \in \{ij,ik,jk\}} c_{pq}(1 - x_{pq}) + \sum_{\substack{pq \in \delta(R) \\ pq \notin \{ij,ik\}}} c_{pq}(x'_{pq} - x_{pq}) + \sum_{pq \notin \delta(R) \cup \{jk\}} c_{pq}(x'_{pq} - x_{pq}) \\
&\leq c_{ijk} + \max_{\substack{x \in Xijk \\ x_{ij}x_{ik}x_{jk} = 0}} \sum_{pq \in \binom{ijk}{2}} c_{pq}(1 - x_{pq}) + \sum_{\substack{pqr \in T_{\delta(R)} \\ pqr \neq ijk}} |c_{pqr}| + \sum_{\substack{pqr \in T^+ \\ pqr \notin T_{\delta(R)}}} c_{pqr} \\
&+ \sum_{\substack{pq \in \delta(R) \\ pq \notin \{ij,ik\}}} |c_{pq}| + \sum_{\substack{pq \in P^+ \\ pq \notin (\delta(R) \cup \{jk\})}} c_{pq} \\
&= c_{ijk} + \max_{\substack{x \in X_{ijk} \\ x_{ij}x_{ik}x_{jk} = 0}} \sum_{pq \in \binom{ijk}{2}} c_{pq}(1 - x_{pq}) + \sum_{\substack{pqr \in T^+ \cup T_{\delta(R)} \\ pqr \neq ijk}} |c_{pqr}| + \sum_{\substack{pq \in P^+ \cup \delta(R) \\ pq \notin \{ij,ik,jk\}}} |c_{pq}| \\
&= -2c^-_{ijk} - 2c^-_{ij} - 2c^-_{ik} - c^-_{jk} - \min_{\substack{x \in X_{ijk} \\ x_{ij}x_{ik}x_{jk} = 0}} \sum_{pq \in \binom{ijk}{2}} c_{pq}x_{pq} \\
&+ \sum_{pqr \in T_{\delta(R)}} c^-_{pqr} + \sum_{pqr \in \binom{S}{3}} c^+_{pqr} + \sum_{pq \in \binom{S}{2}} c^+_{pq} + \sum_{pq \in \delta(R)} c^-_{pq} \leq 0.
\end{aligned}$$

Assumption (6) provides the last inequality. We arrive at the thesis by applying Proposition 2.3 with $Q = \{x \in X_S \mid x_{ij}x_{ik}x_{jk} = 1\}$. □

Proof of Proposition 3.6. Let $\sigma\colon X_S \to X_S$ be defined as

$$\sigma(x) := \begin{cases} x & \text{if } x_{ik} = 1 \\ (\sigma_{ik} \circ \sigma_{\delta(R)})(x) & \text{if } x_{ik} = x_{ij} = 0, x_{jk} = 1 \\ (\sigma_{ik} \circ \sigma_{\delta(R')})(x) & \text{if } x_{ik} = x_{jk} = 0, x_{ij} = 1 \\ (\sigma_{ijk} \circ \sigma_{\delta(ijk)})(x) & \text{if } x_{ik} = x_{ij} = x_{jk} = 0 \end{cases}.$$

We use the notation $x' = \sigma(x)$ for all $x \in X$. We note that the map σ is such that $x'_{ik} = 1$ for all $x \in X_S$. Second, for any $x \in X_S$ such that $x_{ik} = x_{ij} = 0$ and $x_{jk} = 1$, the map $\sigma_{ik} \circ \sigma_{\delta(R)}$ is such that $x'_{ij} = x'_{ik} = x'_{jk} = 1$ and $x'_{pq} = x_{pq}$ for any $pq \notin \delta(R)$. It holds that

$$\begin{aligned}
\phi_c(x') - \phi_c(x) &= c_{ijk} + c_{ij} + c_{ik} + \sum_{\substack{pqr \in T_{\{ij,ik\}} \\ pqr \neq ijk}} c_{pqr} x'_{pq} x'_{pr} x'_{qr} \\
&\quad + \sum_{\substack{pqr \in T_{\delta(R)} \\ pqr \notin T_{\{ij,ik\}}}} c_{pqr} \left(x'_{pq} x'_{pr} x'_{qr} - x_{pq} x_{pr} x_{qr} \right) + \sum_{\substack{pq \in \delta(R) \\ pq \notin \{ij,ik\}}} c_{pq} (x'_{pq} - x_{pq}) \\
&\leq c_{ijk} + c_{ij} + c_{ik} + \sum_{\substack{pqr \in T_{\{ij,ik\}} \cap T^+ \\ pqr \neq ijk}} c_{pqr} + \sum_{\substack{pqr \in T_{\delta(R)} \\ pqr \notin T_{\{ij,ik\}}}} |c_{pqr}| + \sum_{\substack{pq \in \delta(R) \\ pq \notin \{ij,ik\}}} |c_{pq}| \\
&= -c^-_{ijk} - 2c^-_{ij} - 2c^-_{ik} + \sum_{pqr \in T_{\{ij,ik\}}} c^+_{pqr} + \sum_{pqr \in T_{\delta(R)}} |c_{pqr}| - \sum_{pqr \in T_{\{ij,ik\}}} |c_{pqr}| + \sum_{pq \in \delta(R)} |c_{pq}| \\
&= -c^-_{ijk} - 2c^-_{ij} - 2c^-_{ik} - \sum_{pqr \in T_{\{ij,ik\}}} c^-_{pqr} + \sum_{pqr \in T_{\delta(R)}} |c_{pqr}| + \sum_{pq \in \delta(R)} |c_{pq}| \leq 0.
\end{aligned}$$

