Partial Optimality in Cubic Correlation Clustering for General Graphs

David Stein David.Stein1@tu-dresden.de
Faculty of Computer Science, TU Dresden, 01069 Dresden, Germany

Bjoern Andres bjoern.andres@tu-dresden.de
Faculty of Computer Science, TU Dresden, 01069 Dresden, Germany
Center for Scalable Data Analytics and AI Dresden/Leipzig, Dresden, Germany

Silvia Di Gregorio digregorio@lipn.fr
Université Sorbonne Paris Nord, LIPN, CNRS UMR 7030, F-93430 Villetaneuse, France
Faculty of Computer Science, TU Dresden, 01069 Dresden, Germany

Abstract

The higher-order correlation clustering problem for a graph G and costs associated with cliques of G consists in finding a clustering of G so as to minimize the sum of the costs of those cliques whose nodes all belong to the same cluster. To tackle this NP-hard problem in practice, local search heuristics have been proposed and studied in the context of applications. Here, we establish partial optimality conditions for cubic correlation clustering, i.e., for the special case of at most 3-cliques. We define and implement algorithms for deciding these conditions and examine their effectiveness numerically, on two data sets.

Keywords: cubic correlation clustering, cubic clique partitioning, partial optimality, discrete optimization, data reduction

1 Introduction

We study an optimization problem whose feasible solutions are all clusterings of a graph that is given as input. Here, a clustering of a graph $G = (V, E)$ is a partition Π of the node set V such that for any $U \in \Pi$, the subgraph $G[U]$ of G induced by U is connected. In addition to the edges, we are given a subset T of node triples pqr such that $pq \in E$ and $qr \in E$ and $pr \in E$. For any $e \in E$ and any $t \in T$, we are given numbers $c_e, c_t \in \mathbb{R}$ called costs. The objective is to find a clustering Π of G so as to minimize the sum of the costs of those edges and triples whose nodes all belong to the same cluster. More formally:

Definition 1 *For any graph $G = (V, E)$, any $T \subseteq \{pqr \in \binom{V}{3} \mid pq \in E, qr \in E, pr \in E\}$ and any $c\colon T \cup E \cup \{\emptyset\} \to \mathbb{R}$, the instance of the* cubic correlation clustering *problem is*

$$\min_{\Pi \in P_G} \quad c_\emptyset + \sum_{R \in \Pi} \sum_{e \in E \cap \binom{R}{2}} c_e + \sum_{R \in \Pi} \sum_{t \in T \cap \binom{R}{3}} c_t \quad .$$

Here, P_G denotes the set of all clusterings of G.

The cubic correlation clustering problem is NP-hard, as it generalizes the NP-hard clique partitioning problem for complete graphs (Grötschel and Wakabayashi, 1989), specializing to the latter in the case that $c_t = 0$ for all $t \in T$. Potential applications of cubic correlation clustering include the tasks of fitting equilateral triangles to points in a plane, the task of

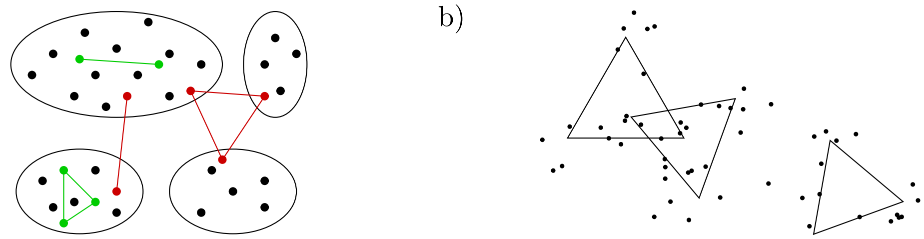

Figure 1: In order to examine the effectiveness of partial optimality conditions empirically, we implement algorithms for deciding these conditions and measure the fraction of fixed variables with respect to a parameter controlling the noise of the problem, for (a) synthetic instances with four clusters and noisy costs, and (b) instances for the task of finding equilateral triangles in a noisy point cloud.

fitting an unknown number of planes containing the origin to a set of points in 3-dimensional space, and the task of fitting an unknown number of lines to a set of points in 2-dimensional space, as discussed, e.g., by Levinkov et al. (2022).

In this article, we ask whether we can compute a partial solution to the problem efficiently, i.e., decide efficiently for some edges or triples whether their elements are in the same cluster or distinct clusters of an optimal clustering. 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 empirically, we implement algorithms for deciding these conditions, and conduct experiments, cf. Figure 1.

This article is an extension of a conference article (Stein et al., 2023) about partial optimality conditions for the cubic correlation clustering problem with respect to a complete graph. Here, we extend the conference article in three ways: Firstly, we transfer all conditions by Stein et al. (2023) from complete to general graphs. Secondly, we generalize the implementation of Stein et al. (2023) from complete to general graphs. Thirdly, we evaluate the effectiveness of these conditions on instances defined with respect to general (also incomplete) graphs. These extensions make the algorithms for partial optimality applicable to larger instances with incomplete graphs. The complete source code of the algorithms and for reproducing the experiments is available at `https://github.com/dsteindd/partial-optimality-in-cubic-correlation-clustering-for-general-graphs`.

2 Related Work

We choose to state the cubic correlation clustering problem (Definition 1) in the form of a non-linear binary program (Proposition 2), 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 applications as accurate mathematical abstractions of intrinsically non-linear tasks (Agarwal et al., 2005; Kappes et al., 2016; Kim et al., 2014; Levinkov et al., 2022; Ochs and Brox, 2012; Purkait et al., 2017).

Like Kim et al. (2014), we introduce costs only for node triples $t \in T$ such that $G[t]$ is complete. This means that we do not introduce costs for node triples for which the induced subgraph $G[t]$ is incomplete (possibly even disconnected). Encoding whether three nodes that do not induce a complete subgraph belong to the same cluster would require additional variables and constraints (Horňáková et al., 2017) that generalize the correlation clustering problem in a different manner than a cubic objective function does.

Being able to efficiently fix some variables to an optimal value and thus reduce 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 and Sutter, 1992; Hammer et al., 1984; Kappes et al., 2013; Kohli et al., 2008; Shekhovtsov, 2014; Shekhovtsov et al., 2015). In particular, we mention here the impressive 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 discrete 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 and 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 correlation clustering problem we consider here generalizes the specialization to complete graphs of both the correlation clustering problem and the multicut problem. We transfer all partial optimality conditions established by Alush and Goldberger (2012) and Lange et al. (2018, 2019) for the correlation clustering problem and the multicut problem to the cubic correlation clustering problem. In addition, we establish new results. Unlike Lange et al. (2018), we do not contribute partial optimality conditions for the max cut problem.

3 Preliminaries

In order to establish partial optimality conditions for the cubic correlation clustering problem (Definition 1), we state this problem in the form of the non-linear integer program introduced by Kim et al. (2014):

Proposition 2 *The instance of the cubic correlation clustering problem (*CP3*) with respect to a graph $G = (V, E)$, a set $T \subseteq \{pqr \in \binom{V}{3} \mid pq \in E, pr \in E, qr \in E\}$ and a function $c\colon T \cup E \cup \{\emptyset\} \to \mathbb{R}$ has the form of the cubic integer program*

$$\min_{x\colon E\to\{0,1\}} \quad c_\emptyset + \sum_{pq\in E} c_{pq}\, x_{pq} + \sum_{pqr\in T} c_{pqr}\, x_{pq}\, x_{pr}\, x_{qr} \tag{1}$$

$$\text{s.t.}: \quad \forall (V_C, E_C) \in \mathrm{cycles}(G)\ \forall e' \in E_C\colon \quad \sum_{e\in E_C\setminus\{e'\}} x_e - x_{e'} \le |E_C| - 2\ , \tag{2}$$

where cycles(G) *is the set containing all cycles in G.*

Proof For each graph partition $\Pi \in P_G$ and every edge $pq \in E$, 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_G of all graph partitions of G and the feasible set X_G of all $x : E \to \{0, 1\}$ that satisfy the above inequalities (2), see (Chopra and Rao, 1993). Under this bijection, the objective functions of Definition 1 and Proposition 2 are equivalent. ■

Below, let CP3(G, c) denote this instance of the problem, ϕ_c its objective function, and X_G its feasible set, i.e., the set of all $x : E \to \{0, 1\}$ that satisfy the inequalities (2).

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

Definition 3 *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 4 *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 $x^* \in Q$.*

Proof 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 $\sigma(x^*) \in Q$. ■

Corollary 5 *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 $\sigma_{\delta(U)}$ that cuts a set $U \subseteq V$ from its complement, and the map σ_U that joins all sets intersecting with a set $U \subseteq V$:

Definition 6 *For any node set V and $U \subseteq V$, the* elementary cut map $\sigma_{\delta(U)}\colon X_G \to X_G$ *is such that for all $x \in X_G$ and all $pq \in E$:*

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

Lemma 7 *For any node set V and $U \subseteq V$, the map $\sigma_{\delta(U)}$ is well-defined, i.e., for all $x \in X_G$ we have that $\sigma_{\delta(U)}(x) \in X_G$.*

Proof Let $x \in X_G$ and let $x' = \sigma_{\delta(U)}(x)$. For the sake of contradiction, we assume that $x' \notin X_G$. Thus, there is a cycle $(V_C, E_C) \in \text{cycles}(G)$ and an edge $e' \in E_C$ such that $\sum_{e\in E_C\setminus\{e'\}} x'_e - x'_{e'} > |E_C| - 2$. This is only possible if $x'_e = 1$ for all $e \in E_C \setminus \{e'\}$ and $x'_{e'} = 0$. By definition of x', we have that $x'_e = x_e = 1$ for all $e \in E_C \setminus \{e'\}$ and all $e \in E_C \setminus \{e'\}$ are not cut by U. For e' we distinguish two cases:

(1) If e' is not cut by U, we have that $x'_{e'} = x_{e'}$. It follows that $0 = x'_{e'} = x_{e'}$. Thus $\sum_{e\in E_C\setminus\{e'\}} x_e - x_{e'} > |E_C| - 2$. This is a violation of $x \in X_G$.

(2) If $|e' \cap U| = 1$, namely e' is cut by U, then e' is the only edge in E_C that has exactly one element in common with U. Therefore (V_C, E_C) cannot be a cycle, which is a contradiction. ■

Definition 8 *For any node sets V and $U \subseteq V$, the* elementary join map $\sigma_U : X_G \to X_G$ *is such that for all $x \in X_G$ and all $pq \in E$:*

$$\sigma_U(x)_{pq} := \begin{cases} 1 & \text{if } pq \in \binom{U}{2} \cap E \\ 1 & \text{if there is a } pq\text{-path } (V_P, E_P) \text{ such that } \forall e \in E_P\colon x_e = 1 \lor e \in \binom{U}{2} \cap E \\ x_{pq} & \text{otherwise} \end{cases}$$

Lemma 9 *For any node set V and $U \subseteq V$, the map σ_U is well-defined, i.e., for all $x \in X_G$ we have $\sigma_U(x) \in X_G$.*

Proof Let $x \in X_G$ and let $x' = \sigma_U(x)$. For the sake of contradiction, we assume that $x' \notin X_G$. Thus there is a cycle $(V_C, E_C) \in \text{cycles}(G)$ and an edge $f \in E_C$ such that

$$\sum_{e \in E_C \setminus \{f\}} x'_e - x'_f > |E_C| - 2 \ . \tag{3}$$

Therefore $x'_e = 1$ for all $e \in E_C \setminus \{f\}$ and $x'_f = 0$. Moreover $x_f = 0$ since $x' \geq x$ by definition of σ_U.

For any $E' \subseteq E$, let $E_{0\to1}(E') := \{e \in E' \mid x_e = 0, x'_e = 1\} \setminus \binom{U}{2}$. Since $x'_f = 0$ we have that $f \notin E_{0\to1}(E_C)$. Next we show that $E_{0\to1}(E_C) = \emptyset$. For the sake of contradiction, let us assume that $E_{0\to1}(E_C) \neq \emptyset$. We pick a cycle $(V_C, E_C) \in \text{cycles}(G)$ with minimal $|E_{0\to1}(E_C)|$ and that fulfills (3). We proceed to construct a different cycle $(V_{C'}, E_{C'}) \in \text{cycles}(G)$ such that $|E_{0\to1}(E_{C'})| = |E_{0\to1}(E_C)| - 1$ that still fulfills (3), which would lead to a contradiction. Let $uv \in E_{0\to1}(E_C)$ be any arbitrary edge. We know that $uv \in E_C \setminus \{f\}$. Therefore $x_{uv} = 0$ and $x'_{uv} = 1$ by definition of $E_{0\to1}(E_C)$. Then, according to the definition of σ_U, there is a uv-path (V_P, E_P) such that $x_e = 1$ for all $e \in E_P \setminus \binom{U}{2}$. Let $(V'_C, E'_C) \in \text{cycles}(G)$ be the symmetric difference of the two cycles (V_C, E_C) and $(V_P, E_P \cup \{uv\})$, i.e., $E_{C'} = (E_C \setminus (E_P \cup \{uv\})) \cup ((E_P \cup \{uv\}) \setminus E_C)$ and $V_{C'} = \{v \in V \mid \exists e \in E_{C'}\colon v \in e\}$. As $E_{0\to1}(E_P \cup \{uv\}) = \{uv\}$, we have that $E_{0\to1}((E_P \cup \{uv\}) \setminus E_C) = \emptyset$. Next, $E_{0\to1}(E_{C'}) = E_{0\to1}(E_C \setminus (E_P \cup \{uv\})) \cup E_{0\to1}((E_P \cup \{uv\}) \setminus E_C) = E_{0\to1}(E_C \setminus (E_P \cup \{uv\})) = E_{0\to1}(E_C) \setminus \{uv\}$. The last equality follows from the fact that $x_e = 1$ for all $e \in E_P \setminus \binom{U}{2}$. Therefore $|E_{0\to1}(E_{C'})| = |E_{0\to1}(E_C)| - 1$ which contradicts minimality of (V_C, E_C) with respect to $|E_{0\to1}(E_C)|$. Thus we must have $E_{0\to1}(E_C) = \emptyset$.

Since $E_{0\to1}(E_C) = \emptyset$ and (3) holds, we have that $x_e = 1$ for all $e \in E_C \setminus \binom{U}{2}$. Thus $(V_C, E_C \setminus \{f\})$ is a path connecting the endpoints of f such that $x_e = 1$ for all $e \in (E_C \setminus \{f\}) \setminus \binom{U}{2}$. Therefore $\sum_{e \in E_C \setminus \{f\}} x_e - x_f = |E_C| - 1 > |E_C| - 2$, which contradicts to $x \in X_G$. ■

4 Partial Optimality Conditions

In this section, we establish partial optimality conditions for the cubic correlation clustering problem by constructing improving maps, starting from the elementary maps $\sigma_{\delta(U)}$ and σ_U defined in Section 3. 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'. For any $U, U', U'' \subseteq V$, let

$$\delta(U, U') := \left\{pq \in E \mid p \in U \land q \in U'\right\} \quad , \quad \delta(U) := \delta(U, V \setminus U),$$
$$T_{UU'U''} := \left\{pqr \in T \mid p \in U \land q \in U' \land r \in U''\right\} .$$

For $I(G) := T \cup E \cup \{\emptyset\}$ and any $c\colon I(G) \to \mathbb{R}$, let

$$E^{\pm} := \{pq \in E \mid c_{pq} \gtreqless 0\} \quad , \quad T^{\pm} := \{pqr \in T \mid c_{pqr} \gtreqless 0\} \, ,$$
$$T_{E'} := \left\{pqr \in T \mid E' \cap \tbinom{pqr}{2} \neq \emptyset\right\} \quad \forall E' \subseteq E \ .$$

For any subset $V_H \subseteq V$, let $E_H = E \cap \binom{V_H}{2}, T_H = T \cap \binom{V_H}{3}$.

4.1 Cut Conditions

Here, we establish partial optimality conditions that imply the existence of an optimal solution x^* to CP3(G, c) such that $x^*_{ij} = 0$ for some $ij \in E$ or $x^* \in \{x \in X_G \mid x_{ij}x_{ik}x_{jk} = 0\}$ for some $ijk \in T$. The following Proposition 10 generalizes Theorem 1 of Alush and Goldberger (2012) to cubic objective functions.

