Industrial and
Systems Engineering

# On Relaxations of the Max $k$-cut Formulations

Ramin Fakhimi[1], Hamidreza Validi[2], Illya V. Hicks[2], Tamás Terlaky[1], and Luis F. Zuluaga[1]

[1]Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA
[2]Department of Industrial, Manufacturing & Systems Engineering, Texas Tech University, TX
[3]Department of Computational and Applied Mathematics, Rice University, Houston, TX

ISE Technical Report 22T-018

LEHIGH UNIVERSITY

# On relaxations of the max $k$-cut problem formulations

Ramin Fakhimi[a], Hamidreza Validi[b], Illya V. Hicks[c], Tamás Terlaky[a], Luis F. Zuluaga[a]

[a] *Industrial and Systems Engineering, Lehigh University, PA, U.S.*
[b] *Industrial, Manufacturing & Systems Engineering, Texas Tech University, TX, U.S.*
[c] *Computational Applied Mathematics & Operations Research, Rice University, TX, U.S.*

## Abstract

A tight continuous relaxation is a crucial factor in solving mixed integer formulations of many NP-hard combinatorial optimization problems. The (weighted) max $k$-cut problem is a fundamental combinatorial optimization problem with multiple notorious mixed integer optimization formulations. In this paper, we explore four existing mixed integer optimization formulations of the max $k$-cut problem and show that the continuous relaxation of a binary quadratic optimization formulation is: (i) stronger than that of two mixed integer linear optimization formulations and (ii) at least as strong as a mixed integer semidefinite optimization formulation. We also conduct a set of experiments on the state-of-the-art solvers to assess the theoretical results in practice. The computational results support our theoretical findings on multiple sets of instances. Our codes and data are available on GitHub.

*Keywords:* the max $k$-cut problem, mixed integer optimization, semidefinite optimization, continuous relaxation

## 1. Introduction

The continuous relaxation of a mixed integer optimization formulation plays a fundamental role in the efficient solution process of not only linear formulations, but also non-linear ones of a mixed integer optimization problem [1]. The (weighted) max $k$-cut problem is a fundamental NP-hard combinatorial optimization problem [2, 3] with multiple mixed integer linear optimization formulations that suffer either from weak relaxation or large size. The max $k$-cut problem has a wide range of applications, including but not limited to statistical physics [4, 5], gas and power networks [6], data clustering [7], and scheduling [8]. Given a graph $G = (V, E)$ with edge weights $w$ and a positive integer number $k \geq 2$, the max $k$-cut problem seeks to find at most $k$ partitions such that the weights of edges with endpoints in different partitions are maximized.

Motivated by the considerable effect of the continuous relaxation strength in solving mixed integer optimization formulations of the max $k$-cut to optimality, we discuss multiple known optimization formulations of the max $k$-cut problem in the literature: (i) a binary quadratic optimization (BQO) formulation [8]; (ii) a vertex-based mixed integer linear optimization (V-MILO) formulation; (iii) an edge-based mixed integer linear optimization (E-MILO) formulation [9]; and (iv) a mixed integer semidefinite optimization (MISDO) formulation [10]. We prove that the continuous relaxation of the BQO formulation is: (i) stronger than that of the V-MILO and the E-MILO formulations and (ii) at least as strong as a mixed integer semidefinite optimization formulation. Further, we conduct a set of computational experiments to assess our theoretical results in practice. Thanks to the recent advancements of state-of-the-art solvers (e.g., Gurobi), the numerical results support most of the theoretical ones. The continuous relaxation of the BQO formulation provides a tighter upper bound compared to the other MILO formulations. It also provides a high-quality upper bound for large-scale instances of the problem that the continuous relaxation of the MISDO formulation struggles to achieve.

## 2. Mixed Integer Optimization Formulations

Motivated by solving a scheduling problem, Carlson and Nemhauser [8] proposed a BQO formulation for the max $k$-cut problem. Let $n \coloneqq |V|$ and $m \coloneqq |E|$ respectively be the number of vertices and edges of graph $G = (V, E)$. Furthermore, we define $P \coloneqq \{1, \ldots, k\}$ as the set of partitions and $w_{uv} \in \mathbb{R}$ as the edge weights for $\{u, v\} \in E$. For every vertex $v \in V$ and every partition $j \in P$, binary variable $x_{vj}$ is one if vertex $v$ is assigned to partition $j$ and zero otherwise. Then, the BQO formulation is as follows.

$$
\begin{aligned}
& \max && \sum_{\{u,v\}\in E} w_{uv}\left(1-\sum_{j\in P} x_{uj}x_{vj}\right) && && \text{(1a)}\\
\text{(BQO)}\quad & \text{s.t.} && \sum_{j\in P} x_{vj} = 1 && \forall v\in V && \text{(1b)}\\
& && x\in\{0,1\}^{n\times k}. && && \text{(1c)}
\end{aligned}
$$

Objective function (1a) maximizes the number of cut edges, and constraints (1b) imply that each vertex must be assigned to exactly one partition. Carlson and Nemhauser [8] proved that an optimal solution of the continuous relaxation of BQO formulation (1) can be converted into an optimal solution of its binary variant with the same objective value.