The last inequality follows from assumption (7). Third, for any $x \in X_S$ such that $x_{ik} = x_{jk} = 0$ and $x_{ij} = 1$, the map $\sigma_{ik} \circ \sigma_{\delta(R')}$ is improving by analogous arguments and assumption (8). Finally, for any $x \in X_S$ such that $x_{ik} = x_{jk} = x_{ij} = 0$, the map $\sigma_{ijk} \circ \sigma_{\delta(ijk)}$ is such that

$$(\sigma_{ijk} \circ \sigma_{\delta(ijk)})_{pq} = \begin{cases} 0 & \text{if } pq \in \delta(ijk) \\ 1 & \text{if } pq \in \{ij, ik, jk\} \\ x_{pq} & \text{otherwise} \end{cases}.$$

Therefore,

$$\begin{aligned}
\phi_c(x') - \phi_c(x) &= c_{ijk} + c_{ij} + c_{ik} + c_{jk} - \sum_{pqr \in T_{\delta(ijk)} \setminus T_{\{ij,ik,jk\}}} c_{pqr} x_{pq} x_{pr} x_{qr} - \sum_{pq \in \delta(ijk)} c_{pq} x_{pq} \\
&\leq c_{ijk} + c_{ij} + c_{ik} + c_{jk} + \sum_{\substack{pqr \in T_{\delta(ijk)} \cap T^- \\ pqr \notin T_{\{ij,ik,jk\}}}} |c_{pqr}| + \sum_{pq \in \delta(ijk) \cap P^-} |c_{pq}| \leq 0.
\end{aligned}$$

The last inequality is true thanks to to assumption (9). Applying Corollary 2.4 concludes the proof. □

Proof of Proposition 3.7. First, we prove a lemma that establishes a relation which will be needed at the end of this proof.

Lemma A.1. *Let $S \neq \emptyset$ and $c \in \mathbb{R}^{\mathcal{I}_S}$. We define $c' \in \mathbb{R}^{\mathcal{I}_S}$ as in (10), (11), (12) for $S_H = S$. Then for any partition $\mathcal{R}$ of S it holds that*

$$\begin{aligned}
\phi_{c'}(x^{\mathcal{R}}) = \frac{1}{2} \sum_{pqr \in \binom{S}{3}} c_{pqr} \prod_{uv \in \binom{pqr}{2}} (1 - x^{\mathcal{R}}_{uv}) + \sum_{pqr \in \binom{S}{3}} c_{pqr} \sum_{uv \in \binom{pqr}{2}} x^{\mathcal{R}}_{uv} \prod_{\substack{u'v' \in \binom{pqr}{2} \\ u'v' \neq uv}} (1 - x^{\mathcal{R}}_{u'v'}) \\
+ \sum_{pq \in \binom{S}{2}} c_{pq} (1 - x^{\mathcal{R}}_{pq})
\end{aligned} \tag{22}$$

$$= \frac{1}{2} \sum_{RR'R'' \in \binom{\mathcal{R}}{3}} \sum_{pqr \in T_{RR'R''}} c_{pqr} + \sum_{RR' \in \binom{\mathcal{R}}{2}} \sum_{pq \in \delta(R,R')} c_{pq} + \sum_{RR' \in \binom{\mathcal{R}}{2}} \left(\sum_{pqr \in T_{RRR'}} c_{pqr} + \sum_{pqr \in T_{RR'R'}} c_{pqr} \right), \quad (23)$$

where $x^{\mathcal{R}}$ denotes the feasible vector corresponding to the partition $\mathcal{R}$ of S.

Proof Lemma A.1. We use the fact that for any partition $\mathcal{R}$ of S and any $pqr \in \binom{S}{3}$ we have that $x^{\mathcal{R}}_{pq} x^{\mathcal{R}}_{qr} = x^{\mathcal{R}}_{pq} x^{\mathcal{R}}_{pr} = x^{\mathcal{R}}_{pr} x^{\mathcal{R}}_{qr} = x^{\mathcal{R}}_{pq} x^{\mathcal{R}}_{pr} x^{\mathcal{R}}_{qr}$. Expanding the inner products and the inner sums leads to

$$\prod_{uv \in \binom{pqr}{3}} \left(1 - x^{\mathcal{R}}_{uv}\right) = 1 - x^{\mathcal{R}}_{pq} - x^{\mathcal{R}}_{pr} - x^{\mathcal{R}}_{qr} + 2 x^{\mathcal{R}}_{pq} x^{\mathcal{R}}_{pr} x^{\mathcal{R}}_{qr},$$

$$\sum_{uv \in \binom{pqr}{2}} x^{\mathcal{R}}_{uv} \prod_{\substack{u'v' \in \binom{pqr}{2} \\ u'v' \notin \{uv\}}} (1 - x^{\mathcal{R}}_{u'v'}) = x^{\mathcal{R}}_{pq} + x^{\mathcal{R}}_{pr} + x^{\mathcal{R}}_{qr} - 3 x^{\mathcal{R}}_{pq} x^{\mathcal{R}}_{qr} x^{\mathcal{R}}_{pr}.$$