Proposition 10 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. If there exists $U \subseteq V$ such that*

$$c_{pq} \geq 0 \quad \forall pq \in \delta(U) \tag{4}$$
$$c_{pqr} \geq 0 \quad \forall pqr \in T_{\delta(U)} \tag{5}$$

then there is an optimal solution x^* to CP3(G, c) *such that $x^*_{ij} = 0$ for all $ij \in \delta(U)$.*

Proof We define $\sigma\colon X_G \to X_G$ such that for all $x \in X_G$ we have

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

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

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

The last equality is due to the fact that those sums vanish by Assumptions (4) and (5). Applying Corollary 5 concludes the proof. ■

This condition can be exploited: When satisfied for a subset U, CP3(G, c) decomposes into two independent subproblems. Firstly,

$$\min_{x \in X_G} \phi_c(x) = \min_{x \in X_{G[U]}} \phi_{c|_{I(G[U])}}(x) + \min_{x \in X_{G[V \setminus U]}} \phi_{c|_{I(G[V \setminus U])}}(x) \ ,$$

where we denote the subgraph of G induced by U by $G[U]$, for any $U \subseteq V$. Secondly, given solutions

$$x' \in \operatorname*{argmin}_{x \in X_{G[U]}} \phi_{c|_{I(G[U])}}(x), \quad x'' \in \operatorname*{argmin}_{x \in X_{G[V \setminus U]}} \phi_{c|_{I(G[V \setminus U])}}(x) \ ,$$

an optimal solution to the problem $\min_{x \in X_G} \phi_c(x)$ is given by the vector $x \in X_G$ such that

$$x_{pq} = \begin{cases} x'_{pq} & \text{if } pq \in E \cap \binom{U}{2} \\ x''_{pq} & \text{if } pq \in E \cap \binom{V \setminus U}{2} \\ 0 & \text{if } pq \in \delta(U) \end{cases} \ .$$

The following Proposition 11, together with Proposition 14 further below, generalizes Theorem 1 of Lange et al. (2019) to cubic objective functions.

Proposition 11 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $ij \in E$. If there exists $U \subseteq V$ such that $ij \in \delta(U)$ and*

$$c^+_{ij} \geq \sum_{pqr \in T_{\delta(U)}} c^-_{pqr} + \sum_{pq \in \delta(U)} c^-_{pq} \ , \tag{6}$$

then there is an optimal solution x^* to CP3(G, c) *such that $x^*_{ij} = 0$.*

Proof Let $\sigma \colon X_G \to X_G$ be constructed as

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

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

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

The last inequality follows from Assumption (6). We conclude the proof by applying Corollary 5. ∎

The following Proposition 12 establishes a partial optimality condition that implies the existence of an optimal solution x^* such that $x^*_{ij} x^*_{ik} x^*_{jk} = 0$ for some $ijk \in T$. Note that one cannot conclude which of the variables x^*_{ij}, x^*_{ik} or x^*_{jk} equal zero.

Proposition 12 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $ijk \in T$ and $U \subseteq V$ such that $ij, ik \in \delta(U)$. If*

$$c^+_{ijk} + c^+_{ij} + c^+_{ik} \geq \sum_{pqr \in T_{\delta(U)}} c^-_{pqr} + \sum_{pq \in \delta(U)} c^-_{pq} \ , \tag{7}$$

there is an optimal solution x^* to CP3(G, c) *such that $x^*_{ij} x^*_{ik} x^*_{jk} = 0$.*

Proof We define $\sigma \colon X_G \to X_G$ as

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

For any $x \in X_G$, we denote $\sigma(x)$ by x'. Firstly, we observe that $x'_{ij} x'_{ik} x'_{jk} = 0$ for all $x \in X_G$. Secondly, for any $x \in X_G$ such that $x_{ij} x_{ik} x_{jk} = 1$, we have

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

The last inequality holds by Assumption (7). Applying Proposition 4 with $Q = \{x \in X_G \mid x_{ij} x_{ik} x_{jk} = 0\}$ concludes the proof. ∎

4.2 Join Conditions

Next, we establish partial optimality conditions that imply the existence of an optimal solution x^* to CP3(G, c) such that $x^*_{ij} = 1$ for some $ij \in E$. This property can be applied to simplify a given instance by joining the nodes i and j. Below, Proposition 14 transfers a result of Lange et al. (2019) to the cubic correlation clustering problem. Lemma 13 is an auxiliary lemma applied in the proof of Proposition 14.

Lemma 13 *Let $G = (V, E)$ be a graph, $c \in \mathbb{R}^{I(G)}$, $U \subseteq V$ and $ij \in \delta(U)$. Moreover, let $x \in X_G$ such that $x_{ij} = 0$ and $\bar{x} = \left(\sigma_{\{i,j\}} \circ \sigma_{\delta(U)}\right)(x)$. Then, $\bar{x}_{pq} = x_{pq}$ for all $pq \notin \delta(U)$.*

Proof Let $\hat{x} = \sigma_{\delta(U)}(x)$. We have that $\hat{x}_{pq} = 0$ for all $pq \in \delta(U)$, and $\hat{x}_{pq} = x_{pq}$ for all $pq \notin \delta(U)$. Let us assume that there exists some $pq \notin \delta(U)$ such that $\bar{x}_{pq} \neq x_{pq}$. First, consider the case where $x_{pq} = \hat{x}_{pq} = 0$ and $\bar{x}_{pq} = 1$. Then, there must be a path from p to i and from q to j such that the edge variables along this path in $\hat{x}$ all have value 1. Analogously, we could have a path from p to j and from q to i such that the edge variables along these paths in $\hat{x}$ all have value 1. Without loss of generality, we only consider the first case. Since at least one of these paths must contain an edge $e \in \delta(U) \setminus \{ij\}$, it follows that $\hat{x}_e$ has value equal to zero. This leads to a contradiction. Secondly, consider the case where $x_{pq} = \hat{x}_{pq} = 1$ and $\bar{x}_{pq} = 0$. We observe that this case is impossible, as the elementary join map never sets to 0 a variable that was previously equal to 1 by definition. ∎

Proposition 14 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $ij \in E$. If there exists $U \subseteq V$ such that $ij \in \delta(U)$ and*

$$2c^-_{ij} + \sum_{pqr \in T_{\{ij\}}} c^-_{pqr} \geq \sum_{pqr \in T_{\delta(U)}} |c_{pqr}| + \sum_{pq \in \delta(U)} |c_{pq}| \ , \tag{8}$$

there is an optimal solution x^* to CP3(G, c) *such that $x^*_{ij} = 1$.*

Proof Let $\sigma \colon X_G \to X_G$ such that for all $x \in X_G$, we have

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

For any $x \in X_G$, let $x' = \sigma(x)$. The map σ is such that $x'_{ij} = 1$ for all $x \in X_G$. We show that σ is improving. In particular, let $x \in X_G$ such that $x_{ij} = 0$. By Lemma 13 we have that $x'_{pq} = x_{pq}$ for all $pq \notin \delta(U)$. Therefore,

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

The last inequality is due to Assumption (8). ∎

While Proposition 14 is about deciding whether two adjacent nodes i, j should end up together, Proposition 15 asks the same question for a triple ijk.

Proposition 15 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $ijk \in T$ and $U \subseteq V$ such that $ij, ik \in \delta(U)$. If*

$$2c^-_{ijk} + 2c^-_{ij} + 2c^-_{ik} + c^-_{jk} - \sum_{pqr\in T} c^+_{pqr} - \sum_{pq\in E} c^+_{pq} + \min_{\substack{x\in X_{G[\{i,j,k\}]}:\\ x_{ij}x_{ik}x_{jk}=0}} \sum_{pq\in\binom{ijk}{2}} c_{pq}x_{pq} \geq \sum_{pqr\in T_{\delta(U)}} c^-_{pqr} + \sum_{pq\in\delta(U)} c^-_{pq} \ , \quad (9)$$

there is an optimal solution x^* to CP3(G, c) *such that $x^*_{ij}x^*_{ik}x^*_{jk} = 1$.*

Proof We construct $\sigma\colon X_G \to X_G$ as follows.

$$\sigma(x) := \begin{cases} x & \text{if } x_{ij}x_{ik}x_{jk} = 1 \\ \left(\sigma_{\{i,j,k\}} \circ \sigma_{\delta(U)}\right)(x) & \text{otherwise} \end{cases}$$

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

In order to see this, let $pq \notin \delta(U)$. We observe that $\sigma_{\delta(U)}(x)_{pq} = x_{pq}$ and that $\sigma_{\{i,j,k\}}(x)_{pq} \geq x_{pq}$ by definitions of $\sigma_{\delta(U)}$ and $\sigma_{\{i,j,k\}}$. Moreover, we make the following observations: For every $pqr \in T_{\delta(U)} \setminus \{ijk\}$ we have that $c_{pqr}\left(x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr}\right) \leq |c_{pqr}|$ since $x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr} \in \{-1, 0, +1\}$. For every $pq \in \delta(U) \setminus \{ij, ik\}$ we have that $c_{pq}\left(x'_{pq} - x'_{pq}\right) \leq |c_{pq}|$ since $x'_{pq} - x_{pq} \in \{-1, 0, 1\}$. For every $pqr \notin T_{\delta(U)}$ we have that $c_{pqr}\left(x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr}\right) \leq \max(0, c_{pqr}) = c^+_{pqr}$ since $x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr} \in \{0, 1\}$. For every $pq \notin \delta(U) \cup \{jk\}$ we have that $c_{pq}\left(x'_{pq} - x_{pq}\right) \leq \max(0, c_{pq}) = c^+_{pq}$ since $x'_{pq} - x_{pq} \in \{0, 1\}$. It follows that

$$\begin{aligned}
\phi_c(x') - \phi_c(x) = &\sum_{\substack{pqr\in T_{\delta(U)}:\\ pqr\neq ijk}} c_{pqr}(x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr}) + c_{ijk} + \sum_{pqr\notin T_{\delta(U)}} c_{pqr}(x'_{pq}x'_{pr}x'_{qr} - x_{pq}x_{pr}x_{qr}) \\
&+ \sum_{pq\in\{ij,ik,jk\}} c_{pq}(1 - x_{pq}) + \sum_{\substack{pq\in\delta(U):\\ pq\notin\{ij,ik\}}} c_{pq}(x'_{pq} - x_{pq}) + \sum_{pq\notin\delta(U)\cup\{jk\}} c_{pq}(x'_{pq} - x_{pq}) \\
\leq\ & c_{ijk} + \max_{\substack{x\in X_{G[\{i,j,k\}]}:\\ x_{ij}x_{ik}x_{jk}=0}} \sum_{pq\in\binom{ijk}{2}} c_{pq}(1 - x_{pq}) + \sum_{\substack{pqr\in T_{\delta(U)}:\\ pqr\neq ijk}} |c_{pqr}| + \sum_{\substack{pqr\in T^+:\\ pqr\notin T_{\delta(U)}}} c_{pqr} + \sum_{\substack{pq\in\delta(U):\\ pq\notin\{ij,ik\}}} |c_{pq}| + \sum_{\substack{pq\in E^+:\\ pq\notin(\delta(U)\cup\{jk\})}} c_{pq} \\
=\ & c_{ijk} + \max_{\substack{x\in X_{G[\{i,j,k\}]}:\\ x_{ij}x_{ik}x_{jk}=0}} \sum_{pq\in\binom{ijk}{2}} c_{pq}(1 - x_{pq}) + \sum_{\substack{pqr\in T^+\cup T_{\delta(U)}:\\ pqr\neq ijk}} |c_{pqr}| + \sum_{\substack{pq\in E^+\cup\delta(U):\\ pq\notin\{ij,ik,jk\}}} |c_{pq}| \\
=\ & -2c^-_{ijk} - 2c^-_{ij} - 2c^-_{ik} - c^-_{jk} - \min_{\substack{x\in X_{G[\{i,j,k\}]}:\\ x_{ij}x_{ik}x_{jk}=0}} \sum_{pq\in\binom{ijk}{2}} c_{pq}x_{pq} + \sum_{pqr\in T_{\delta(U)}} c^-_{pqr} + \sum_{pqr\in T} c^+_{pqr} + \sum_{pq\in E} c^+_{pq} + \sum_{pq\in\delta(U)} c^-_{pq} \leq 0.
\end{aligned}$$

Assumption (9) provides the last inequality. We arrive at the thesis by applying Proposition 4 with $Q = \{x \in X_G \mid x_{ij}x_{ik}x_{jk} = 1\}$. ■

The following Proposition 16 expands Corollary 1 of Lange et al. (2019) to the cubic correlation clustering problem.

Proposition 16 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $ijk \in T$ and $U, U' \subseteq V$ such that $ij, ik \in \delta(U)$ and $jk, ik \in \delta(U')$. If all of the following conditions hold, there exists an optimal solution x^* to* CP3(G, c) *such that $x^*_{ik} = 1$.*

$$c^-_{ijk} + 2c^-_{ij} + 2c^-_{ik} + \sum_{pqr\in T_{\{ij,ik\}}} c^-_{pqr} \geq \sum_{pqr\in T_{\delta(U)}} |c_{pqr}| + \sum_{pq\in\delta(U)} |c_{pq}| \quad (10)$$

$$c^-_{ijk} + 2c^-_{jk} + 2c^-_{ik} + \sum_{pqr \in T_{\{jk,ik\}}} c^-_{pqr} \geq \sum_{pqr \in T_{\delta(U')}} |c_{pqr}| + \sum_{pq \in \delta(U')} |c_{pq}| \tag{11}$$

$$c_{ijk} + c_{ij} + c_{ik} + c_{jk} \leq - \sum_{\substack{pqr \in T_{\delta(ijk)} \cap T^-: \\ pqr \notin T_{\{ij,ik,jk\}}}} |c_{pqr}| - \sum_{pq \in \delta(ijk) \cap E^-} |c_{pq}| \tag{12}$$

Proof Let $\sigma\colon X_G \to X_G$ be defined as

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

We use the notation $x' = \sigma(x)$ for all $x \in X_G$. Firstly, the map σ is such that $x'_{ik} = 1$ for all $x \in X_G$. Secondly, we consider $x \in X_G$ such that $x_{ik} = x_{ij} = 0$ and $x_{jk} = 1$. The map $\sigma_{\{i,k\}} \circ \sigma_{\delta(U)}$ is such that $x'_{ij} = x'_{ik} = x'_{jk} = 1$ and $x'_{pq} = x_{pq}$ for any $pq \notin \delta(U)$. The difference of the objective values for such $x \in X_G$ is given by

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

The last inequality follows from Assumption (10). Thirdly, we consider $x \in X_G$ such that $x_{ik} = x_{jk} = 0$ and $x_{ij} = 1$. The map $\sigma_{\{i,k\}} \circ \sigma_{\delta(U')}$ is improving by analogous arguments and Assumption (11). Finally, we consider $x \in X_G$ such that $x_{ik} = x_{jk} = x_{ij} = 0$. The map $\sigma_{\{i,j,k\}} \circ \sigma_{\delta(ijk)}$ is such that

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

Therefore,

$$\phi_c(x') - \phi_c(x) = c_{ijk} + c_{ij} + c_{ik} + c_{jk} - \sum_{pqr \in T_{\delta(ijk)} \setminus T_{\{ij,ik,jk\}}} c_{pqr} x_{pq} x_{pr} x_{qr} - \sum_{pq \in \delta(ijk)} c_{pq} x_{pq}$$

$$\leq c_{ijk} + c_{ij} + c_{ik} + c_{jk} + \sum_{\substack{pqr \in T_{\delta(ijk)} \cap T^-: \\ pqr \notin T_{\{ij,ik,jk\}}}} |c_{pqr}| + \sum_{pq \in \delta(ijk) \cap E^-} |c_{pq}| \overset{(12)}{\leq} 0 \ .$$

Applying Corollary 5 concludes the proof. ■

Next, we discuss a generalization of Theorem 2 of Lange et al. (2019). To this end, let $V_H \subseteq V$, $E_H = E \cap \binom{V_H}{2}$ and $T_H = T \cap \binom{V_H}{3}$. We define $c' \in \mathbb{R}^{I(G[V_H])}$ by the equations written below.

$$c'_\emptyset = \frac{1}{2} \sum_{pqr \in T_H} c_{pqr} + \sum_{pq \in E_H} c_{pq} \tag{13}$$

$$\forall pq \in E_H\colon \quad c'_{pq} = -c_{pq} + \frac{1}{2} \sum_{\substack{r \in V_H \setminus \{p,q\}: \\ pqr \in T_H}} c_{pqr} \tag{14}$$

$$\forall pqr \in T_H\colon \quad c'_{pqr} = -2c_{pqr} \tag{15}$$

Proposition 18 studies subsets $V_H \subseteq V$ such that $\mathbb{1}_{E_H}$ is a trivial solution to the problem $\max_{x \in X_{G[V_H]}} \phi_{c'}(x)$, i.e. $\max_{x \in X_{G[V_H]}} \phi_{c'}(x) = 0$, since

$$\begin{aligned}
\phi_{c'}\left(\mathbb{1}_{E_H}\right) &= \sum_{pqr \in T_H} c'_{pqr} + \sum_{pq \in E_H} c'_{pq} + c'_\emptyset \\
&= -2 \sum_{pqr \in T_H} c_{pqr} + \sum_{pq \in E_H} \Big(-c_{pq} + \frac{1}{2} \sum_{\substack{r \in V_H: \\ pqr \in T_H}} c_{pqr} \Big) + \frac{1}{2} \sum_{pqr \in T_H} c_{pqr} + \sum_{pq \in E_H} c_{pq} \\
&= -\frac{3}{2} \sum_{pqr \in T_H} c_{pqr} + \frac{1}{2} \sum_{pq \in E_H} \sum_{\substack{r \in V_H: \\ pqr \in T_H}} c_{pqr} = -\frac{3}{2} \sum_{pqr \in T_H} c_{pqr} + \frac{3}{2} \sum_{pqr \in T_H} c_{pqr} = 0 \ .
\end{aligned}$$

Lemma 17 is an auxiliary lemma applied in the proof of Proposition 18.