**Theorem 1** (Carlson and Nemhauser [8])**.** *An optimal solution of the* BQO *formulation* (1) *is also optimal for its continuous relaxation.*

One can linearize the BQO formulation (1) to develop a MILO formulation of the max $k$-cut problem that is called the vertex-based MILO (V-MILO) formulation in this paper. For every edge $\{u,v\}\in E$, binary variable $y_{uv}$ is one if the endpoints of edge $\{u,v\}$ belong to different partitions; that is $\{u,v\}$ is a *cut edge*, and zero otherwise.

$$
\begin{aligned}
& \max && \sum_{\{u,v\}\in E} w_{uv}y_{uv} && && \text{(2a)}\\
& \text{s.t.} && \sum_{j\in P} x_{vj} = 1 && \forall v\in V && \text{(2b)}\\
\text{(V-MILO)}\quad & && x_{uj}-x_{vj}\le y_{uv} && && \\
& && x_{vj}-x_{uj}\le y_{uv} && \forall\{u,v\}\in E,\ j\in P && \text{(2c)}\\
& && x_{uj}+x_{vj}+y_{uv}\le 2 && \forall\{u,v\}\in E,\ j\in P && \text{(2d)}\\
& && x\in\{0,1\}^{n\times k} && && \text{(2e)}\\
& && y\in\{0,1\}^{m}. && && \text{(2f)}
\end{aligned}
$$

Objective function (2a) maximizes the total weight of cut edges. Constraints (2b) imply that every vertex is assigned to exactly one partition. Constraints (2c) imply that if endpoints of an edge belong to different partitions, then it is a cut edge. Constraints (2d) imply that if the endpoints of an edge belong to the same partition, then the edge cannot be a cut edge. Despite the reasonable size of formulation (2) having $kn + m$ variables and $n+3km$ constraints, it suffers from weak continuous relaxation and symmetry issues [6].

Another classical MILO formulation is a large edge-based MILO (E-MILO) formulation with $\binom{n}{2}$ variables and $3\binom{n}{3} + \binom{n}{k+1}$ constraints [9, 11]. Although the continuous relaxation of this formulation provides a relatively tight upper bound in practice, classical solvers struggle to solve even medium-size instances of the max $k$-cut problem to optimality [12]. For every set $S$, we employ $\binom{S}{2}$ to denote all subsets of $S$ with size 2. For every pair of vertices $\{u, v\} \in \binom{V}{2}$, we define binary variable $z_{uv}$ as follows: $z_{uv}$ is one if vertices $u$ and $v$ belong to the same partition, and zero otherwise.

$$\max \quad \sum_{\{u,v\}\in E} w_{uv}(1 - z_{uv}) \tag{3a}$$

$$\text{s.t.} \quad z_{uv} + z_{vw} \leq 1 + z_{uw}$$

$$z_{uw} + z_{uv} \leq 1 + z_{vw}$$

$$\text{(E-MILO)} \quad z_{vw} + z_{uw} \leq 1 + z_{uv} \qquad \forall \{u, v, w\} \subseteq V \tag{3b}$$

$$\sum_{\{u,v\}\in\binom{Q}{2}} z_{uv} \geq 1 \qquad \forall Q \subseteq V, |Q| = k + 1 \tag{3c}$$

$$z \in \{0, 1\}^{\binom{n}{2}}. \tag{3d}$$

Objective function (3a) maximizes the total weight of cut edges. Constraints (3b) imply that for every set $\{u, v, w\} \subseteq V$, if pairs $\{u, v\}$ and $\{v, w\}$ belong to a partition, then vertices $u$ and $w$ also belong to the same partition. Constraints (3c) imply that vertex set $V$ must be partitioned into at most $k$ partitions. Because of the large number of constraints (3c), one can add them on the fly [13]. Chopra and Rao [11] conducted a polyhedral study on the max $k$-cut problem and proposed several facet-defining inequalities for the E-MILO formulation. They also studied the V-MILO and E-MILO formulations for the min $k$-cut problem and developed multiple facet-defining inequalities for both formulations [9].