By plugging in and collecting terms we conclude the proof for equality (22). Equality (23) follows instead from the following observations:

$$\prod_{ab \in \binom{pqr}{3}} \left(1 - x^{\mathcal{R}}_{ab}\right) = 1 \Leftrightarrow \exists\, RR'R'' \in \binom{\mathcal{R}}{3} \colon pqr \in T_{RR'R''},$$

$$\sum_{ab \in \binom{pqr}{2}} x^{\mathcal{R}}_{ab} \prod_{a'b' \in \binom{pqr}{2} \setminus \{ab\}} (1 - x^{\mathcal{R}}_{a'b'}) = 1 \Leftrightarrow \exists\, RR' \in \binom{\mathcal{R}}{2} \colon \left(pqr \in T_{RRR'} \vee pqr \in T_{RR'R'}\right),$$

$$1 - x^{\mathcal{R}}_{pq} = 1 \Leftrightarrow \exists\, RR' \in \binom{\mathcal{R}}{2} \colon pq \in \delta(R, R').$$

This concludes the proof. ◇

We define $\sigma \colon X_S \to X_S$ as

$$\sigma(x) := \begin{cases} x & \text{if } x_{ij} = 1 \\ (\sigma_{S_H} \circ \sigma_{\delta(S_H)})(x) & \text{otherwise} \end{cases}.$$

Let $x' = \sigma(x)$ for any $x \in X_S$. It is easy to see that $x'_{ij} = 1$ for all $x \in X_S$. Similarly to before, we show that σ is an improving map. For any $x \in X_S$ such that $x_{ij} = 1$ we have that $\phi_c(x') - \phi_c(x) = 0$ by definition of x'. Now, we consider $x \in X_S$ such that $x_{ij} = 0$. Let $P_H = \binom{S_H}{2}$ and $T_H = \binom{S_H}{3}$. We denote the restriction of x containing only components corresponding to elements in P_H by $x|_{P_H}$. Let $\mathcal{R}$ be the partition of S such that $x = x^{\mathcal{R}}$, and let $\mathcal{R}_H$ be the induced partition of S_H such that $x|_{P_H} = x^{\mathcal{R}_H}$. Since $x_{ij} = 0$, there exist $R_1, R_2 \in \mathcal{R}_H$ such that $i \in R_1$, $j \in R_2$. We have that

$$x'_{pq} = \begin{cases} 1 & \text{if } pq \in P_H \\ 0 & \text{if } pq \in \delta(S_H) \\ x_{pq} & \text{otherwise} \end{cases}.$$

Therefore, it follows that

$$\phi_c(x') - \phi_c(x) = \sum_{pq \in P_H} c_{pq}(1 - x_{pq}) + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq} x_{pr} x_{qr}) - \sum_{pq \in \delta(S_H)} c_{pq} x_{pq} - \sum_{pqr \in T_{\delta(S_H)}} c_{pqr} x_{pq} x_{pr} x_{qr}. \quad (24)$$

In order to find an upper bound for the sums over P_H and T_H, we show that there exists a subset $R \subset S_H$ with $i \in R$ and $j \in S_H \setminus R$ such that

$$\sum_{pq \in P_H} c_{pq} \left(1 - x_{pq}\right) + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq} x_{pr} x_{qr}) \leq \sum_{pq \in \delta(R, S_H \setminus R)} c_{pq} + \sum_{pqr \in T_{\delta(R, V_H \setminus R)} \cap T_H} c_{pqr}. \quad (25)$$

For the sake of contradiction, we assume that there is no such $R \subset S_H$. For any $\mathcal{R}' \subset \mathcal{R}_H$ let $R_{\mathcal{R}'} = \bigcup_{P' \in \mathcal{R}'} P'$. Furthermore we define $t\colon \binom{\mathcal{R}_H}{2} \cup \binom{\mathcal{R}_H}{3} \to \mathbb{R}$ and $p\colon \binom{\mathcal{R}_H}{2} \to \mathbb{R}$ as

$$\begin{aligned} t_{RR'R''} &= \sum_{pqr \in T_{RR'R''}} c_{pqr}, && \forall RR'R'' \in \tbinom{\mathcal{R}_H}{3} \\ t_{RR'} &= \sum_{pqr \in T_{RRR'} \cup T_{RR'R'}} c_{pqr}, && \forall RR' \in \tbinom{\mathcal{R}_H}{2} \\ p_{RR'} &= \sum_{pq \in \delta(R,R')} c_{pq}, && \forall RR' \in \tbinom{\mathcal{R}_H}{2}. \end{aligned}$$

Therefore, let $\mathcal{R}' \subset \mathcal{R}_H$ with $R_1 \in \mathcal{R}'$ and $R_2 \notin \mathcal{R}'$. We observe that this implies that $i \in R_{\mathcal{R}'}$ and $j \notin R_{\mathcal{R}'}$, since $\mathcal{R}_H$ is a partition of H. It holds that

$$\sum_{pq \in P_H} c_{pq}\,(1 - x_{pq}) + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) > \sum_{pq \in \delta(R_{\mathcal{R}'}, S_H \setminus R_{\mathcal{R}'})} c_{pq} + \sum_{pqr \in T_{\delta(R_{\mathcal{R}'}, S_H \setminus R_{\mathcal{R}'})} \cap T_H} c_{pqr}. \tag{26}$$