Lemma 17 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Define $c' \in \mathbb{R}^{I(G)}$ as in* (13), (14), (15) *for $V_H = V$, $E_H = E$ and $T_H = T$. Then for any graph partition $\Pi \in P_G$:*

$$\phi_{c'}(x^\Pi) = \frac{1}{2} \sum_{pqr \in T} c_{pqr} \prod_{uv \in \binom{pqr}{2}} (1 - x^\Pi_{uv}) + \sum_{pqr \in T} c_{pqr} \sum_{uv \in \binom{pqr}{2}} x^\Pi_{uv} \prod_{\substack{u'v' \in \binom{pqr}{2} \\ u'v' \neq uv}} (1 - x^\Pi_{u'v'}) + \sum_{pq \in E} c_{pq} (1 - x^\Pi_{pq}) \tag{16}$$

$$= \frac{1}{2} \sum_{UU'U'' \in \binom{\Pi}{3}} \sum_{pqr \in T_{UU'U''}} c_{pqr} + \sum_{UU' \in \binom{\Pi}{2}} \sum_{pq \in \delta(U,U')} c_{pq} + \sum_{UU' \in \binom{\Pi}{2}} \Big(\sum_{pqr \in T_{UUU'}} c_{pqr} + \sum_{pqr \in T_{UU'U'}} c_{pqr} \Big) \ , \tag{17}$$

where x^Π denotes the feasible vector corresponding to the graph partition Π.

Proof We use the fact that for any graph partition $\Pi \in P_G$ and any $pqr \in T$ we have that $x^\Pi_{pq} x^\Pi_{qr} = x^\Pi_{pq} x^\Pi_{pr} = x^\Pi_{pr} x^\Pi_{qr} = x^\Pi_{pq} x^\Pi_{pr} x^\Pi_{qr}$ thanks to the transitivity property, i.e. given a 3-clique either the three nodes are all apart, or only two are together, or they are all together. Expanding the inner products and the inner sums leads to $\prod_{uv \in \binom{pqr}{2}} \left(1 - x^\Pi_{uv}\right) = 1 - x^\Pi_{pq} - x^\Pi_{pr} - x^\Pi_{qr} + 2x^\Pi_{pq} x^\Pi_{pr} x^\Pi_{qr}$, and $\sum_{uv \in \binom{pqr}{2}} x^\Pi_{uv} \prod_{\substack{u'v' \in \binom{pqr}{2}: \\ u'v' \notin \{uv\}}} (1 - x^\Pi_{u'v'}) = x^\Pi_{pq} + x^\Pi_{pr} + x^\Pi_{qr} - 3x^\Pi_{pq} x^\Pi_{qr} x^\Pi_{pr}$. By substituting and collecting terms, we conclude the proof for Equation (16).

Equation (17) follows instead from the following observations:

$$\prod_{ab \in \binom{pqr}{3}} \left(1 - x^{\Pi}_{ab}\right) = 1 \quad \Leftrightarrow \quad \exists\, UU'U'' \in \tbinom{\Pi}{3} \colon pqr \in T_{UU'U''}$$

$$\sum_{ab \in \binom{pqr}{2}} x^{\Pi}_{ab} \prod_{a'b' \in \binom{pqr}{2} \setminus \{ab\}} (1 - x^{\Pi}_{a'b'}) = 1 \quad \Leftrightarrow \quad \exists\, UU' \in \tbinom{\Pi}{2} \colon\ (pqr \in T_{UUU'} \vee pqr \in T_{UU'U'})$$

$$1 - x^{\Pi}_{pq} = 1 \quad \Leftrightarrow \quad \exists\, UU' \in \tbinom{\Pi}{2} \colon pq \in \delta(U, U')\ .$$

This concluses the proof. ■

Proposition 18 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $V_H \subseteq V$, $E_H = E \cap \binom{V_H}{2}$, $T_H = T \cap \binom{V_H}{3}$ and $ij \in E_H$. Define $c' \in \mathbb{R}^{I(G[V_H])}$ as in* (13), (14), (15)*. If $\max_{x \in X_{G[V_H]}} \phi_{c'}(x) = 0$ and for all $U \subseteq V_H$ with $i \in U$ and $j \in V_H \setminus U$ we have*

$$\sum_{pq \in \delta(V_H) \cap E^-} c_{pq} + \sum_{pqr \in T_{\delta(V_H)} \cap T^-} c_{pqr} \geq \sum_{pq \in \delta(U, V_H \setminus U)} c_{pq} + \sum_{pqr \in T_{\delta(U, V_H \setminus U)} \cap T_H} c_{pqr}\ , \tag{18}$$

then there exists an optimal solution x^* to CP3(G, c) *such that $x^*_{ij} = 1$.*

Proof We define $\sigma \colon X_G \to X_G$ as

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

Let $x' = \sigma(x)$ for any $x \in X_G$. It is easy to see that $x'_{ij} = 1$ for all $x \in X_G$. Similarly to before, we show that σ is an improving map. For any $x \in X_G$ such that $x_{ij} = 1$ we have $\phi_c(x') - \phi_c(x) = 0$, by definition of x'. Now, we consider $x \in X_G$ such that $x_{ij} = 0$. We let $x|_{E_H}$ denote the restriction of x containing only components corresponding to edges in E_H. Let Π be the graph partition of V such that $x = x^{\Pi}$, and let Π_H be the induced graph partition of V_H such that $x|_{E_H} = x^{\Pi_H}$. Since $x_{ij} = 0$, there exist $U_1, U_2 \in \Pi_H$ such that $i \in U_1$, $j \in U_2$. We have for all $pq \in E$:

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

Therefore, it follows that

$$\phi_c(x') - \phi_c(x) = \sum_{pq \in E_H} c_{pq}(1 - x_{pq}) + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) - \sum_{pq \in \delta(V_H)} c_{pq}x_{pq} - \sum_{pqr \in T_{\delta(V_H)}} c_{pqr}x_{pq}x_{pr}x_{qr}. \tag{19}$$

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

$$\sum_{pq \in E_H} c_{pq}\,(1 - x_{pq}) + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) \leq \sum_{pq \in \delta(U, V_H \setminus U)} c_{pq} + \sum_{pqr \in T_{\delta(U, V_H \setminus U)} \cap T_H} c_{pqr}. \tag{20}$$

For the sake of contradiction, we assume that there is no such $U \subset V_H$. For any $\Pi' \subset \Pi_H$, let $R_{\Pi'} = \bigcup_{P' \in \Pi'} P'$. Furthermore, we define $t \colon \binom{\Pi_H}{2} \cup \binom{\Pi_H}{3} \to \mathbb{R}$ and $e \colon \binom{\Pi_H}{2} \to \mathbb{R}$ such that

$$t_{UU'U''} = \sum_{pqr \in T_{UU'U''}} c_{pqr}\ \forall UU'U'' \in \tbinom{\Pi_H}{3}, \qquad t_{UU'} = \sum_{pqr \in T_{UUU'} \cup T_{UU'U'}} c_{pqr}\ \forall UU' \in \tbinom{\Pi_H}{2}$$

$$e_{UU'} = \sum_{pq\in\delta(U,U')} c_{pq} \; \forall UU' \in \tbinom{\Pi_H}{2}.$$

Therefore, let $\Pi' \subset \Pi_H$ with $U_1 \in \Pi'$ and $U_2 \notin \Pi'$. We observe that this implies that $i \in R_{\Pi'}$ and $j \notin R_{\Pi'}$, since Π_H is a partition of V_H. By the assumption (20) of the proof by contradiction, we have that

$$\sum_{pq\in E_H} c_{pq}\,(1-x_{pq}) + \sum_{pqr\in T_H} c_{pqr}(1-x_{pq}x_{pr}x_{qr}) > \sum_{pq\in\delta(R_{\Pi'},V_H\setminus R_{\Pi'})} c_{pq} + \sum_{pqr\in T_{\delta(R_{\Pi'},V_H\setminus R_{\Pi'})}\cap T_H} c_{pqr}\;. \tag{21}$$

We evaluate the terms in (21) one-by-one, and express them as sums over elements in Π' and $\Pi_H\setminus\Pi'$. Firstly, we observe that for any $pq\in E_H$ we have $x_{pq}=0$ if and only if there exist $UU'\in\binom{\Pi_H}{2}$ such that $pq\in\delta(U,U')$. Therefore, $\sum_{pq\in E_H} c_{pq}\,(1-x_{pq}) = \sum_{UU'\in\binom{\Pi_H}{2}} e_{UU'}$ whereas $\sum_{pq\in\delta(R_{\Pi'},V_H\setminus R_{\Pi'})} c_{pq} = \sum_{U\in\Pi'}\sum_{U'\in\Pi_H\setminus\Pi'} e_{UU'}$. For the first sum, we use the decomposition $\binom{\Pi_H}{2} = \binom{\Pi'}{2} \cup \{UU' \mid U\in\Pi' \wedge U'\in\Pi_H\setminus\Pi'\} \cup \binom{\Pi_H\setminus\Pi'}{2}$, where the subsets are mutually disjoint. Consequently:

$$\sum_{pq\in E_H} c_{pq}\,(1-x_{pq}) - \sum_{pq\in\delta(R_{\Pi'},V_H\setminus R_{\Pi'})} c_{pq} = \sum_{UU'\in\binom{\Pi'}{2}} e_{UU'} + \sum_{UU'\in\binom{\Pi_H\setminus\Pi'}{2}} e_{UU'}\;. \tag{22}$$

Secondly, for any $pqr\in T_H$, we have $x_{pq}x_{pr}x_{qr}=0$ if and only if there exist $UU'\in\binom{\Pi_H}{2}$ such that $pqr\in T_{UUU'}\cup T_{UU'U'}$ or there exist $UU'U''\in\binom{\Pi_H}{3}$ such that $pqr\in T_{UU'U''}$. Therefore,

$$\sum_{pqr\in T_H} c_{pqr}\,(1-x_{pq}x_{pr}x_{qr}) = \sum_{UU'U''\in\binom{\Pi_H}{3}} t_{UU'U''} + \sum_{UU'\in\binom{\Pi_H}{2}} t_{UU'} \tag{23}$$

whereas

$$\sum_{\substack{pqr\in T_{\delta(R_{\Pi'},V_H\setminus R_{\Pi'})}\\ pqr\in T_H}} c_{pqr} = \sum_{UU'\in\binom{\Pi'}{2}}\sum_{\substack{U''\in\Pi_H\\ U''\notin\Pi'}} t_{UU'U''} + \sum_{U\in\Pi'}\sum_{U'U''\in\binom{\Pi_H\setminus\Pi'}{2}} t_{UU'U''} + \sum_{U\in\Pi'}\sum_{\substack{U'\in\Pi_H\\ U'\notin\Pi'}} t_{UU'}. \tag{24}$$

For the first sum, we use the decomposition

$$\tbinom{\Pi_H}{3} = \tbinom{\Pi'}{3} \cup \left\{UU'U'' \mid UU'\in\tbinom{\Pi'}{2} \wedge U''\in\Pi_H\setminus\Pi'\right\} \cup \left\{UU'U'' \mid U\in\Pi' \wedge U'U''\in\tbinom{\Pi_H\setminus\Pi'}{2}\right\} \cup \tbinom{\Pi_H\setminus\Pi'}{3} \tag{25}$$

where again the subsets are mutually disjoint. By (23) together with the decomposition (25), and (24) it follows that

$$\begin{aligned}&\sum_{pqr\in T_H} c_{pqr}\,(1-x_{pq}x_{pr}x_{qr}) - \sum_{pqr\in T_{\delta(R_{\Pi'},V_H\setminus R_{\Pi'})}\cap T_H} c_{pqr}\\ &= \sum_{UU'U''\in\binom{\Pi'}{3}} t_{UU'U''} + \sum_{UU'U''\in\binom{\Pi_H\setminus\Pi'}{3}} t_{UU'U''} + \sum_{UU'\in\binom{\Pi'}{2}} t_{UU'} + \sum_{UU'\in\binom{\Pi_H\setminus\Pi'}{2}} t_{UU'}\;.\end{aligned} \tag{26}$$

Combining (21), (22) and (26) yields

$$\begin{aligned}0 &< \sum_{pq\in E_H} c_{pq}(1-x_{pq}) + \sum_{pqr\in T_H} c_{pqr}(1-x_{pq}x_{pr}x_{qr}) - \sum_{pq\in\delta(R_{\Pi'},V_H\setminus R_{\Pi'})} c_{pq} - \sum_{pqr\in T_{\delta(R_{\Pi'},V_H\setminus R_{\Pi'})}\cap T_H} c_{pqr}\\ &= \sum_{UU'\in\binom{\Pi'}{2}} e_{UU'} + \sum_{UU'\in\binom{\Pi_H\setminus\Pi'}{2}} e_{UU'} + \sum_{UU'U''\in\binom{\Pi'}{3}} t_{UU'U''} + \sum_{UU'U''\in\binom{\Pi_H\setminus\Pi'}{3}} t_{UU'U''}\end{aligned}$$

$$+ \sum_{UU' \in \binom{\Pi'}{2}} t_{UU'} + \sum_{UU' \in \binom{\Pi_H \setminus \Pi'}{2}} t_{UU'} =: W_{\Pi'} \ . \tag{27}$$

Let $k = |\Pi_H|$ be the number of components in the partition Π_H, and $W_{\Pi'}$ the right-hand side of the last inequality. Recall that $U_1, U_2 \in \Pi_H$, $U_1 \in \Pi'$, and $U_2 \notin \Pi'$. As $W_{\Pi'} > 0$, 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 triples of subsets all belonging either to Π' or to $\Pi_H \setminus \Pi'$, we observe that there must exist at least another subset of nodes in Π_H different from U_1 and U_2. In fact, each sum is indexed over at least two subsets of Π_H that are either both of them in Π' or in $\Pi_H \setminus \Pi'$. Hence, at least one set between Π' and $\Pi_H \setminus \Pi'$ has two elements. Given also that neither Π' nor $\Pi \setminus \Pi'$ can be empty, it follows that $k \geq 3$.

We calculate

$$W = \sum_{\substack{\Pi' \subseteq \Pi_H : \\ U_1 \in \Pi', U_2 \notin \Pi'}} W_{\Pi'}. \tag{28}$$

We need this in order to contradict $\max_{x \in X_{G[V_H]}} \phi_{c'}(x) = 0$.

First, we consider the special case $k = 3$. Without loss of generality, let $\Pi_H = \{U_1, U_2, U_3\}$. We observe that there are only two subsets of Π_H that can serve as Π' in the sum (28) are $\Pi'_1 = \{U_1\}$ and $\Pi'_2 = \{U_1, U_3\}$. We have that $W_{\Pi'_1} = e_{U_2U_3} + t_{U_2U_3}$, $W_{\Pi'_2} = e_{U_1U_3} + t_{U_1U_3}$. Therefore, we arrive at $0 < W = e_{U_1U_3} + e_{U_2U_3} + t_{U_1U_3} + t_{U_2U_3} = \phi_{c'}(x^{\Pi''})$, where $\Pi'' = (\Pi_H \setminus \{U_1, U_2\}) \cup \{U_1 \cup U_2\}$ is the partition obtained by merging U_1 and U_2. The last equality follows from Lemma 17. That contradicts $\max_{x \in X_{G[V_H]}} \phi_{c'}(x) = 0$.

Next, we consider the case $k \geq 4$. We evaluate the values of all the sums in (27). We start by looking at the possible values $e_{UU'}$ and $t_{UU'}$. We observe that there are four different scenarios for their subscripts.