Further, Wang and Hijazi [12] propose a reduced E-MILO (RE-MILO) formulation that is constructed as follows: (i) graph $G$ is extended to a chordal graph, (ii) all maximal cliques of the chordal graph are found, (iii) binary variables $z$ are created *only* for the edge set of the chordal graph, and (iv) constraints (3b)–(3c) are added *only* for the maximal cliques. The number of variables and constraints in their formulation is fewer than or equal to that of the E-MILO formulation (3). However, for dense graphs in which the chordalized graph is complete, they are the same as the E-MILO formulation. They show that their formulation outperforms the E-MILO formulation when the chordalized graph is sparse.

We also provide an existing mixed integer semidefinite optimization (MISDO) formulation [14]. For every $(u, v) \in V \times V$, binary variable $Z_{uv}$ is one if vertices $u$ and $v$ belong to the same partition. Then the formulation is as follows.

$$\text{(MISDO)} \quad \begin{aligned} \max \quad & \sum_{\{u,v\}\in E} w_{uv}(1 - Z_{uv}) && && \text{(4a)} \\ \text{s.t.} \quad & Z_{vv} = 1 && \forall v \in V && \text{(4b)} \\ & kZ \succeq ee^T && && \text{(4c)} \\ & Z \in \{0,1\}^{n\times n}. && && \text{(4d)} \end{aligned}$$

Interested readers are encouraged to refer to [15, 16, 12, 17] for more details on semidefinite optimization and mixed integer semidefinite optimization formulations of the max $k$-cut.

## 3. A Theoretical Comparison of Relaxations

In this section, we provide theoretical comparisons between the continuous relaxations of BQO formulation (1) and formulations (2)-(4). For analysis purposes, we introduce $y$ variables to the BQO formulation (1): for every edge $\{u, v\} \in E$, variable $y_{uv}$ equals one if $\{u, v\}$ is a cut edge.

$$y_{uv} = 1 - \sum_{j\in P} x_{uj}x_{vj} \qquad \forall \{u, v\} \in E. \tag{5}$$

Furthermore, we define the set of lifted continuous relaxation of the BQO formulation as follows.

$$\mathcal{R}^{\mathrm{y}}_{\mathrm{BQO}} := \Big\{ (x, y) \in [0, 1]^{n\times k} \times \mathbb{R}^m \;\Big|\; (x, y) \text{ satisfies constraints (1b) and (5)} \Big\}.$$

The following remark shows that we do not need to impose 0-1 bounds on $y$ variables.

**Remark 1.** *Constraints $y \in [0,1]^m$ are implied by the* BQO *formulation* (1) *and constraints* (5).

To see this, consider a point $(\hat{x}, \hat{y}) \in \mathcal{R}^{\mathrm{y}}_{\mathrm{BQO}}$. For every edge $\{u, v\} \in E$, we have

$$\hat{y}_{uv} = 1 - \sum_{j \in P} \hat{x}_{uj}\hat{x}_{vj} \geq 1 - \sum_{j \in P} \hat{x}_{uj} = 1 - 1 = 0.$$

The first equality holds by constraints (5). The inequality holds because, for any partition $j \in P$, we have $x_{vj} \leq 1$. The second equality holds by constraints (1b). Furthermore, we have

$$\hat{y}_{uv} = 1 - \sum_{j \in P} \hat{x}_{uj}\hat{x}_{vj} \leq 1 - 0 = 1.$$

The first equality holds by constraints (5). The inequality holds because for any partition $j \in P$, and any vertex $v \in V$, we have $x_{vj} \geq 0$.

First, we prove Lemma 1 that will be used in our further analyses. We define $[n] \coloneqq \{1, \dots, n\}$ for every $n \in \mathbb{Z}_{++}$.

**Lemma 1.** *Let $a \in [0,1]^n$ with $n \geq 2$. Then, we have*

$$1 - \sum_{i \in [n]} a_i + \sum_{\{i,j\} \in \binom{[n]}{2}} a_i a_j \geq 0. \tag{6}$$

*Proof.* We prove the claim by induction. First, we show that the inequality holds for the base case $n = 2$. In this case, we have

$$1 - a_1 - a_2 + a_1 a_2 = (1 - a_1)(1 - a_2) \geq 0. \tag{7}$$

The inequality (7) holds because for every $i \in \{1, 2\}$, we have $1 - a_i \geq 0$.