We evaluate the terms in (26) one-by-one, and express them as sums over elements in $\mathcal{R}'$ and $\mathcal{R}_H \setminus \mathcal{R}'$. First, we observe that for any $pq \in P_H$ we have that $x_{pq} = 0$ if and only if there exist $RR' \in \binom{\mathcal{R}_H}{2}$ such that $pq \in \delta(R, R')$. Therefore,

$$\sum_{pq \in P_H} c_{pq}\,(1 - x_{pq}) = \sum_{RR' \in \binom{\mathcal{R}_H}{2}} p_{RR'},$$

whereas

$$\sum_{pq \in \delta(R_{\mathcal{R}'}, S_H \setminus R_{\mathcal{R}'})} c_{pq} = \sum_{R \in \mathcal{R}'} \sum_{R' \in \mathcal{R}_H \setminus \mathcal{R}'} p_{RR'}.$$

For the first sum we use the decomposition

$$\tbinom{\mathcal{R}_H}{2} = \tbinom{\mathcal{R}'}{2} \cup \{RR' \mid R \in \mathcal{R}' \wedge R' \in \mathcal{R}_H \setminus \mathcal{R}'\} \cup \tbinom{\mathcal{R}_H \setminus \mathcal{R}'}{2}, \tag{27}$$

where the subsets are mutually disjoint. Consequently, it holds that

$$\sum_{pq \in P_H} c_{pq}\,(1 - x_{pq}) - \sum_{pq \in \delta(W_{\mathcal{R}'}, V_H \setminus W_{\mathcal{R}'})} c_{pq} = \sum_{RR' \in \binom{\mathcal{R}'}{2}} p_{RR'} + \sum_{RR' \in \binom{\mathcal{R}_H \setminus \mathcal{R}'}{2}} p_{RR'}. \tag{28}$$

Second, for any $pqr \in T_H$ we have that $x_{pq}x_{pr}x_{qr} = 0$ if and only if there exist $RR' \in \binom{\mathcal{R}_H}{2}$ such that $pqr \in T_{RRR'} \cup T_{RR'R'}$ or there exist $RR'R'' \in \binom{\mathcal{R}_H}{3}$ such that $pqr \in T_{RR'R''}$. Therefore,

$$\sum_{pqr \in T_H} c_{pqr}\,(1 - x_{pq}x_{pr}x_{qr}) = \sum_{RR'R'' \in \binom{\mathcal{R}_H}{3}} t_{RR'R''} + \sum_{RR' \in \binom{\mathcal{R}_H}{2}} t_{RR'},$$

whereas

$$\begin{aligned} \sum_{pqr \in T_{\delta(R_{\mathcal{R}'}, S_H \setminus R_{\mathcal{R}'})} \cap T_H} c_{pqr} = &\sum_{RR' \in \binom{\mathcal{R}'}{2}} \sum_{R'' \in \mathcal{R}_H \setminus \mathcal{R}'} t_{RR'R''} + \sum_{R \in \mathcal{R}'} \sum_{R'R'' \in \binom{\mathcal{R}_H \setminus \mathcal{R}'}{2}} t_{RR'R''} \\ &+ \sum_{R \in \mathcal{R}'} \sum_{R' \in \mathcal{R}_H \setminus \mathcal{R}'} t_{RR'}. \end{aligned}$$

For the first sum we use the decomposition

$$\tbinom{\mathcal{R}_H}{3} = \tbinom{\mathcal{R}'}{3} \cup \left\{RR'R'' \mid RR' \in \tbinom{\mathcal{R}'}{2} \wedge R'' \in \mathcal{R}_H \setminus \mathcal{R}'\right\} \cup \left\{RR'R'' \mid R \in \mathcal{R}' \wedge R'R'' \in \tbinom{\mathcal{R}_H \setminus \mathcal{R}'}{2}\right\} \cup \tbinom{\mathcal{R}_H \setminus \mathcal{R}'}{3}, \tag{29}$$

where again the subsets are mutually disjoint. By (27) and (29) , it follows that

$$\sum_{pqr \in T_H} c_{pqr}\,(1 - x_{pq}x_{pr}x_{qr}) - \sum_{pqr \in T_{\delta(R_{\mathcal{R}'}, S_H \setminus R_{\mathcal{R}'})} \cap T_H} c_{pqr} = \sum_{RR'R'' \in \binom{\mathcal{R}'}{3}} t_{RR'R''} + \sum_{RR'R'' \in \binom{\mathcal{R}_H \setminus \mathcal{R}'}{3}} t_{RR'R''}$$

$$+ \sum_{RR' \in \binom{\mathcal{R}'}{2}} t_{RR'} + \sum_{RR' \in \binom{\mathcal{R}_H \setminus \mathcal{R}'}{2}} t_{RR'}. \tag{30}$$

Combining (26), (28) and (30) yields

$$\begin{aligned}
0 &< \sum_{pq \in P_H} c_{pq}(1 - x_{pq}) + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) - \sum_{pq \in \delta(R_{\mathcal{R}'}, S_H \setminus R_{\mathcal{R}'})} c_{pq} - \sum_{pqr \in T_{\delta(R_{\mathcal{R}'}, S_H \setminus R_{\mathcal{R}'})} \cap T_H} c_{pqr} \\
&= \sum_{RR' \in \binom{\mathcal{R}'}{2}} p_{RR'} + \sum_{RR' \in \binom{\mathcal{R}_H \setminus \mathcal{R}'}{2}} p_{RR'} + \sum_{RR'R'' \in \binom{\mathcal{R}'}{3}} t_{RR'R''} + \sum_{RR'R'' \in \binom{\mathcal{R}_H \setminus \mathcal{R}'}{3}} t_{RR'R''} \\
&+ \sum_{RR' \in \binom{\mathcal{R}'}{2}} t_{RR'} + \sum_{RR' \in \binom{\mathcal{R}_H \setminus \mathcal{R}'}{2}} t_{RR'} =: S_{\mathcal{R}'}.
\end{aligned}$$