1. For any $UU' \in \binom{\Pi_H}{2}$ and $\{U, U'\} \cap \{U_1, U_2\} = \emptyset$, there are exactly 2^{k-3} subsets $\Pi' \subseteq \Pi_H$ such that $e_{UU'}$ or $t_{UU'}$ occurs in $W_{\Pi'}$ and $U_1 \in \Pi', U_2 \notin \Pi'$, namely 2^{k-4} subsets Π' for which we have additionally $U, U' \in \Pi'$ plus 2^{k-4} subsets Π' for which we have additionally $U, U' \notin \Pi'$.

2. For any $U \in \Pi_H \setminus \{U_1, U_2\}$ there are exactly 2^{k-3} subsets $\Pi' \subseteq \Pi_H$ such that e_{U_1U} or t_{U_1U} occurs in $W_{\Pi'}$ and $U_1 \in \Pi', U_2 \notin \Pi'$, namely the 2^{k-3} subsets Π' for which we additionally have $U \in \Pi'$.

3. For any $U \in \Pi_H \setminus \{U_1, U_2\}$ there are exactly 2^{k-3} subsets $\Pi' \subseteq \Pi_H$ such that e_{U_2U} or t_{U_2U} occurs in $W_{\Pi'}$ and $U_1 \in \Pi', U_2 \notin \Pi'$, namely the 2^{k-3} subsets Π' for which we additionally have $U \notin \Pi'$.

4. There is no $\Pi' \subseteq \Pi_H$ such that $e_{U_1U_2}$ or $t_{U_1U_2}$ occurs in $W_{\Pi'}$ with $U_1 \in \Pi', U_2 \notin \Pi'$.

We now turn to the remaining sums in (27), namely the ones that have three subsets as indices. We observe that for these there are again four possibilities.

1. For any $UU'U'' \in \binom{\Pi_H}{3}$ and $\{U, U', U''\} \cap \{U_1, U_2\} = \emptyset$, there are exactly 2^{k-4} subsets $\Pi' \subseteq \Pi_H$ such that $t_{UU'U''}$ occurs in $W_{\Pi'}$ and $U_1 \in \Pi', U_2 \notin \Pi'$, namely 2^{k-5} subsets Π' for which we have additionally $U, U', U'' \in \Pi'$ plus 2^{k-5} subsets Π' for which we have additionally $U, U', U'' \notin \Pi'$.

2. For any $UU' \in \binom{\Pi_H}{2}$ and $\{U, U'\} \cap \{U_1, U_2\} = \emptyset$, there are exactly 2^{k-4} subsets $\Pi' \subseteq \Pi_H$ such that $t_{U_1 UU'}$ occurs in $W_{\Pi'}$ and $U_1 \in \Pi', U_2 \notin \Pi'$, namely the 2^{k-4} subsets Π' for which we have additionally $U, U' \in \Pi'$.
3. For any $UU' \in \binom{\Pi_H}{2}$ and $\{U, U'\} \cap \{U_1, U_2\} = \emptyset$, there are exactly 2^{k-4} subsets $\Pi' \subseteq \Pi_H$ such that $t_{U_2 UU'}$ occurs in $W_{\Pi'}$ and $U_1 \in \Pi', U_2 \notin \Pi'$, namely the 2^{k-4} subsets Π' for which we have additionally $U, U' \notin \Pi'$.
4. There is no $\Pi' \subseteq \Pi_H$ such that $t_{U_1 U_2 U}$ occurs in $W_{\Pi'}$ for any $U \in \Pi_H \setminus \{U_1, U_2\}$ for which $U_1 \in \Pi', U_2 \notin \Pi'$.

At this point, we put all of this together in (28):

$$\begin{aligned}
0 < W = &\sum_{\substack{\Pi' \subseteq \Pi_H: \\ U_1 \in \Pi', U_2 \notin \Pi'}} W_{\Pi'} = 2^{k-3} \sum_{UU' \in \binom{\Pi_H}{2}} e_{UU'} - 2^{k-3} e_{U_1 U_2} + 2^{k-4} \sum_{UU'U'' \in \binom{\Pi_H}{3}} t_{UU'U''} \\
&- 2^{k-4} \sum_{\substack{U \in \Pi_H: \\ U \notin \{U_1, U_2\}}} t_{U_1 U_2 U} + 2^{k-3} \sum_{UU' \in \binom{\Pi_H}{2}} t_{UU'} - 2^{k-3} t_{U_1 U_2} \\
&= 2^{k-3} \sum_{UU' \in \binom{\Pi''}{2}} e_{UU'} + 2^{k-4} \sum_{UU'U'' \in \binom{\Pi''}{3}} t_{UU'U''} + 2^{k-3} \sum_{UU' \in \binom{\Pi''}{2}} t_{UU'} = 2^{k-3} \phi_{c'}(x^{\Pi''}) \ ,
\end{aligned}$$

where $\Pi'' = (\Pi_H \setminus \{U_1, U_2\}) \cup \{U_1 \cup U_2\}$ is the partition obtained by merging U_1 and U_2. The last equality follows from Lemma 17. This contradicts $\max_{x \in X_{G[V_H]}} \phi_{c'}(x) = 0$. Therefore, this implies that there exists a subset $U \subset V_H$ with $i \in U$ and $j \in V_H \setminus U$ such that inequality (20) is fulfilled.

Let $U \subset V_H$ be a subset for which (20) holds. Therefore, we have that

$$\begin{aligned}
\phi_c(x') - \phi_c(x) &\overset{(19)}{=} \sum_{pq \in E_H} c_{pq}(1 - x_{pq}) + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) - \sum_{pq \in \delta(V_H)} c_{pq}x_{pq} - \sum_{pqr \in T_{\delta(V_H)}} c_{pqr}x_{pq}x_{pr}x_{qr} \\
&\overset{(20)}{\leq} \sum_{pq \in \delta(U, V_H \setminus U)} c_{pq} + \sum_{pqr \in T_{\delta(U, V_H \setminus U)} \cap T_H} c_{pqr} - \sum_{pq \in \delta(V_H) \cap E^-} c_{pq} - \sum_{pqr \in T_{\delta(V_H)} \cap T^-} c_{pqr} \overset{(18)}{\leq} 0.
\end{aligned}$$

Consequently, the map σ is improving. By applying Corollary 5, we conclude the proof. ■

We are unaware of an efficient method for finding subsets $V_H \subseteq V$ and $U \subseteq V$ for which the conditions (18) are satisfied. For subsets V_H with $|V_H| \in \{2, 3\}$, two corollaries of Proposition 18 provide efficiently-verifiable partial optimality conditions:

Corollary 19 *Let $G = (V, E)$ a graph, $c \in \mathbb{R}^{I(G)}$ and $ij \in E$. If $c_{ij} \leq \sum_{pq \in \delta(ij) \cap E^-} c_{pq} + \sum_{pqr \in T_{\delta(ij)} \cap T^-} c_{pqr}$, then there exists an optimal solution x^* to* CP3(G, c) *such that $x^*_{ij} = 1$.*

Corollary 20 *Let $G = (V, E)$ a graph, $c \in \mathbb{R}^{I(G)}$ and $ijk \in T$. If*

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

then there exists an optimal solution x^* to CP3(G, c) *such that $x^*_{ik} = 1$.*

We now present the last main partial optimality condition of this article.

Proposition 21 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $V_H \subseteq V$, $E_H = E \cap \binom{V_H}{2}$ and $T_H = T \cap \binom{V_H}{3}$. If for every $ij \in E_H$ we have*

$$\max_{\substack{x \in X_G:\\ x_{ij}=0}} \Big\{ \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) + \sum_{pq \in E_H} c_{pq}(1 - x_{pq}) \Big\} \leq \min_{\substack{x \in X_G:\\ x_{ij}=0}} \Big\{ \sum_{pqr \in T_{\delta(V_H)}} c_{pqr}x_{pq}x_{pr}x_{qr} + \sum_{pq \in \delta(V_H)} c_{pq}x_{pq} \Big\} \tag{29}$$

then there is an optimal solution x^* to CP3(G, c) *such that $x^*_{ij} = 1$, $\forall ij \in E_H$.*

Proof We define $\sigma \colon X_G \to X_G$ such that

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

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

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

Therefore, it follows that

$$\begin{aligned} \phi_c(x') - \phi_c(x) &= \sum_{pq \in E_H} c_{pq}(1 - x_{pq}) - \sum_{pq \in \delta(V_H)} c_{pq}x_{pq} + \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) - \sum_{pqr \in T_{\delta(V_H)}} c_{pqr}x_{pq}x_{pr}x_{qr} \\ &\leq \max_{\substack{x \in X_G:\\ x_{ij}=0}} \Big\{ \sum_{pqr \in T_H} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) + \sum_{pq \in E_H} c_{pq}(1 - x_{pq}) \Big\} - \min_{\substack{x \in X_G:\\ x_{ij}=0}} \Big\{ \sum_{pqr \in T_{\delta(V_H)}} c_{pqr}x_{pq}x_{pr}x_{qr} + \sum_{pq \in \delta(V_H)} c_{pq}x_{pq} \Big\} \\ &\overset{(29)}{\leq} 0. \end{aligned}$$

■

We are unaware of an efficient method for deciding (29) for arbitrary subsets $V_H \subseteq V$ and costs $c \in \mathbb{R}^{I(G)}$. Yet, Corollary 22 below describes one setting in which a suitable subset can be searched for heuristically, in polynomial time.

Corollary 22 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $V_H \subseteq V$, $E_H = E \cap \binom{V_H}{2}$ and $T_H = T \cap \binom{V_H}{3}$. If*

$$c_{pq} \leq 0 \qquad \forall pq \in E_H, \tag{30}$$

$$c_{pqr} \leq 0 \qquad \forall pqr \in T_H, \tag{31}$$

$$\max_{\substack{U \subset V_H:\\ U \neq \emptyset}} \Big\{ \sum_{pqr \in T_{\delta(U)} \cap T_H} c_{pqr} + \sum_{pq \in \delta(U, V_H \setminus U)} c_{pq} \Big\} \leq \sum_{pqr \in T_{\delta(V_H)} \cap T^-} c_{pqr} + \sum_{pq \in \delta(V_H) \cap E^-} c_{pq} \ , \tag{32}$$

then there is an optimal solution x^* to CP3(G, c) *such that $x^*_{ij} = 1$, $\forall ij \in E_H$.*

Proof The claim follows from the following three observations. Let (30) and (31) be satisfied. Firstly, for any $ij \in E_H$, we have that the left-hand side of (29) is equal to $\max_{\substack{U \subsetneq V_H:\\ i \in U,\, j \notin U}} \sum_{pqr \in T_{\delta(U)} \cap T_H} c_{pqr} + \sum_{pq \in \delta(U, V_H \setminus U)} c_{pq}$. This means that the maximizer is given by a feasible $x \in X_G$ whose restriction to E_H corresponds to a partition Π of V_H into two subsets. To see this, note that for any $ij \in E_H$, instead of maximizing the left-hand side of (29) over $x \in X_G$ such that $x_{ij} = 0$ we can equivalently maximize over all $x' \in X_{G[V_H]}$ such that $x'_{ij} = 0$. Now, let us assume that there exists $ij \in E_H$ such that the maximizer is given by a feasible $x' \in X_{G[V_H]}$ corresponding to a partition Π' of V_H into more than two subsets. Without loss of generality, let $U_1, U_2 \in \Pi'$ such that $i \in U_1$, $j \in U_2$. Then, the vector $x'' \in X_{G[V_H]}$ corresponding to the partition $\Pi'' = \left\{U_1, \bigcup_{U \in \Pi' \setminus \{U_1\}} U\right\}$ has objective value at least the objective value of x'. This follows from the facts that all the costs are non-positive and that Π' is a refinement of Π''. Secondly, by using the trivial lower bound of the right-hand side of (29) we have that it is at least $\sum_{pqr \in T_{\delta(V_H)} \cap T^-} c_{pqr} + \sum_{pqr \in \delta(V_H) \cap E^-} c_{pq}$. Thirdly, for any $ij \in E_H$ we have

$$\max_{\substack{U \subset V_H:\\ U \neq \emptyset}} \Bigg\{ \sum_{pqr \in T_{\delta(U)} \cap T_H} c_{pqr} + \sum_{pq \in \delta(U, V_H \setminus U)} c_{pq} \Bigg\} \geq \max_{\substack{U \subsetneq V_H:\\ i \in U,\, j \notin U}} \sum_{pqr \in T_{\delta(U)} \cap T_H} c_{pqr} + \sum_{pq \in \delta(U, V_H \setminus U)} c_{pq} \ .$$

These three observations, together with satisfaction of (30)–(32), imply that (29) holds for every $ij \in E_H$. Hence, Corollary 22 specializes Proposition 21 and the claim follows. ■

5 Deciding Partial Optimality Efficiently

Next, we describe algorithms for deciding the partial optimality conditions introduced in the previous section. This includes exact algorithms and heuristics, and we discuss their time complexity. We denote by $\deg G = \max_{p \in V} |\{q \in V \setminus \{p\} \mid pq \in E\}|$, and $\deg^{(3)} G = \max_{pq \in E} |\{r \in V \setminus \{p, q\} \mid pqr \in T\}|$, the maximum number of adjacent vertices of a vertex or the maximum number of triples an edge is a part of.

We start by examining Proposition 10. Observe that it suffices to find any subset $U \subseteq V$ that fulfills Proposition 10. Because of that, we focus our attention on minimal subsets, with respect to set inclusion.

Proposition 23 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. The minimal subsets $U \subseteq V$ with respect to set inclusion that satisfy (4)–(5) are exactly the vertex sets of the connected components of the graph $G' = (V, E')$ with $E' = \{pq \in E | c_{pq} < 0 \vee \exists r \in V \text{ s.t. } pqr \in T \colon c_{pqr} < 0\}$.*

Proof First we show that every $U \subseteq V$ inducing a connected component in G' is such that (4)–(5) are satisfied. Let us assume that there is some $U \subseteq V$ inducing a connected component in G', but (4)–(5) do not hold. In this case, there is some $pq \in \delta(U)$ such that $c_{pq} < 0$ or $pqr \in T_{\delta(U)}$ such that $c_{pqr} < 0$. Without loss of generality, let us assume that $p \notin U$ and $q \in U$ in both cases. Then we have that $pq \in E'$ by the definition of E'. Since $U \cup \{p\}$ also induces a connected subgraph of G', this contradicts the assumption that U induces a (maximal) connected component of G'.

Second we show that every minimal $U \subseteq V$ with respect to set inclusion that satisfies (4)–(5) induces a connected component of G'. For the sake of contradiction, let us assume that U does not induce a connected component of G'. By definition of E', we have that $\delta(U) \cap E' = \emptyset$, since U satisfies (4)–(5). Therefore every connected component of the subgraph $G'[U]$ is also a connected component of the underlying graph G'. If $G'[U]$ has exactly one connected component, then U induces a connected component of G', which contradicts our assumption. We now consider the case that the subgraph $G'[U]$ has more than one connected components.

Without loss of generality, let $U' \subset U$ be a vertex set inducing a connected component in $G'[U]$ and thus also in G'. For all edges $pq \in \delta(U')$ we have that $pq \notin E'$ because otherwise U' would not induce a connected component of G'. Therefore for all $pq \in \delta(U')$ we have that $c_{pq} > 0$ and $c_{pqr} > 0$ for all $pqr \in T$. Since $T_{\delta(U')}$ is the union of all sets $\{pqr \in T \mid pq \in \delta(U')\}$, we have that U' fulfills (4)–(5), which contradicts the minimality of U with respect to set inclusion. ∎

Hence, we can use any algorithm that computes connected components in a graph. In this work we use a disjoint set data structure with path compression and union-by-size, starting with at most $|V|$ singleton sets, proceeding with at most $|E|$ called of union, followed by precisely $|V|$ calls of find. This has time complexity $\mathcal{O}((|V| + |E|)\alpha(|V|))$ (Tarjan and van Leeuwen, 1984) where α denotes the inverse Ackermann function.

Partial optimality according to Propositions 11–16 is conditional to the existence of an edge $ij \in E$ or a triple $ijk \in T$, together with a subset $U \subseteq V$ (in case of Proposition 16 even a second subset $U' \subseteq V$ independent of U) for which specific conditions are satisfied, namely (6)–(12). For every triple, we decide (12) explicitly by enumerating all edges incident to one of the triple vertices and then enumerating all triples this edge is a part of, in time $\mathcal{O}(\deg G \cdot \deg^{(3)} G)$. For every pair or triple, we reduce the search of subsets U or U' satisfying (6)–(11) *with maximum margin* to the min *st*-cut problem (in Appendix A). In order to decide partial optimality efficiently, we solve the dual max *st*-flow problems by the push-relabel algorithm with FIFO vertex selection rule (Goldberg and Tarjan, 1988). Note that any subsets U, or U', that satisfy (6)–(11) would suffice. We choose to use an exact method that returns subsets U or U' that fulfill (6)–(11) with maximum margin.