Let $k = |\mathcal{R}_H|$, and $S_{\mathcal{R}'}$ the right-hand side of the last inequality. Recall that $R_1, R_2 \in \mathcal{R}_H$, $R_1 \in \mathcal{R}'$, and $R_2 \notin \mathcal{R}'$. As $S_{\mathcal{R}'} > 0$, it follows that at least one of the sums in its definition must not be vacuous. Moreover, since its sums are indexed by pairs or triplets of subsets all belonging either to $\mathcal{R}'$ or to $\mathcal{R}_H \setminus \mathcal{R}'$, we observe that there must exist at least another subset of elements in $\mathcal{R}_H$ different from R_1 and R_2. Hence, $k \geq 3$. We calculate

$$\mathcal{S} = \sum_{\substack{\mathcal{R}' \subseteq \mathcal{R}_H : \\ R_1 \in \mathcal{R}', R_2 \notin \mathcal{R}'}} S_{\mathcal{R}'}.$$

We need this in order to contradict $\max_{x \in X_{S_H}} \phi_{c'}(x) = 0$. For any $RR' \in \binom{\mathcal{R}_H}{2} \setminus \{R_1R_2\}$, there are exactly 2^{k-3} subsets $\mathcal{R}' \subseteq \mathcal{R}_H$ such that $p_{RR'}$ or $t_{RR'}$ occurs in $S_{\mathcal{R}'}$ and $R_1 \in \mathcal{R}', R_2 \notin \mathcal{R}'$. There is no $\mathcal{R}' \subseteq \mathcal{R}_H$ such that $p_{R_1R_2}$ or $t_{R_1R_2}$ occurs in $S_{\mathcal{R}'}$ with $R_1 \in \mathcal{R}', R_2 \notin \mathcal{R}'$. For any $RR'R'' \in \binom{\mathcal{R}_H}{3} \setminus \{R_1R_2R \mid R \in \mathcal{R}_H \setminus \{R_1, R_2\}\}$, there are exactly $\lfloor 2^{k-4} \rfloor$ subsets $\mathcal{R}' \subseteq \mathcal{R}_H$ such that $t_{RR'R''}$ occurs in $S_{\mathcal{R}'}$ and $R_1 \in \mathcal{R}', R_2 \notin \mathcal{R}'$. There is no $\mathcal{R}' \subseteq \mathcal{R}_H$ such that $t_{R_1R_2R}$ occurs in $S_{\mathcal{R}'}$ for any $R \in \mathcal{R}_H \setminus \{R_1, R_2\}$ for which $R_1 \in \mathcal{R}', R_2 \notin \mathcal{R}'$. Therefore,

$$\begin{aligned}
0 < \mathcal{S} &= 2^{k-3} \sum_{RR' \in \binom{\mathcal{R}_H}{2}} p_{RR'} - 2^{k-3} p_{R_1R_2} + \lfloor 2^{k-4} \rfloor \sum_{RR'R'' \in \binom{\mathcal{R}_H}{3}} t_{RR'R''} - \lfloor 2^{k-4} \rfloor \sum_{\substack{R \in \mathcal{R}_H \\ R \notin \{R_1, R_2\}}} t_{R_1R_2R} \\
&+ 2^{k-3} \sum_{RR' \in \binom{\mathcal{R}_H}{2}} t_{RR'} - 2^{k-3} t_{R_1R_2} \\
&= 2^{k-3} \sum_{RR' \in \binom{\mathcal{R}''}{2}} p_{RR'} + \lfloor 2^{k-4} \rfloor \sum_{RR'R'' \in \binom{\mathcal{R}''}{3}} t_{RR'R''} + 2^{k-3} \sum_{RR' \in \binom{\mathcal{R}''}{2}} t_{RR'} = 2^{k-3} \phi_{c'}(x^{\mathcal{R}''}),
\end{aligned}$$

where $\mathcal{R}'' = (\mathcal{R}_H \setminus \{R_1, R_2\}) \cup \{R_1 \cup R_2\}$ is the partition obtained by merging R_1 and R_2. The last equality follows from Lemma A.1. That contradicts $\max_{x \in X_{S_H}} \phi_{c'}(x) = 0$. Therefore, this implies that there exists a subset $R \subset S_H$ with $i \in R$ and $j \in S_H \setminus R$ such that inequality (25) is fulfilled.

Let $R \subset S_H$ be a subset such that (25) holds. Therefore, we have that

$$\begin{aligned}
\phi_c(x') - \phi_c(x) &\overset{(24)}{=} \sum_{pq \in P_H} c_{pq}(1 - x_{pq}) + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) - \sum_{pq \in \delta(S_H)} c_{pq}x_{pq} \\
&- \sum_{pqr \in T_{\delta(S_H)}} c_{pqr}x_{pq}x_{pr}x_{qr} \\
&\overset{(25)}{\leq} \sum_{pq \in \delta(R, S_H \setminus R)} c_{pq} + \sum_{pqr \in T_{\delta(R, S_H \setminus R)} \cap T_H} c_{pqr} - \sum_{pq \in \delta(S_H) \cap P^-} c_{pq} \\
&- \sum_{pqr \in T_{\delta(S_H)} \cap T^-} c_{pqr} \overset{(13)}{\leq} 0.
\end{aligned}$$

Consequently, the map p is improving. By applying Corollary 2.4 we conclude the proof. □

Proof of Proposition 3.10. We define $\sigma\colon X_S \to X_S$ such that

$$\sigma(x) := \begin{cases} x & \text{if } x_{ij} = 1, \forall ij \in \binom{R}{2} \\ \left(\sigma_R \circ \sigma_{\delta(R)}\right)(x) & \text{otherwise} \end{cases}.$$

Let $x' = \sigma(x)$, for every $x \in X_S$. First, it holds that $x'_{ij} = 1$, for every $ij \in \binom{R}{2}$. Second, we show that σ is an improving map. Let $x \in X_S$ such that $x_{ij} = 1$, for all $ij \in \binom{R}{2}$. In this case, we have that $\phi_c(x') = \phi_c(x)$ by definition of x'. Now, let us consider the complementary case, i.e. let $x \in X_S$ such that there exists $ij \in \binom{R}{2}$ for which $x_{ij} = 0$. Then,

$$x'_{pq} = \begin{cases} 1 & \text{if } pq \in \binom{R}{2} \\ 0 & \text{if } pq \in \delta(R) \\ x_{pq} & \text{otherwise} \end{cases}.$$

Therefore, it follows that

$$\begin{aligned} \phi_c(x') - \phi_c(x) = & \sum_{pq \in \binom{R}{2}} c_{pq}(1 - x_{pq}) - \sum_{pq \in \delta(R)} c_{pq} x_{pq} + \sum_{pqr \in \binom{R}{3}} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) \\ & - \sum_{pqr \in T_{\delta(R)}} c_{pqr} x_{pq} x_{pr} x_{qr} \\ & \leq \max_{\substack{x \in X_S \\ x_{ij}=0}} \Big\{ \sum_{pqr \in \binom{R}{2}} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) + \sum_{pq \in \binom{R}{2}} c_{pq}(1 - x_{pq}) \Big\} \\ & - \min_{\substack{x \in X_S \\ x_{ij}=0}} \Big\{ \sum_{pqr \in T_{\delta(R)}} c_{pqr} x_{pq} x_{pr} x_{qr} + \sum_{pq \in \delta(R)} c_{pq} x_{pq} \Big\} \overset{(14)}{\leq} 0. \end{aligned}$$

This concludes the proof. □

Proof of Proposition 4.1. We start by observing that Algorithm 1 always terminates. If it returns a nontrivial partition $\mathcal{R}$, then $\mathcal{R}$ contains a subset R that satisfies (1)–(2) by construction. Therefore, let us assume that the output of Algorithm 1 is the trivial partition $\mathcal{R} = \{S\}$. If indeed there exists no nontrivial subset of S for which (1)–(2) hold, then Algorithm 1 is returning the correct output. Next, we consider the case in which there exists a nontrivial subset of S that satisfies the assumptions of Proposition 3.1, but Algorithm 1 still returns the trivial partition. We prove that this cannot happen. Let $R \subseteq S$ be a nontrivial subset of S for which (1)–(2) are satisfied. Note that such a set must exist by the assumptions of this case. Moreover we have that both R and $S \setminus R$ are not empty. Two cases can arise at this point: Algorithm 1 starts either from an element of R or from an item of $S \setminus R$. Let Algorithm 1 start sampling from R. The fact that $\mathcal{R} = \{S\}$ implies that $\exists pq \in \delta(R)$ such that $c_{pq} < 0$ or $\exists pqr \in T_{\delta(R)}$ such that $c_{pqr} < 0$ by definition of Algorithm 1. However, this is in contradiction with the assumption that R satisfies (1)–(2). Since the second scenario is symmetrical, we again reach a contradiction by applying an analogous reasoning. Therefore, we have showed that if there exists a nontrivial subset of S that fulfills (1)–(2), then Algorithm 1 finds such a subset. □

Proof of Proposition 4.2. Let $R \subseteq S$. Observe that

$$\begin{aligned} \sum_{pqr \in T_{\delta(R)}} c_{pqr} &= \sum_{pq \in \binom{R}{2}} \sum_{r \in S \setminus R} c_{pqr} + \sum_{pq \in \binom{S \setminus R}{2}} \sum_{r \in R} c_{pqr} \\ &= \frac{1}{2} \sum_{p \in R} \sum_{q \in R \setminus \{p\}} \sum_{r \in S \setminus R} c_{pqr} + \frac{1}{2} \sum_{p \in S \setminus R} \sum_{q \in S \setminus (R \cup \{p\})} \sum_{r \in R} c_{pqr} \\ &= \frac{1}{2} \sum_{p \in R} \sum_{q \in S \setminus R} \left(\sum_{r \in R \setminus \{p\}} c_{pqr} + \sum_{r \in S \setminus (R \cup \{q\})} c_{pqr} \right) = \frac{1}{2} \sum_{pq \in \delta(R)} \sum_{r \in S \setminus \{p,q\}} c_{pqr}. \end{aligned}$$