As mentioned already in Section 4.2, we are unaware of an efficient method for finding subsets that satisfy the conditions of Proposition 18 or 21. Regarding Proposition 18, we resort to the special case of Corollary 19 that we test for each pair in time $\mathcal{O}(\deg G \cdot \deg^{(3)} G)$, and to the special case of Corollary 20 that we test for each triple in time $\mathcal{O}(\deg G \cdot \deg^{(3)} G)$. Regarding Proposition 21, we employ the special case of Corollary 22 and search heuristically for a witness U of (32), as follows: In an outer loop, we iterate over all pairs $ij \in E$ with $c_{ij} \leq 0$, and initialize $U = \{i, j\}$. For each of these initializations of U, we add elements to U for which the costs of all pairs and triples inside U is non-positive. Upon termination of this inner loop, we take U to be a candidate. By construction of U, all coefficients on the left-hand side of (32) are non-positive. By applying Proposition 25 to the left-hand side of (32), this problem takes the form of a min cut problem with non-negative capacities that we solve exactly using the algorithm by Stoer and Wagner (1997). As soon as we find a candidate U that fulfills (32), we merge the nodes in U as described in Section 6.2 and obtain a smaller instance of CP3. Then, we proceed as described in Section 6.3.

6 Combining Partial Optimality Conditions

Next, we discuss how we apply partial optimality conditions iteratively and why this requires special attention. Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Furthermore, let $Q_1, Q_2 \subseteq X_G$. If there is an optimal solution $x_1^* \in X_G$ to CP3(G, c) such that $x_1^* \in Q_1$, and an optimal solution $x_2^* \in X_G$ to CP3(G, c) such that $x_2^* \in Q_2$, then there is not necessarily an optimal solution $x^* \in X_G$ to CP3(G, c) such that $x^* \in Q_1 \cap Q_2$. Thus, we cannot naïvely apply the partial optimality conditions simultaneously (see Example 1).

Example 1 *Let* $V = \{1,2,3\}$*,* $E = \{12,13,23\}$*,* $T = \{123\}$ *and* $c \in \mathbb{R}^{I(G)}$ *such that* $c_{123} = 5$*,* $c_{12} = c_{13} = c_{23} = -2$*,* $c_\emptyset = 0$*. Then,* $\min_{x \in X_G} \phi_c(x) = -2$*. Furthermore, let* $Q_1 = \{x \in X_G \mid x_{12} = 1\}$ *and* $Q_2 = \{x \in X_G \mid x_{13} = 1\}$*. It follows that* $Q_1 \cap Q_2 = \{x \in X_G \mid x_{12} = 1, x_{13} = 1\} = \{(1,1,1)\}$*. The set of optimal solutions is the set of all* $x \in X_G$ *for which there is exactly one* $pq \in E$ *with* $x_{pq} = 1$*. Thus, the feasible* $x' \in X_G$ *such that* $x'_{12} = x'_{13} = x'_{23} = 1$ *is not optimal.*

6.1 Cut Conditions

Let us assume that we have found partially optimal assignments $E_0 \subseteq E$ by applying Proposition 11 and partially optimal constraints $T_0 \subseteq T$ by applying Proposition 12, i.e., we know that there exists a solution x^* to CP3 such that $x^*_{pq} = 0$ for all $pq \in E_0$ and $x^*_{pq}x^*_{qr}x^*_{pr} = 0$ for all $pqr \in T_0$. Moreover, we denote by $X_G^{E_0,T_0} = \{x \in X_G \mid \forall pq \in E_0 \colon x_{pq} = 0 \wedge \forall pqr \in T_0 \colon x_{pq}x_{pr}x_{qr} = 0\}$ the set of feasible vectors $x \in X_G$ that fulfill the partially optimal assignments and partially optimal constraints. Thus we have $\min_{x \in X_G} \phi_c(x) = \min_{x \in X_G^{E_0,T_0}} \phi_c(x)$. In order not to lose already established partial optimality, we may from now on only apply improving maps $\sigma' \colon X_G^{E_0,T_0} \to X_G^{E_0,T_0}$. This is the case for the elementary cut map restricted to $X_G^{E_0,T_0}$, i.e., $\sigma_{\delta(U)}(X_G^{E_0,T_0}) \subseteq X_G^{E_0,T_0}$ for every $U \subseteq V$. Since the cut conditions Proposition 11 and Proposition 12 use only the elementary cut map or the identity map to improve solutions, we can apply these cut conditions simultaneously.

On the contrary, it can happen that the image of $X_G^{E_0,T_0}$ under σ_U is not a subset of $X_G^{E_0,T_0}$ for some $U \subseteq V$, i.e., $\sigma_U(X_G^{E_0,T_0}) \nsubseteq X_G^{E_0,T_0}$. In particular, $\sigma_V(X_G^{E_0,T_0}) = \{\mathbb{1}_V\} \nsubseteq X_G^{E_0,T_0}$ for non-empty $E_0 \neq \emptyset$ and $T_0 \neq \emptyset$. Therefore, we cannot apply join conditions if we want to maintain already-established partial optimality due to Propositions 11 and 12.

6.2 Join Conditions

Let us assume the existence of an optimal solution x^* to CP3(G,c) such that $x^*_{ij} = 1$ for some $ij \in E$. We define $X_G|_{x_{ij}=1} = \{x \in X_G \mid x_{ij} = 1\}$. Then, we have

$$\min_{x \in X_G} \phi_c(x) = \min_{x \in X_G|_{x_{ij}=1}} \phi_c(x) \ . \tag{33}$$

Let $V' = V \setminus \{j\}, E' = \left(E \cap \binom{V' \setminus \{i\}}{2}\right) \cup \{pi \in \binom{V'}{2} \mid pi \in E \vee pj \in E\}$ and $G' = (V', E')$, that is, we contract i and j. Now, we relate feasible vectors of $X_G|_{x_{ij}=1}$ to feasible vectors of $X_{G'}$. We observe that for any $x \in X_G|_{x_{ij}=1}$ we have $\forall p \in V \setminus \{i,j\} \colon x_{pi} = x_{pj}$ whenever $pi, pj \in E$ by transitivity. We define $\varphi_{ij} \colon X_G|_{x_{ij}=1} \to X_{G'}$ as

$$\begin{aligned} \varphi_{ij}(x)_{pi} &= \begin{cases} x_{pi} & \text{if } pi \in E \wedge pj \notin E \\ x_{pj} & \text{if } pi \notin E \wedge pj \in E \\ x_{pi} = x_{pj} & \text{if } pi \in E \wedge pj \in E \end{cases} && \forall p \in V' \setminus \{i\} \colon pi \in E' \\ \varphi_{ij}(x)_{pq} &= x_{pq} && \forall pq \in E \cap \tbinom{V' \setminus \{i\}}{2} \ . \end{aligned}$$

It is easy to see that φ_{ij} is bijective. Proposition 24 below shows that solving the right-hand side of (33) is equivalent to solving a smaller instance of the original problem.

Proposition 24 *Let $G = (V, E)$ a graph and $c \in \mathbb{R}^{I(G)}$. Moreover, let $ij \in E$ and $V' = V \setminus \{j\}$, $E' = \left(E \cap \binom{V' \setminus \{i\}}{2}\right) \cup \{pi \in \binom{V'}{2} \mid pi \in E \vee pj \in E\}$ and $G' = (V', E')$. Define $c' \in \mathbb{R}^{I(G')}$ such that*

$$
\begin{aligned}
c'_{pqr} &= c_{pqr} && \forall pqr \in T \cap \tbinom{V' \setminus \{i\}}{3} \\
c'_{pqi} &= \sum_{r \in \{i,j\} \colon pqr \in T} c_{pqr} && \forall pq \in E \cap \tbinom{V' \setminus \{i\}}{2} \colon pqi \in T \vee pqj \in T \\
c'_{pq} &= c_{pq} && \forall pq \in E \cap \tbinom{V' \setminus \{i\}}{2} \\
c'_{pi} &= \sum_{q \in \{i,j\} \colon pq \in E} c_{pq} && \forall p \in V' \setminus \{i\} \colon (pi \in E \vee pj \in E) \wedge pij \notin T \\
c'_{pi} &= c_{pi} + c_{pj} + c_{pij} && \forall p \in V' \setminus \{i\} \colon pij \in T \\
c'_{\emptyset} &= c_{\emptyset} + c_{ij} \ .
\end{aligned}
$$

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

Proof Let $x \in X_G|_{x_{ij}=1}$ and $x' = \varphi_{ij}(x)$. We show that $\phi_c(x) = \phi_{c'}(x')$. Note that we have

$$
\begin{aligned}
T' &= \{pqr \in \tbinom{V'}{3} \mid pq \in E' \wedge pr \in E' \wedge qr \in E'\} \\
&= \left(T \cap \tbinom{V' \setminus \{i\}}{3}\right) \cup \left\{pqi \in \tbinom{V'}{3} \mid pq \in \left(E \cap \tbinom{V' \setminus \{i\}}{2}\right) \wedge (pqi \in T \vee pqj \in T)\right\}
\end{aligned}
$$

We use the fact that $x_{pi} = x_{pj}$ for all $p \in V \setminus \{i, j\}$ such that $pi, pj \in E$, and $x_{ij} = 1$. It follows

$$
\begin{aligned}
&\phi_{c'}(x') = \sum_{pqr \in T'} c'_{pqr} x'_{pq} x'_{pr} x'_{qr} + \sum_{pq \in E'} c'_{pq} x'_{pq} + c'_{\emptyset} \\
&= \sum_{pqr \in T \cap \binom{V' \setminus \{i\}}{3}} c'_{pqr} x'_{pq} x'_{pr} x'_{qr} + \sum_{\substack{pq \in E \cap \binom{V' \setminus \{i\}}{2} \colon \\ pqi \in T \vee pqj \in T}} c'_{pqi} x'_{pq} x'_{pi} x'_{pj} + \sum_{pq \in E \cap \binom{V' \setminus \{i\}}{2}} c'_{pq} x'_{pq} + \sum_{\substack{p \in V' \setminus \{i\} \colon \\ pi \in E \vee pj \in E}} c'_{pi} x'_{pi} + c'_{\emptyset} \\
&= \sum_{\substack{pq \in E \cap \binom{V \setminus \{i,j\}}{2} \colon \\ pqi \in T \vee pqj \in T}} c'_{pqi} x'_{pi} x'_{qi} x'_{pq} + \sum_{pqr \in T \cap \binom{V \setminus \{i,j\}}{3}} c'_{pqr} x'_{pq} x'_{pr} x'_{qr} + \sum_{\substack{p \in V \setminus \{i,j\} \colon \\ pi \in E \vee pj \in E}} c'_{pi} x'_{pi} + \sum_{pq \in E \cap \binom{V \setminus \{i,j\}}{2}} c'_{pq} x'_{pq} + c'_{\emptyset} \\
&= \sum_{\substack{pq \in E \cap \binom{V \setminus \{i,j\}}{2} \colon \\ pqi \in T}} c_{pqi} x_{pi} x_{qi} x_{pq} + \sum_{\substack{pq \in E \cap \binom{V \setminus \{i,j\}}{2} \colon \\ pqj \in T}} c_{pqj} x_{pj} x_{qj} x_{pq} + \sum_{pqr \in T \cap \binom{V \setminus \{i,j\}}{3}} c_{pqr} x_{pq} x_{pr} x_{qr} \\
&\quad + \sum_{pq \in E \cap \binom{V \setminus \{i,j\}}{2}} c_{pq} x_{pq} + \sum_{\substack{p \in V \setminus \{i,j\} \colon \\ pij \in T}} c_{pij} x_{pi} + \sum_{\substack{p \in V \setminus \{i,j\} \colon \\ pi \in E}} c_{pi} x_{pi} + \sum_{\substack{p \in V \setminus \{i,j\} \colon \\ pj \in E}} c_{pj} x_{pj} + c_{\emptyset} + c_{ij} x_{ij} \\
&= \sum_{pqr \in T} c_{pqr} x_{pq} x_{pr} x_{qr} + \sum_{pq \in E} c_{pq} x_{pq} + c_{\emptyset} = \phi_c(x).
\end{aligned}
$$

Therefore, we have $\min_{x \in X_G|_{x_{ij}=1}} \phi_c(x) = \min_{x \in X_G|_{x_{ij}=1}} \phi_{c'}(\varphi_{ij}(x)) = \min_{x \in X_{G[V']}} \phi_{c'}(x)$. ■

6.3 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 Proposition 10). Firstly, we apply Proposition 10, which leaves us with independent sub-problems. Secondly, we apply our join conditions until we find an edge $ij \in E$ or triple $ijk \in T$ to join, starting from Corollary 22 and then moving on to Propositions 14 and 16, Corollaries 19 and 20 and Proposition 15, in this order. Thirdly, we apply the remaining cut conditions, which can be applied jointly as we have seen in Section 6.1.

7 Experiments

We examine the effect of the algorithms empirically on two data sets. For both of them, 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 20 instances. We apply all partial optimality conditions jointly, as described in Section 6.3, and we also evaluate the effect of each condition separately. All algorithms are implemented in C++ and run on one core of an Intel Core i9-12900KF equipped with 64 GB of RAM. Note that compared to Stein et al. (2023) the hardware has changed. In the interest of objectivity, we run the experiments from Stein et al. (2023) on the new hardware and report the results in Appendix B.

7.1 Partition Data Set

Instances of the partition data set are defined with respect to a graph $G = (V, E)$ with $|V| = 8n$ nodes, where $n \in \mathbb{N}$, and with respect to a graph partition $\Pi_G = \{U_1, U_2, U_3, U_4\}$ into four sets with $|U_1| = n$, $|U_2| = |U_3| = 2n$ and $|U_4| = 3n$ elements, as depicted in Figure 1a, and with respect to three design parameters, $p_E \in [0, 1]$, $\alpha \in [0, 1]$ and $\beta \in [0, 1]$. We remark that we choose the same probability for all the edges in E, namely p_E.

The graph G is constructed randomly in the following way: Firstly, we insert every pair $pq \in \binom{V}{2}$ as edge to G at random with probability p_E. Secondly, for every induced subgraph $G[U_i]$, we iterate over all pairs $pq \in \binom{U_i}{2}$ and check if p and q are path-connected in $G[U_i]$. If this is not the case, then we add the edge pq as an edge to G.

We define costs for all edges E and for all triples $T = \{pqr \in \binom{V}{3} \mid \{pq, qr, pr\} \subseteq E\}$ that form a 3-clique in the graph. Both the costs of edges and triples are drawn 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 Π_G, and with standard deviation $\sigma = \sigma_0 + \alpha(\sigma_1 - \sigma_0)$ with $\sigma_0 = 0.1$ and $\sigma_1 = 0.4$. Costs of edges are multiplied by $1 - \beta$, and costs of triples by β. For small α, the costs of pairs and triples whose elements belong to the same set of the partition Π_G are well-separated from the costs of those pairs and triples whose elements belong to different sets of the partition Π_G. For α close to 1 this is not the case anymore. Therefore, the higher α is, the harder the problem becomes. The higher β is, the more important the costs of triples become.