□

Proof of Proposition 4.3. Let $R \subseteq S$ such that $i \in R$ and $j \notin R, \forall j \in S_0$. We define $y \in \{0,1\}^S$ such that $y = \mathbb{1}_R$. Then, we have that $y_i = 1$ and $y_j = 0, \forall j \in S_0$. Moreover, it follows that

$$\begin{aligned}
\sum_{pq \in \delta(R)} c_{pq} &= \sum_{pq \in \binom{S}{2}} c_{pq}\left(y_p(1-y_q) + y_q(1-y_p)\right) = \sum_{pq \in \binom{S}{2}} c_{pq}\left(y_p + y_q - 2y_p y_q\right) \\
&= -2 \sum_{pq \in \binom{S}{2}} c_{pq} y_p y_q + \sum_{\substack{p,q \in S \\ p \neq q}} c_{pq} y_p \\
&= -2 \sum_{pq \in \binom{S'}{2}} c_{pq} y_p y_q - 2 \sum_{p \in S'} c_{pi} y_p + \sum_{p \in S'} \sum_{q \in S \setminus \{p\}} c_{pq} y_p + \sum_{q \in S \setminus \{i\}} c_{qi} \\
&= -2 \sum_{pq \in \binom{S'}{2}} c_{pq} y_p y_q + \sum_{p \in S'} \left(-2c_{pi} + \sum_{q \in S \setminus \{p\}} c_{pq} \right) y_p + \sum_{q \in S \setminus \{i\}} c_{qi} \\
&= \sum_{pq \in \binom{S'}{2}} c'_{pq} y_p y_q + \sum_{p \in S'} c'_p y_p + c'_\emptyset.
\end{aligned}$$

This concludes the proof. □

Proof of Proposition 5.1. Let $x \in X_S|_{x_{ij}=1}$. We show that $\phi_c(x) = \phi_{c'}(\varphi_{ij}(x))$. Let $x' = \varphi_{ij}(x)$. We use the fact that $x_{pi} = x_{pj}, \forall p \in S \setminus \{i,j\}$, and $x_{ij} = 1$. It follows that

$$\begin{aligned}
\phi_{c'}(x') &= \sum_{pqr \in \binom{S'}{3}} c'_{pqr} x'_{pq} x'_{pr} x'_{qr} + \sum_{pq \in \binom{S}{2}} c'_{pq} x'_{pq} + c'_\emptyset \\
&= \sum_{pq \in \binom{S \setminus \{i,j\}}{2}} c'_{pqi} x'_{pi} x'_{qi} x'_{pq} + \sum_{pqr \in \binom{S \setminus \{i,j\}}{3}} c'_{pqr} x'_{pq} x'_{pr} x'_{qr} + \sum_{p \in S \setminus \{i,j\}} c'_{pi} x'_{pi} + \sum_{pq \in \binom{V \setminus \{i,j\}}{2}} c'_{pq} x'_{pq} + c'_\emptyset \\
&= \sum_{pq \in \binom{S \setminus \{i,j\}}{2}} (c_{pqi} + c_{pqj}) x_{pi} x_{qi} x_{pq} + \sum_{pqr \in \binom{S \setminus \{i,j\}}{3}} c_{pqr} x_{pq} x_{pr} x_{qr} + \sum_{pq \in \binom{S \setminus \{i,j\}}{2}} c_{pq} x_{pq} \\
&+ \sum_{p \in S \setminus \{i,j\}} (c_{pi} + c_{pj} + c_{pij}) x_{pi} + c'_\emptyset \\
&= \sum_{pq \in \binom{S \setminus \{i,j\}}{2}} c_{pqi} x_{pi} x_{qi} x_{pq} + \sum_{pq \in \binom{S \setminus \{i,j\}}{2}} c_{pqj} x_{pj} x_{qj} x_{pq} + \sum_{pqr \in \binom{S \setminus \{i,j\}}{3}} c_{pqr} x_{pq} x_{pr} x_{qr} \\
&+ \sum_{p \in V} c_{pij} x_{ij} x_{pi} x_{pj} + \sum_{pq \in \binom{S \setminus \{i,j\}}{2}} c_{pq} x_{pq} + \sum_{p \in V \setminus \{i,j\}} c_{pi} x_{pi} + \sum_{p \in S \setminus \{i,j\}} c_{pj} x_{pj} + c_{ij} x_{ij} + c_\emptyset \\
&= \sum_{pqr \in \binom{S}{3}} c_{pqr} x_{pq} x_{pr} x_{qr} + \sum_{pq \in S} c_{pq} x_{pq} + c_\emptyset = \phi_c(x).
\end{aligned}$$

Therefore we have that

$$\min_{x \in X_S|_{x_{ij}=1}} \phi_c(x) = \min_{x \in X_S|_{x_{ij}=1}} \phi_{c'}(\varphi_{ij}(x)) = \min_{x \in X_{S'}} \phi_{c'}(x).$$

This concludes the proof. □

A.2. Reduction of QPBO to Min-st-Cut

Lemma A.2. *Let $S \neq \emptyset$ and $c \in \mathbb{R}^{S \cup \binom{S}{2}}$. We define $c' \in \mathbb{R}^{S \cup \binom{S}{2}}$ as $c'_p = c_p + \frac{1}{2}\sum_{q \in S \setminus \{p\}} c_{pq}$, for every $p \in S$, $c'_{pq} = -\frac{1}{2} c_{pq}$, for every $pq \in \binom{S}{2}$. Then, for any $y \in \{0,1\}^V$ we have that*

$$\sum_{pq \in \binom{S}{2}} c_{pq} y_p y_q + \sum_{p \in S} c_p y_p = \sum_{p \in S} \sum_{q \in S \setminus \{p\}} c'_{pq} y_p (1 - x_q) + \sum_{\substack{p \in S \\ c'_p > 0}} c'_p y_p - \sum_{\substack{p \in S \\ c'_p < 0}} c'_p (1 - y_p) + \sum_{\substack{p \in S \\ c'_p < 0}} c'_p.$$

Proof. Let $y \in \{0,1\}^S$. We have that

$$\begin{aligned}\sum_{pq\in\binom{S}{2}} c_{pq}y_py_q + \sum_{p\in S} c_py_p &= \frac{1}{2}\sum_{p\in S}\sum_{q\in S\setminus\{p\}} c_{pq}y_py_q + \sum_{p\in V} c_py_p \\ &= -\frac{1}{2}\sum_{p\in S}\sum_{q\in S\setminus\{p\}} c_{pq}y_p(1-y_q) + \frac{1}{2}\sum_{p\in S}\sum_{q\in S\setminus\{p\}} c_{pq}y_p + \sum_{p\in S} c_py_p \\ &= -\frac{1}{2}\sum_{p\in S}\sum_{q\in S\setminus\{p\}} c_{pq}y_p(1-y_q) + \sum_{p\in S}\left(c_p + \frac{1}{2}\sum_{q\in S\setminus\{p\}} c_{pq}\right)y_p \\ &= \sum_{p\in S}\sum_{q\in S\setminus\{p\}} c'_{pq}y_p(1-y_q) + \sum_{p\in V} c'_py_p = \sum_{p\in S}\sum_{q\in S\setminus\{p\}} c'_{pq}y_p(1-y_q) \\ &+ \sum_{\substack{p\in V\\ c'_p>0}} c'_py_p - \sum_{\substack{p\in S\\ c'_p<0}} c'_p(1-y_p) + \sum_{\substack{p\in S\\ c'_p<0}} c'_p.\end{aligned}$$

We therefore reach the thesis. □

We reduce this problem to solving an instance of min-st-cut. If $c_{pq} \le 0$, $pq \in \binom{S}{2}$, and therefore $c'_{pq} \ge 0$, $\forall pq \in \binom{S}{2}$, in Lemma A.2 the resulting instance can be solved efficiently.

Proposition A.3. *Let $S \neq \emptyset$ and $c \in \mathbb{R}^{S\cup\binom{S}{2}}$. We define $\phi_c\colon \{0,1\}^S \to \mathbb{R}$ such that for all $y \in \{0,1\}^S$ it holds that*

$$\phi_c(y) = \sum_{p\in S}\sum_{q\in S\setminus\{p\}} c_{pq}y_p(1-y_q) + \sum_{\substack{p\in S\\ c_p>0}} c_py_p - \sum_{\substack{p\in S\\ c_p<0}} c_p(1-y_p).$$

Furthermore, we define $S' = S \cup \{s,t\}$, $P' \subset S' \times S'$ such that

$$\begin{aligned}(s,p) \in P' &\Leftrightarrow c_p < 0, \quad \forall p \in S,\\ (p,t) \in P' &\Leftrightarrow c_p > 0, \quad \forall p \in S,\\ (p,q) \in P' \wedge (q,p) \in P', &\quad \forall pq \in \tbinom{S}{2},\end{aligned}$$

and $c' \in \mathbb{R}^{P'}$ such that

$$\begin{aligned}c'_{(s,p)} &= -c_p, \ \forall (s,p) \in P',\\ c'_{(p,t)} &= c_p, \quad \forall (p,t) \in P',\\ c'_{(p,q)} = c'_{(q,p)} &= c_{pq}, \quad \forall pq \in \tbinom{S}{2}.\end{aligned}$$

Moreover, we define the function $\varphi_{c'} : \{0,1\}^{S'} \to \mathbb{R}$ such that for all $y \in \{0,1\}^{S'}$ it holds that

$$\varphi_{c'}(y) = \sum_{(p,q)\in P'} c'_{(p,q)}y_p(1-y_q).$$

Then we have that

$$\min_{x\in\{0,1\}^S} \phi_c(x) = \min_{\substack{y\in\{0,1\}^{S'}\\ y_s=1\\ y_t=0}} \varphi_{c'}(y).$$

Proof. First, the map $\chi\colon \{0,1\}^S \to \{y \in \{0,1\}^{S'} \mid y_s = 1 \wedge y_t = 0\}$ such that $\chi(y)_s = 1$, $\chi(y)_t = 0$ and $\chi(y)_p = y_p$ for any $p \in S$ is bijective. Second, for any $y \in \{0,1\}^S$ it holds that

$$\varphi_{c'}(\chi(y)) = \sum_{(p,q)\in P'} c'_{pq}\chi(y)_p(1-\chi(y)_q) = \sum_{p\in S}\sum_{q\in S\setminus\{p\}} c'_{(p,q)}y_p(1-y_q) + \sum_{\substack{p\in S\\ c_p>0}} c'_{(p,t)}y_p + \sum_{\substack{p\in S\\ c_p<0}} c'_{(s,p)}(1-y_p)$$

$$= \sum_{p \in S} \sum_{q \in V \setminus \{p\}} c_{pq} y_p (1 - y_q) + \sum_{\substack{p \in S \\ c_p > 0}} c_p y_p - \sum_{\substack{p \in S \\ c_p < 0}} c_p (1 - y_p) = \phi_c(y)$$

This concludes the proof. □