The percentage of edges and triples fixed by applying all conditions jointly, as described in Section 6.3, and with respect to α is shown in Figure 2 (Rows 1–2). It can be seen from this figure that the percentage of fixed variables decreases with increasing α. As α rises, the runtime increases but remains below one minute for all the instances (Row 2). Varying β does not affect the overall trend. However, the percentage of fixed variables decreases as soon as triple costs are introduced. The percentage of variables fixed by applying Proposition 10, Proposition 11 and Corollary 22 separately is shown in Figures 3 and the percentage of triples fixed by applying Proposition 12 separately is shown in Figure 4. 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

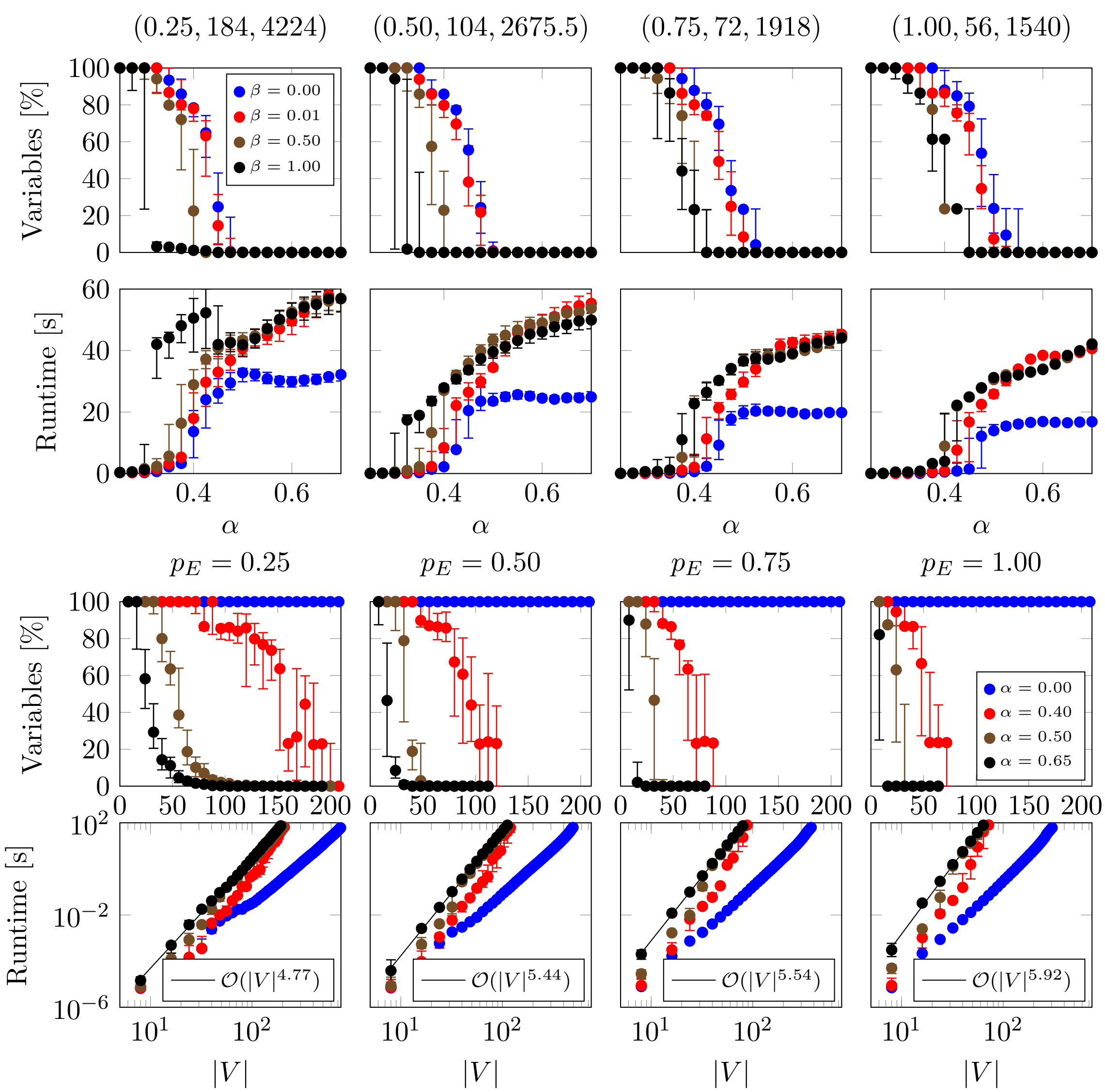

Figure 2: We report above for the partition data set the percentage of fixed variables and runtime after applying all conditions jointly, as described in Section 6.3, with respect to α (Rows 1–2) and with respect to $|V|$ (Rows 3–4). From left to right, the sparsity of the graph decreases as written in the column title, $(p_E, |V|, \bar{E})$ (Rows 1–2) and p_E (Rows 3–4).

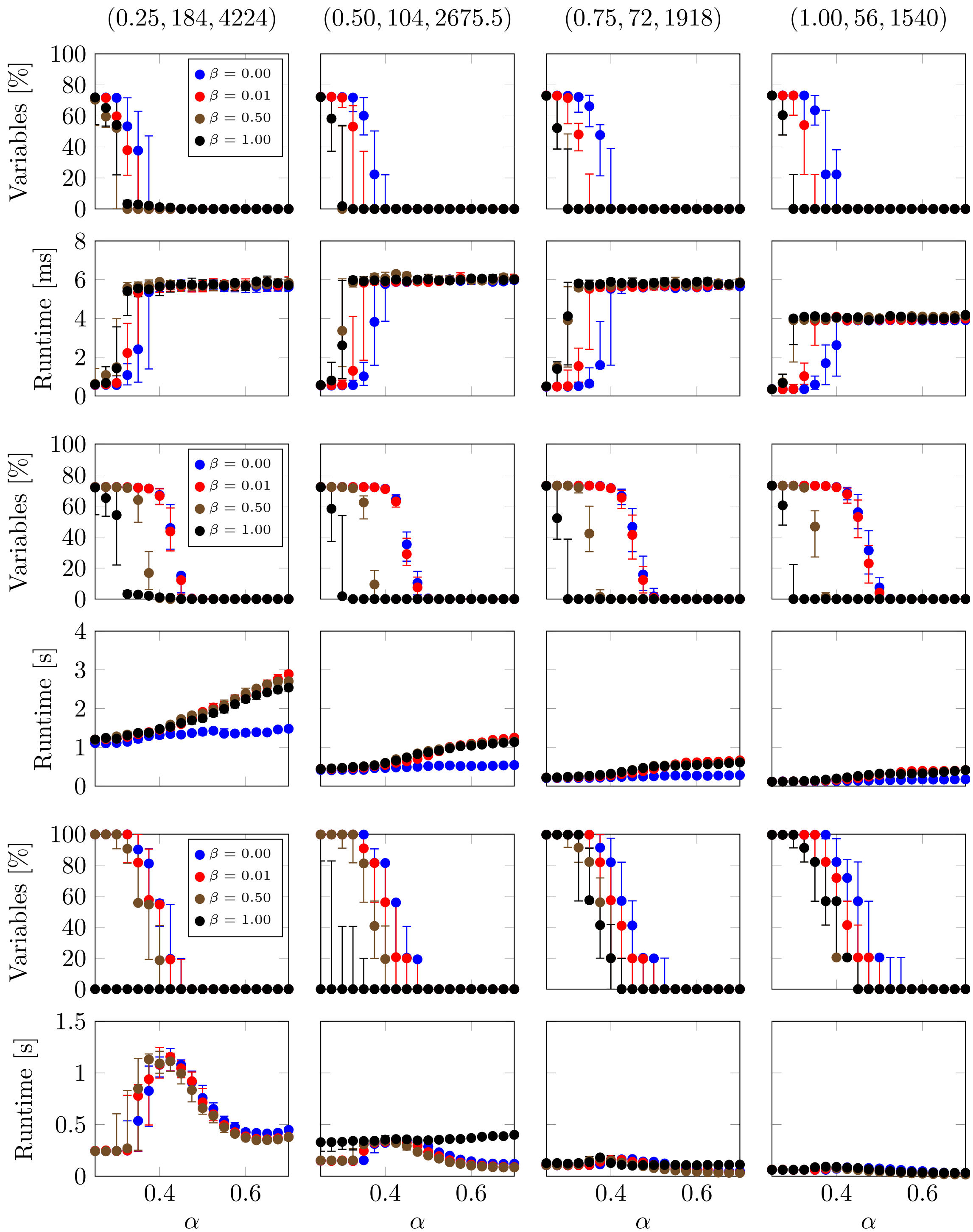

Figure 3: We report above for the partition data set the percentage of fixed variables and runtime after applying Proposition 10 (Rows 1–2), Proposition 11 (Rows 3–4) and Corollary 22 (Rows 5–6), separately. From left to right, the sparsity of the graph decreases as written in the column title $(p_E, |V|, \bar{E})$.

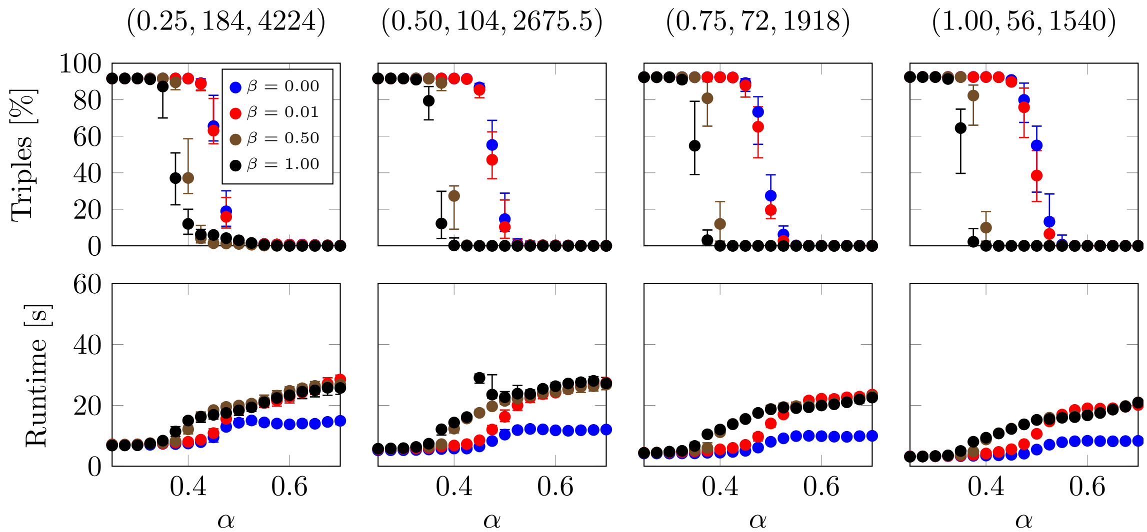

Figure 4: We report above for the partition data set the percentage of fixed triples and runtime after applying Proposition 12. From left to right, the sparsity of the graph decreases as written in the column title $(p_E, |V|, \bar{E})$.

one join condition provides partial optimality in this case: Corollary 22. Interestingly, this is the one statement that fixes the most variables for almost all instances of this data set.

For $\beta = 0.5$ and with respect to the instance size, the runtime and percentage of variables fixed by applying all conditions jointly are shown in Figure 2 (Rows 3–4). It can be seen that, as the instance size increases, the number of fixed variables declines while the runtime increases. The runtime for complete graphs, $p_E = 1.0$, and $\alpha \in \{0.5, 0.65\}$ converges to a $\mathcal{O}(|V|^{5.92})$ trend for increasing instance size. As we sparsify the graphs, i.e., for decreasing p_E, the percentage of fixed variables increases and the runtime converges to a $\mathcal{O}(|V|^{4.77})$ trend for $p_E = 0.25$ and $\alpha \in \{0.5, 0.65\}$. For complete graphs this is close to the expected worst-case complexity of $\mathcal{O}(|V|^6)$ as we need to solve cubically many max-flow problems if no variables can be fixed by our algorithm. Therefore, we observe that we can handle larger sparse graphs, induced by larger instances. This suggests that this approach might be able to handle instances of size interesting for practical applications.

7.2 Geometric Data Set

Next, we consider a data set 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 a Gaussian distribution with mean $\vec{a}$ and covariance matrix $\sigma^2 \mathbb{1}$. This defines the set of nodes, V. We build a graph $G = (V, E)$ from these points in the following way. First we add all pairs $pq \in \binom{V}{2}$ as edges if p and q were drawn from vertices of the same triangle. Second, with respect to a parameter $k \in \mathbb{N}$, we compute for each point $\vec{a}_p$ its k nearest neighbors, V_p^n, and its k farthest neighbors, V_p^f. Then we add for every $p \in V$ the pairs $\{pq \mid q \in V_p^n \cup V_p^f\}$ as edges to the graph G. Adding both the k nearest and the k farthest neighbors is motivated by the following intuition. Three points that are mutually close likely are drawn from the same triangle. For three points that are mutually far apart we can make the decision whether these points belong to the same triangle or to different triangles based on the deviation of interior angles from $\pi/3$. For three points that are neither mutually close or far apart, this decision is ambiguous. Therefore, we prefer to include both edges between far apart and close points. We let $T = \{pqr \in \binom{V}{3} \mid pq \in E, qr \in E, pr \in E\}$, i.e., we introduce costs for all triples that form a 3-clique.

For any three points $\vec{a}_p, \vec{a}_q, \vec{a}_r$ such that $pqr \in T$, let $\varphi_p, \varphi_q, \varphi_r$ be the interior angles of the triangle spanned by these points, and let d_{pqr}^{max} and d_{pqr}^{min} be the maximum and minimum length of edges in this triangle. If the three points are mutually close, $d_{pqr}^{max} \leq 4\sigma$, we reward

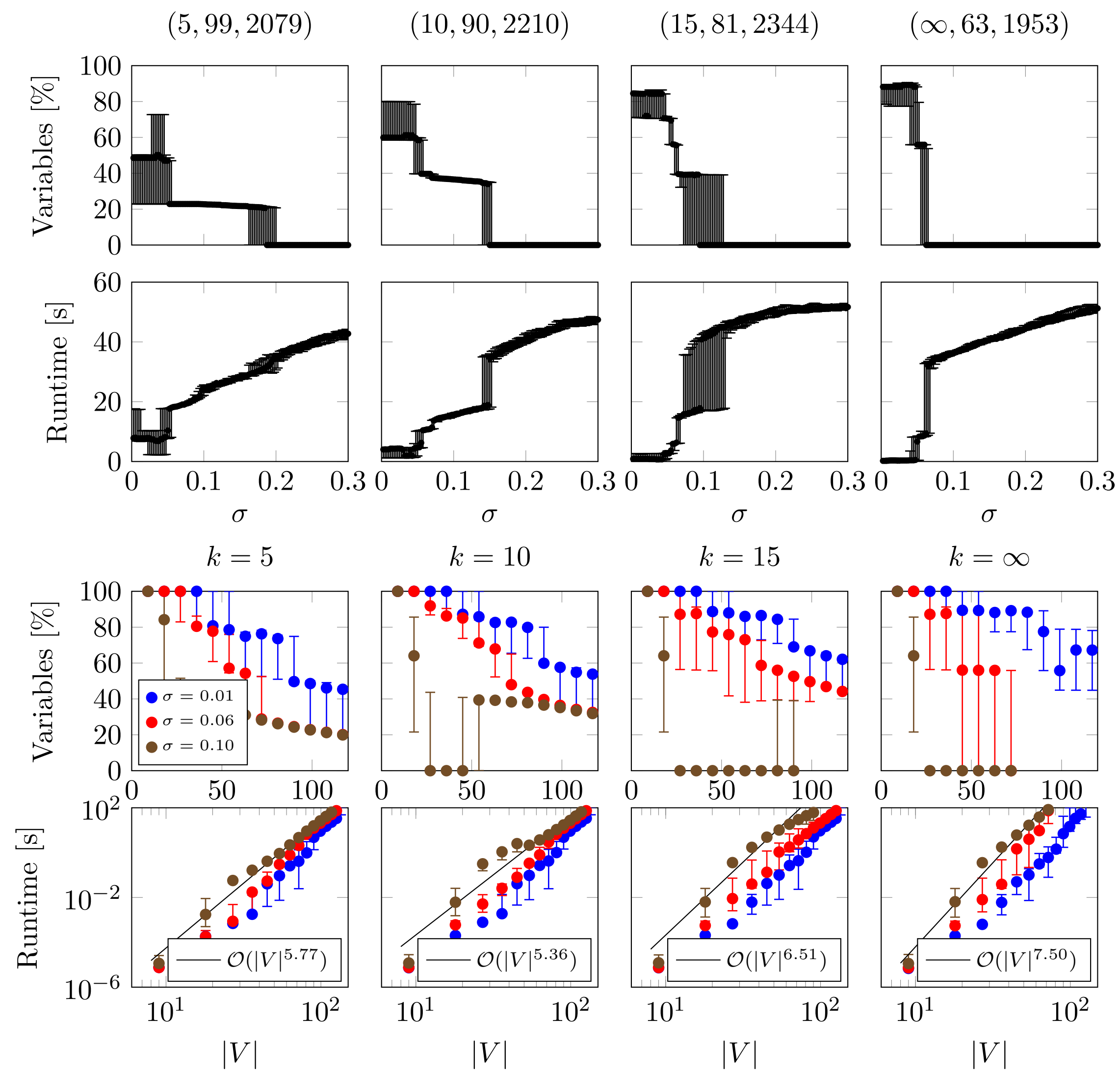

Figure 5: We report above for the geometric data set the percentage of fixed variables and runtime after applying all conditions jointly, as described in Section 6.3, with respect to σ (Rows 1–2) and with respect to $|V|$ (Rows 3–4). From left to right, the sparsity of the graph decreases as written in the column title, $(k, |V|, \bar{E})$ (Rows 1–2) and k (Rows 3–4).

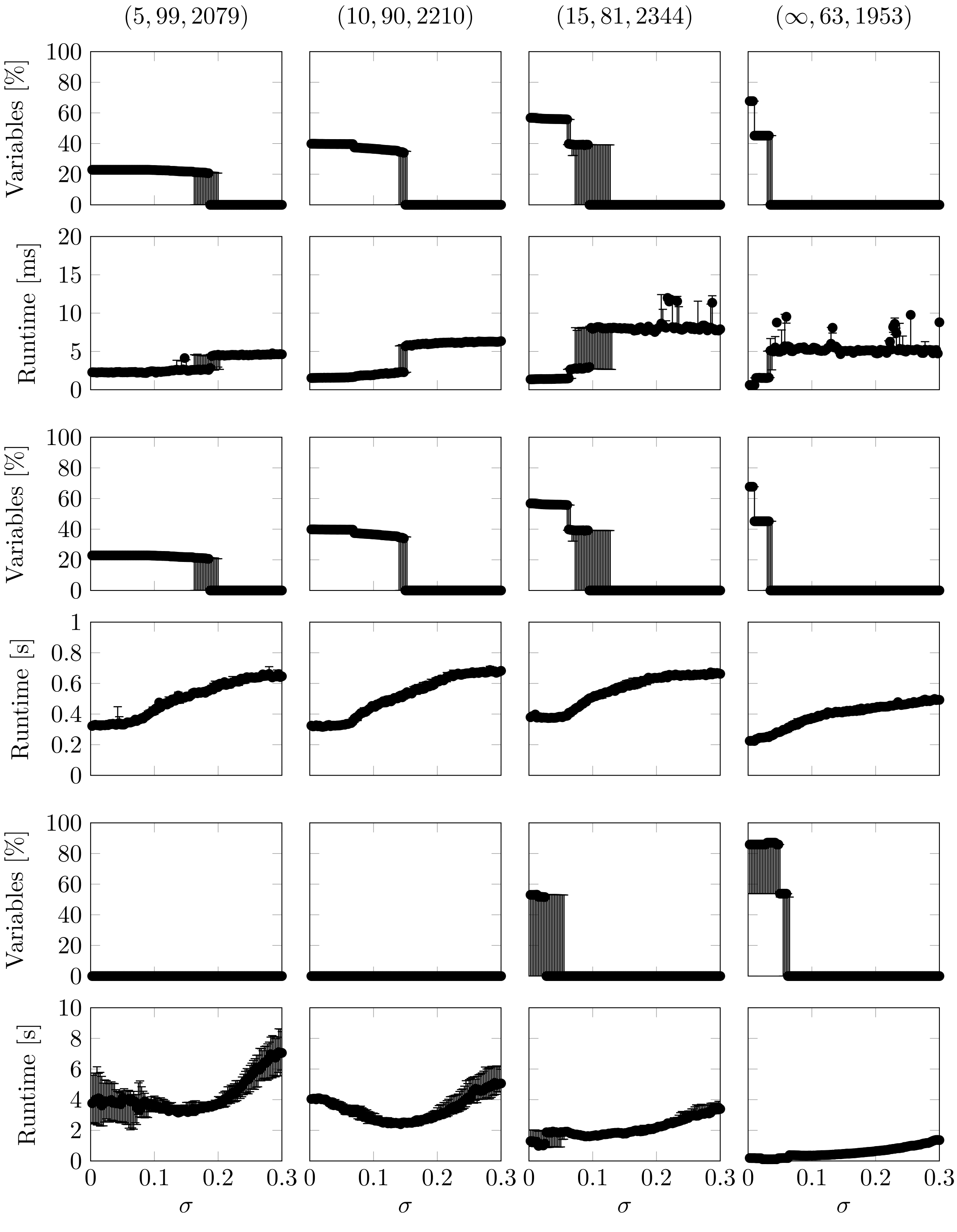

Figure 6: We report above for the geometric data set the percentage of fixed variables and runtime after applying Proposition 10 (Rows 1–2), Proposition 11 (Rows 3–4) and Corollary 22 (Rows 5–6), separately. From left to right, the sparsity of the graph decreases as written in the column title $(k, |V|, \bar{E})$.

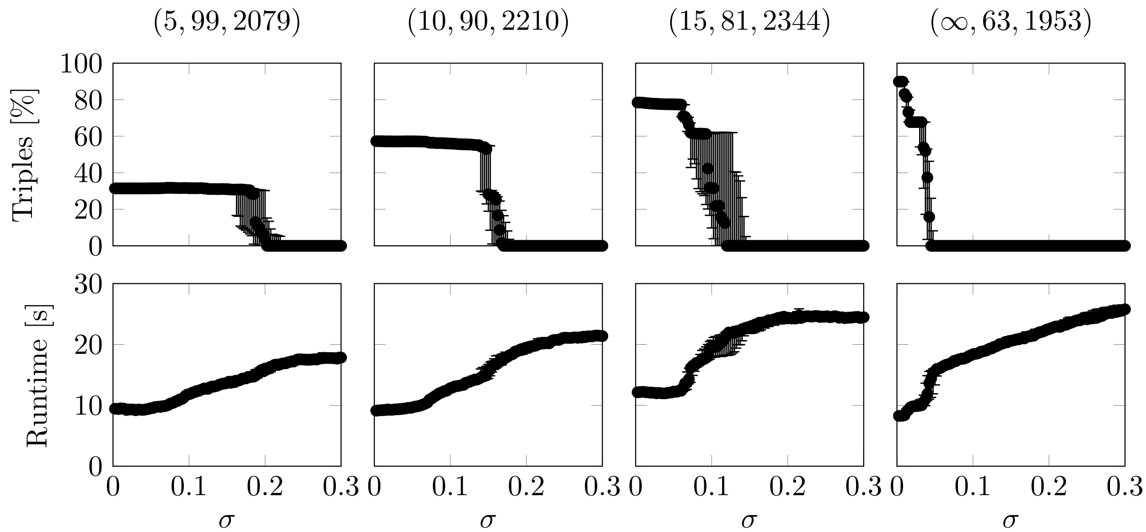

Figure 7: We report above for the geometric data set the percentage of fixed triples and runtime after applying Proposition 12. From left to right, the sparsity of the graph decreases as written in the column title $(k, |V|, \bar{E})$.

solutions in which these belong to the same set by letting $c_{pqr} = -1 + \frac{d_{pqr}^{max}}{4\sigma}$. If only two points are close, $d_{pqr}^{max} > 4\sigma$ and $d_{pqr}^{min} \leq 4\sigma$, we let $c_{pqr} = 0$. If the three points are mutually far apart, $d_{pqr}^{min} > 4\sigma$, 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 $\delta_{pqr} \leq \frac{\pi}{6}$, we let $c_{pqr} = -1 + \frac{6\delta_{pqr}}{\pi}$. Otherwise, $c_{pqr} = \frac{6}{7}\frac{\delta_{pqr} - \frac{\pi}{6}}{\pi}$.

The percentage of variables fixed by applying all conditions jointly, as described in Section 6.3, and with respect to σ is reported in Figure 5 (Rows 1–2). Here, the hardness of the instances is embodied by σ. Depending on the sparsity of the instances, the number of nodes ranges from $|V| = 54$ (for $k = \infty$) to $|V| = 90$ (for $k = 5$). As σ increases, the percentage of fixed variables decreases. The runtime increases, as can be seen from Figure 5 (Row 2), and stays below one minute for all these instances. We also observe that for sparser instances (lower k) we can fix less variables at the same noise level σ. The percentage of variables fixed by applying Proposition 10, Proposition 11 and Corollary 22 separately is shown in Figures 6 and the percentage of triples fixed by applying Proposition 12 separately is shown in Figure 7. Also here, all the cut conditions are effective whereas the only useful join condition is Corollary 22. Moreover, Corollary 22 is the most effective for complete graphs. In contrast, Propositions 10 and 11 are more effective for sparse instances.

The runtime and percentage of variables fixed by applying all conditions jointly and with respect to the instance size are shown in Figure 5 (Rows 3–4). Similar to the partition data set, we see that, as the instance size increases, the number of fixed variables declines while the runtime increases. For some σ, our conditions can fix more variables in instances defined on complete graphs than in instances defined on sparse graphs. For increasing instance size, the runtime for $\sigma = 0.1$ and $k = 5$ converges to a $\mathcal{O}(|V|^{5.77})$ trend and for $k = 10$ to a $\mathcal{O}(|V|^{5.36})$ trend. Again, this is close to the expected worst-case complexity of $\mathcal{O}(|V|^6)$. Moreover, we are able to handle larger sparse graphs, induced by larger instances.

8 Connection to the Cubic Multicut Problem

In this section, we turn to a natural question that arises in the context of correlation clustering and multicuts. When the objective function is linear, it is well-known that the two problems are equivalent, as mentioned also in the introduction. In particular, the feasible vectors of the two problems are in a 1-to-1 correspondence given by the affine map $x \mapsto 1 - x$. This implies that the optimal value of the (linear) correlation clustering problem is equal to a constant minus the optimal value of the (linear) multicut problem. It is therefore reasonable to ask ourselves whether the same behaviour occurs also in the nonlinear setting. One important thing to realize is that, while the cubic term in (1) tells us exactly what we want, i.e., it is equal to one if and only if all the three corresponding points are in the same node set, this expressiveness property does not transfer to the higher-order multicut problem. In fact, if we simply replace the variables with their affine transformation $x \mapsto 1 - x$, the cubic term becomes $(1 - x_{pq})(1 - x_{pr})(1 - x_{qr})$ for $pqr \in T$. We observe that this product equals one if and only if all the correlation clustering variables x_{pq}, x_{pr}, x_{qr} have value zero, meaning that the three points end up in three different node subsets. However, we would desire the cubic term in the multicut problem to evaluate 1 also when two points out of three are together but the third one is in another node subset. This suggests that the higher-order multicut problem cannot be expressed naturally without the introduction of additional variables, contrary to the higher-order correlation clustering problem. Next, we define the variables of the cubic multicut problem, and show the relationship between the optimal values of this problem and of the cubic correlation clustering.

$$\forall pq \in E, \qquad z_{pq} := 1 - x_{pq} = \begin{cases} 1, & \text{if } p, q \text{ are in different subsets,} \\ 0, & \text{if } p, q \text{ otherwise,} \end{cases}$$

$$\forall pqr \in T, \qquad y_{pqr} := 1 - x_{pq}x_{pr}x_{qr} = \begin{cases} 1, & \text{if } p, q, r \text{ are not all together,} \\ 0, & \text{if } p, q, r \text{ otherwise.} \end{cases}$$

The cubic multicut problem w.r.t. to a graph $G = (V, E)$, triples T, and costs $c \in \mathbb{R}^{I(G)}$ is:

$$\begin{aligned} \min_{\substack{z: E \to \{0,1\} \\ y: T \to \{0,1\}}} \quad & \sum_{pqr \in T} c_{pqr} y_{pqr} + \sum_{pq \in E} c_{pq} z_{pq} + c_\emptyset \\ \text{s.t.} \quad & \forall (V_C, E_C) \in \text{cycles}(G)\, \forall e' \in E_C \colon \quad z_{e'} \le \sum_{e \in E_C \setminus \{e'\}} z_e \\ & \forall pqr \in T \colon \quad y_{pqr} \le z_{pq} + z_{qr} + z_{pr} \\ & \forall pqr \in T \colon \quad y_{pqr} \ge z_{pq} \ \text{ and } \ y_{pqr} \ge z_{qr} \ \text{ and } \ y_{pqr} \ge z_{pr} \end{aligned} \tag{34}$$

We denote the objective value of (1) by obj_{CP3} and the value of (34) by obj_{CMCP}:

$$\begin{aligned} \text{obj}_{\text{CP3}} &= \sum_{pqr \in T} c_{pqr} x_{pq} x_{pr} x_{qr} + \sum_{pq \in E} c_{pq} x_{pq} + c_\emptyset \\ &= \sum_{pqr \in T} c_{pqr} - \sum_{pqr \in T} c_{pqr}(1 - x_{pq}x_{pr}x_{qr}) + \sum_{pq \in E} c_{pq} - \sum_{pq \in E} c_{pq}(1 - x_{pq}) - c_\emptyset + 2c_\emptyset \\ &= \sum_{pqr \in T} c_{pqr} - \sum_{pqr \in T} c_{pqr} y_{pqr} + \sum_{pq \in E} c_{pq} - \sum_{pq \in E} c_{pq} z_{pq} - c_\emptyset + 2c_\emptyset = -\text{obj}_{\text{CMCP}} + C, \end{aligned}$$

where C is a constant (specifically, $C = \sum_{pqr \in T} c_{pqr} + \sum_{pq \in E} c_{pq} + 2c_\emptyset$).

We conclude that the two problems are equivalent, similarly to the linear case. However, the two problems have a different number of variables in the non-linear setting. Hence, all partial optimality results obtained for one problem can be transferred to the other problem once we flip the sign of the costs and the binary values of the variables.

9 Conclusion

We establish partial optimality conditions for the cubic correlation clustering problem for general graphs. In particular, we generalize the conditions and the implementation established by Stein et al. (2023) for the case of complete graphs to that of general graphs. We empirically evaluate these conditions on two data sets and make the following observations: Firstly, we observe for the specific instances investigated in this article that all cut conditions are effective and the only join condition that is effective is Proposition 21 (in its simplified form of Corollary 22). In the interest of objectivity, we discuss all conditions generalized from prior work unconditionally and report also the ineffectiveness of some conditions. Secondly, we observe that our implementation can handle larger instances, with respect to the number of elements, as we sparsify the instances.

One question we do not answer here is in which order to apply certified joins. The fraction of fixed variables depends on this order, and looking into heuristics for choosing an order appears promising. Toward future work, it would be interesting to see whether partial optimality conditions are completely or partially transferable to higher-order correlation clustering or to the case where triple costs additionally depend on node labels.

Acknowledgments and Disclosure of Funding

David Stein acknowledges partial support by the Federal Ministry of Education and Research of Germany through grant 16LW0079K. Bjoern Andres acknowledges partial support by the Federal Ministry of Education and Research of Germany through grant 16KIS2332K.

Appendix A. Reductions to Minimum Cut Problems

Here, we discuss how, for a given pair or triple, we reduce the search for subsets $U, U' \subseteq V$ that satisfy (6)–(11) maximally to the min st-cut problem. In any of these cases, we have $G = (V, E)$ be a graph and $c \in \mathbb{R}^{I(G)}$ such that $c_{pqr} \geq 0$ for all $pqr \in T$, and $c_{pq} \geq 0$ for all $pq \in E$. Moreover, in the cases of (6) and (8) we have an edge $ij \in E$ such that $i \in U$ and $j \notin U$, while in the cases of (7), (9), (10) and (11) we have a triple $ijk \in T$ such that $ij, ik \in \delta(U)$. Note that in the latter case we can assume that $i \in U$ and $j, k \notin U$ without loss of generality. We handle both cases jointly by denoting the set of vertices not in U by $V_0 \subseteq V$. The search for subsets that satisfy (6)–(11) maximally then takes the form:

$$\min_{\substack{U \subseteq V: \\ i \in U, \\ j \notin U, \forall j \in V_0}} \quad \sum_{pqr \in T_{\delta(U)}} c_{pqr} + \sum_{pq \in \delta(U)} c_{pq} \quad . \tag{35}$$

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

Proposition 25 *Let $G = (V, E)$ be a graph, $c \in \mathbb{R}^{I(G)}$ and $U \subseteq V$. Then $\sum_{pqr \in T_{\delta(U)}} c_{pqr} = \frac{1}{2} \sum_{pq \in \delta(U)} \sum_{\substack{r \in V \setminus \{p,q\}: \\ pqr \in T}} c_{pqr}$.*

Proof Let $U \subseteq V$. Observe that

$$\begin{aligned}
\sum_{pqr \in T_{\delta(U)}} c_{pqr} &= \sum_{pq \in \binom{U}{2}} \sum_{\substack{r \in V \setminus U: \\ pqr \in T}} c_{pqr} + \sum_{pq \in \binom{V \setminus U}{2}} \sum_{\substack{r \in U: \\ pqr \in T}} c_{pqr} \\
&= \frac{1}{2} \sum_{p \in U} \sum_{q \in U \setminus \{p\}} \sum_{\substack{r \in V \setminus U: \\ pqr \in T}} c_{pqr} + \frac{1}{2} \sum_{p \in V \setminus U} \sum_{q \in V \setminus (U \cup \{p\})} \sum_{\substack{r \in U: \\ pqr \in T}} c_{pqr} \\
&= \frac{1}{2} \sum_{p \in U} \sum_{q \in V \setminus U} \Big(\sum_{\substack{r \in U \setminus \{p\}: \\ pqr \in T}} c_{pqr} + \sum_{\substack{r \in V \setminus (U \cup \{q\}): \\ pqr \in T}} c_{pqr} \Big) = \frac{1}{2} \sum_{pq \in \delta(U)} \sum_{\substack{r \in V \setminus \{p,q\}: \\ pqr \in T}} c_{pqr}.
\end{aligned}$$

■

Consequently, (35) is equivalent to $\min_{\substack{U \subseteq V: \\ i \in U, \\ j \notin U, \forall j \in V_0}} \sum_{pq \in \delta(U)} c'_{pq}$, where $c'_{pq} = c_{pq} + \frac{1}{2} \sum_{\substack{r \in V \setminus \{p,q\} \\ pqr \in T}} c_{pqr}$ for all $pq \in E$. Next, we reduce the above minimization problem to a *quadratic unconstrained binary optimization problem*, by applying the following proposition:

Proposition 26 *Let $G = (V, E)$ be a graph and $c \in \mathbb{R}^E$. Moreover, let $i \in V$ and $V_0 \subseteq V \setminus \{i\}$. Furthermore, let $V' = V \setminus (V_0 \cup \{i\})$ and $c' \colon (E \cap \binom{V'}{2}) \cup V' \cup \{\emptyset\} \to \mathbb{R}$ such that*

$$c'_p = \sum_{q \in V \setminus \{p\}: pq \in E} c_{pq} - 2c_{pi} \; \forall p \in V' \colon pi \in E, \qquad c'_p = \sum_{q \in V \setminus \{p\}: pq \in E} c_{pq} \; \forall p \in V' \colon pi \notin E$$

$$c'_{pq} = -2c_{pq} \; \forall pq \in E \cap \tbinom{V'}{2}, \qquad c'_\emptyset = \sum_{q \in V \setminus \{i\}: qi \in E} c_{qi}.$$

Then:

$$\min_{\substack{U \subseteq V: \\ i \in U, \\ j \notin U, \forall j \in V_0}} \sum_{pq \in \delta(U)} c_{pq} = \min_{y \in \{0,1\}^{V'}} \sum_{pq \in E \cap \binom{V'}{2}} c'_{pq} y_p y_q + \sum_{p \in V'} c'_p y_p + c'_\emptyset \quad . \tag{36}$$

Proof Let $U \subseteq V$ such that $i \in U$ and $\forall j \in V_0 \colon j \notin U$. We define $y \in \{0,1\}^V$ such that $y = \mathbb{1}_U$. Then, we have $y_i = 1$ and $\forall j \in V_0 \colon y_j = 0$. Moreover, it follows

$$
\begin{aligned}
\sum_{pq \in \delta(U)} c_{pq} &= \sum_{pq \in E} c_{pq}\left(y_p(1-y_q) + y_q(1-y_p)\right) = \sum_{pq \in E} c_{pq}\left(y_p + y_q - 2y_py_q\right) = -2\sum_{pq\in E} c_{pq}y_py_q + \sum_{\substack{p,q\in V:\\ p\neq q,\\ pq\in E}} c_{pq}y_p \\
&= -2 \sum_{pq \in E\cap\binom{V'}{2}} c_{pq}y_py_q - 2\sum_{\substack{p\in V':\\ pi\in E}} c_{pi}y_p + \sum_{p\in V'}\sum_{\substack{q\in V\setminus\{p\}:\\ pq\in E}} c_{pq}y_p + \sum_{\substack{q\in V\setminus\{i\}:\\ qi\in E}} c_{qi} \\
&= -2 \sum_{pq \in E\cap\binom{V'}{2}} c_{pq}y_py_q + \sum_{\substack{p\in V':\\ pi\in E}} \Big(-2c_{pi} + \sum_{\substack{q\in V\setminus\{p\}:\\ pq\in E}} c_{pq}\Big)y_p + \sum_{\substack{p\in V':\\ pi\notin E}}\sum_{\substack{q\in V\setminus\{p\}:\\ pq\in E}} c_{pq}y_p + \sum_{\substack{q\in V\setminus\{i\}:\\ qi\in E}} c_{qi} \\
&= \sum_{pq\in\binom{V'}{2}} c'_{pq}y_py_q + \sum_{p\in V'} c'_p y_p + c'_\emptyset. \qquad \blacksquare
\end{aligned}
$$

For the instances of (36) that arise from deciding (6)–(11), we have $\forall pq \in E \cap \binom{V'}{2} \colon c'_{pq} \leq 0$. Thus, the right-hand side of (36) is *submodular* and can be minimized in strongly polynomial time (Boros et al., 2008; Kolmogorov and Zabin, 2004). For completeness we describe the reduction of quadratic unconstrained binary optimization to an instance of min-*st*-cut in detail: Lemma 27 reformulates the objective function of quadratic unconstrained binary optimization. Proposition 28 shows a reduction of this reformulation to min-*st*-cut.

Lemma 27 *Let $G = (V, E)$ be a graph and $c \in \mathbb{R}^{V\cup E}$. We define $c' \colon \mathbb{R}^{V\cup E}$ as*

$$
c'_p = c_p + \frac{1}{2}\sum_{\substack{q\in V\setminus\{p\}:\\ pq\in E}} c_{pq} \quad \forall p \in V, \qquad c'_{pq} = -\frac{1}{2}c_{pq} \quad \forall pq \in E.
$$

Then, for any $y \in \{0,1\}^V$ we have that

$$
\sum_{pq\in E} c_{pq}y_py_q + \sum_{p\in V} c_py_p = \sum_{p\in V}\sum_{\substack{q\in V\setminus\{p\}:\\ pq\in E}} c'_{pq}y_p(1-y_q) + \sum_{\substack{p\in V:\\ c'_p>0}} c'_py_p - \sum_{\substack{p\in V:\\ c'_p<0}} c'_p(1-y_p) + \sum_{\substack{p\in V:\\ c'_p<0}} c'_p \ .
$$

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

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

Proposition 28 *Let $G = (V, E)$ be a graph and $c \in \mathbb{R}^{V \cup E}$. We define $\phi_c \colon \{0,1\}^V \to \mathbb{R}$ such that for all $y \in \{0,1\}^V$ it holds that*

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

Furthermore, we define $V' = V \cup \{s, t\}$, $E' \subset V' \times V'$ such that $(s, p) \in E' \Leftrightarrow c_p < 0$, $\forall p \in V$; $(p, t) \in E' \Leftrightarrow c_p > 0$, $\forall p \in V$; and $(p, q) \in E' \wedge (q, p) \in E'$, $\forall pq \in E$. We also define $c' \in \mathbb{R}^{E'}$ as $c'_{(s,p)} = -c_p$, $\forall (s, p) \in E'$; $c'_{(p,t)} = c_p$, $\forall (p, t) \in E'$; and $c'_{(p,q)} = c'_{(q,p)} = c_{pq}$, $\forall pq \in E$. Moreover, we define the function $\varphi_{c'} : \{0,1\}^{V'} \to \mathbb{R}$ such that for all $y \in \{0,1\}^{V'}$ it holds that $\varphi_{c'}(y) = \sum_{(p,q) \in E'} c'_{(p,q)} y_p (1 - y_q)$. Then we have that

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

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

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

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

If $c_{pq} \leq 0$, $pq \in E$, and therefore $c'_{pq} \geq 0$, $\forall pq \in E$, in Lemma 27 the resulting instance can be solved efficiently.

Appendix B. Reproduction of Experiments from Stein et al. (2023)

We run experiments from Stein et al. (2023) on the same hardware as used for our experiments in this manuscript. Moreover, we impose a runtime limit of $T = 60$s.

B.1 Partition Data Set

In Figure 8 and Figure 10 (Row 1) we show the same experiments as described in Section 7.1 for complete graphs, but with the code from Stein et al. (2023). We observe a slightly better runtime trend of $\mathcal{O}(|V|^{5.79})$ than for our new implementation applied to complete graphs.

B.2 Geometric Data Set

In Figure 9 and Figure 10 (Row 2) we show the same experiments as described in Section 7.2 for complete graphs, but with the code from Stein et al. (2023). We observe a worse runtime behavior of $\mathcal{O}(|V|^{7.45})$, in contrast to the behavior of $\mathcal{O}(|V|^{5.8})$ reported in Stein et al. (2023). We conjecture that the reason for this difference lies in the different runtime limits imposed, i.e., in this manuscript we impose a runtime limit of 60s and in Stein et al. (2023) it is 1000s. Therefore, we assume that the runtimes for small $|V|$ affect the runtime trend disproportionately. Indeed, if we ignore the first two points in Figure 9d), then we obtain a runtime trend of $\mathcal{O}(|V|^{5.54})$ that is closer to $\mathcal{O}(|V|^{5.8})$ reported in Stein et al. (2023).

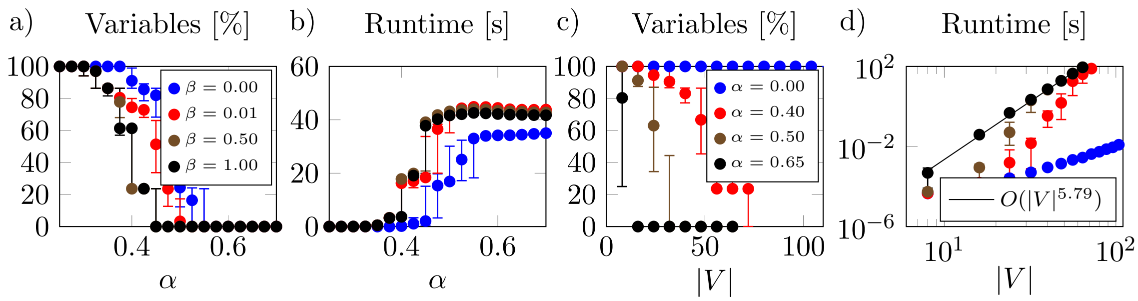

Figure 8: We report above for the partition dataset the percentage of fixed variables and the runtime after applying all conditions jointly, as described in Section 6.3, for the implementation from Stein et al. (2023) for complete graphs and on the new hardware. a-b) show these with respect to α and $|V| = 56$. c-d) show these with respect to $|V|$.

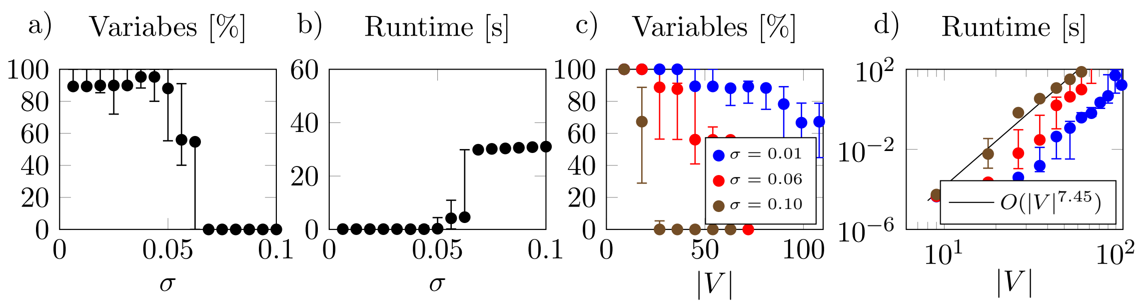

Figure 9: We report above for the geometric dataset the percentage of fixed variables and the runtime after applying all conditions jointly, as described in Section 6.3, for the implementation from Stein et al. (2023) for complete graphs and on the new hardware. a-b) show these with respect to σ and $|V| = 54$. c-d) show these with respect to $|V|$.

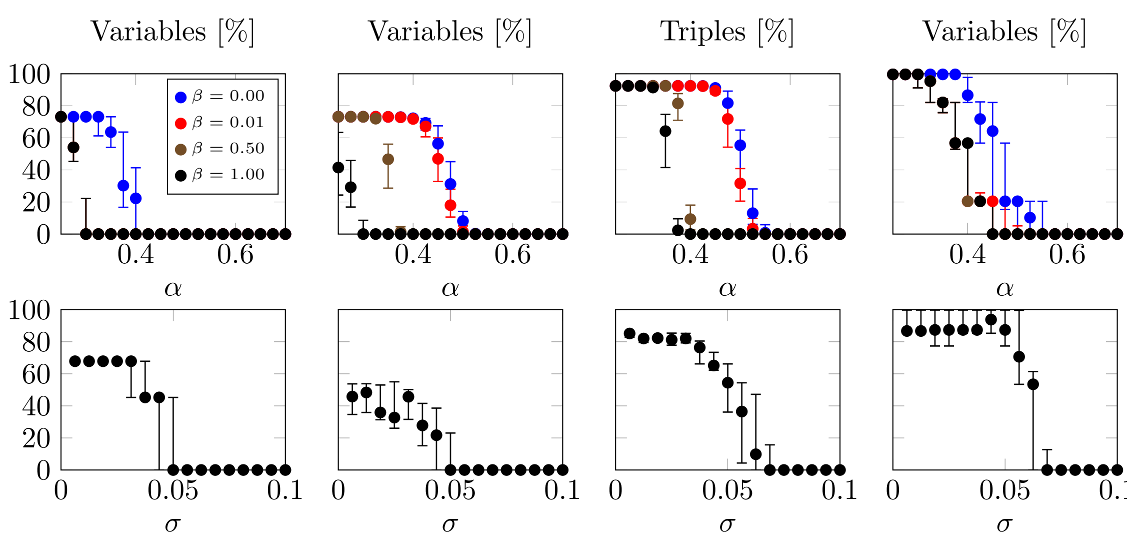

Figure 10: We report above the percentage of fixed variables and triples when applying Propositions 10 to 12 and Corollary 22 (Columsn 1–4), separately, and for the implementation from Stein et al. (2023) for complete graphs and on the new hardware. Row 1 shows these for the partition data set with 56 elements. Row 2 shows these for the geometric data set with 54 elements.

References

Warren Philip Adams, Julie Bowers Lassiter, and Hanif D. Sherali. Persistency in 0-1 polynomial programming. *Mathematics of Operations Research*, 23(2):359–389, 1998. doi: 10.1287/moor.23.2.359.

Sameer Agarwal, Jongwoo Lim, Lihi Zelnik-Manor, Pietro Perona, David Kriegman, and Serge Belongie. Beyond pairwise clustering. In *CVPR*, 2005. doi: 10.1109/CVPR.2005.89.

Amir Alush and Jacob Goldberger. 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.

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

Endre Boros, Peter Ladislaw Hammer, Richard Sun, and Gabriel Tavares. 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.

Sunil Chopra and Mendu Rammohan Rao. The partition problem. *Math. Program.*, 59 (1–3):87–115, 1993. doi: 10.1007/BF01581239.

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

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

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

Andrea Horňáková, Jan-Hendrik Lange, and Bjoern Andres. Analysis and optimization of graph decompositions by lifted multicuts. In *ICML*, 2017. URL `https://dl.acm.org/doi/10.5555/3305381.3305540`.

Jörg Hendrik Kappes, Markus Speth, Gerhard Reinelt, and Christoph Schnörr. 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.

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

Sungwoong Kim, Chang Dong Yoo, Sebastian Nowozin, and Pushmeet Kohli. 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.

Pushmeet Kohli, Alexander Shekhovtsov, Carsten Rother, Vladimir Kolmogorov, and Philip Torr. On partial optimality in multi-label mrfs. In *ICML*, 2008. doi: 10.1145/1390156.1390217.

Vladimir Kolmogorov and Ramin Zabin. 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.

Jan-Hendrik Lange, Andreas Karrenbauer, and Bjoern Andres. Partial optimality and fast lower bounds for weighted correlation clustering. In *ICML*, 2018. URL `https://proceedings.mlr.press/v80/lange18a.html`.

Jan-Hendrik Lange, Bjoern Andres, and Paul Swoboda. Combinatorial persistency criteria for multicut and max-cut. In *CVPR*, 2019. doi: 10.1109/CVPR.2019.00625.

Evgeny Levinkov, Amirhossein Kardoost, Bjoern Andres, and Margret Keuper. Higher-order multicuts for geometric model fitting and motion segmentation. *Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2022. doi: 10.1109/TPAMI.2022.3148795.

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

Pulak Purkait, Tat-Jun Chin, Alireza Sadri, and David Suter. 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.

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

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

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

David Stein, Silvia Di Gregorio, and Bjoern Andres. Partial optimality in cubic correlation clustering. In *ICML*, 2023. URL `https://proceedings.mlr.press/v202/stein23a.html`.

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

Robert Endre Tarjan and Jan van Leeuwen. Worst-case analysis of set union algorithms. *Journal of the ACM*, 31(2):245–281, 1984. doi: 10.1145/62.2160.

Nate Veldt. 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`.