Now suppose that inequality (6) holds for $n = s \geq 2$ (induction hypothesis). It suffices to show that it also holds for $n = s + 1$.

$$0 \leq \left(1 - \sum_{i \in [s]} a_i + \sum_{\{i,j\} \in \binom{[s]}{2}} a_i a_j\right)(1 - a_{s+1}) \tag{8a}$$

$$= 1 - \sum_{i \in [s+1]} a_i + \sum_{\{i,j\} \in \binom{[s+1]}{2}} a_i a_j - a_{s+1} \sum_{\{i,j\} \in \binom{[s]}{2}} a_i a_j \tag{8b}$$

$$\leq 1 - \sum_{i \in [s+1]} a_i + \sum_{\{i,j\} \in \binom{[s+1]}{2}} a_i a_j. \tag{8c}$$

Inequality (8a) holds by induction hypothesis and because $1 - a_{s+1} \geq 0$. Inequality (8c) holds as $-a_{s+1} \sum_{\{i,j\} \in \binom{[s]}{2}} a_i a_j \leq 0$. □

Now we define the polytope of the continuous relaxation of the V-MILO formulation (2) as follows.

$$\mathcal{R}_{\text{V-MILO}} := \Big\{ (x, y) \in [0,1]^{n \times k} \times [0,1]^m \;\Big|\; (x, y) \text{ satisfies constraints (2b)–(2d)} \Big\}.$$

Theorem 2 shows that the continuous relaxation of the BQO formulation is stronger than that of the V-MILO formulation.

**Theorem 2.** $\mathcal{R}^{\text{y}}_{\text{BQO}} \subset \mathcal{R}_{\text{V-MILO}}$.

*Proof.* Let point $(\hat{x}, \hat{y}) \in \mathcal{R}^{\text{y}}_{\text{BQO}}$. First, we are to show that $(\hat{x}, \hat{y}) \in \mathcal{R}_{\text{V-MILO}}$. We show that $(\hat{x}, \hat{y})$ satisfies constraints (2c). For every edge $\{u, v\} \in E$ and every partition $j \in P$, we have

$$\hat{y}_{uv} = 1 - \sum_{i \in P} \hat{x}_{ui} \hat{x}_{vi} \tag{9a}$$
$$= \sum_{i \in P} \hat{x}_{ui} - \sum_{i \in P} \hat{x}_{ui} \hat{x}_{vi} \tag{9b}$$
$$= \hat{x}_{uj} + \sum_{i \in P \setminus \{j\}} \hat{x}_{ui} - \hat{x}_{uj} \hat{x}_{vj} - \sum_{i \in P \setminus \{j\}} \hat{x}_{ui} \hat{x}_{vi} \tag{9c}$$
$$\geq \hat{x}_{uj} + \sum_{i \in P \setminus \{j\}} \hat{x}_{ui} - \hat{x}_{vj} - \sum_{i \in P \setminus \{j\}} \hat{x}_{ui} \hat{x}_{vi} \tag{9d}$$
$$= \hat{x}_{uj} - \hat{x}_{vj} + \sum_{i \in P \setminus \{j\}} \hat{x}_{ui} (1 - \hat{x}_{vi}) \tag{9e}$$
$$\geq \hat{x}_{uj} - \hat{x}_{vj}. \tag{9f}$$

Equality (9a) holds by constraints (5). Equality (9b) follows from constraint (1b). Inequality (9d) holds because $\hat{x}_{uj} \leq 1$. Inequality (9f) holds because $\sum_{i \in P \setminus \{j\}} \hat{x}_{ui} (1 - \hat{x}_{vi}) \geq 0$.

Finally, we show that $(\hat{x}, \hat{y})$ satisfies constraints (2d). For every edge $\{u, v\} \in E$ and every partition $j \in P$, we have

$$\hat{y}_{uv} = 1 - \sum_{i \in P} \hat{x}_{ui}\hat{x}_{vi} \tag{10a}$$
$$= 1 - \hat{x}_{uj}\hat{x}_{vj} - \sum_{i \in P \setminus \{j\}} \hat{x}_{ui}\hat{x}_{vi} \tag{10b}$$
$$\leq 2 - 1 - \hat{x}_{uj}\hat{x}_{vj} \tag{10c}$$
$$\leq 2 - (\hat{x}_{uj} + \hat{x}_{vj}). \tag{10d}$$

Equality (10a) holds by constraints (5). Inequality (10c) holds by inequality (7) in Lemma 1 and $\sum\limits_{i \in P \setminus \{j\}} \hat{x}_{ui}\hat{x}_{vi} \geq 0$.

Now, we are to show that there exists a point $(\hat{x}, \hat{y}) \in \mathcal{R}_{\text{V-MILO}}$ such that $(\hat{x}, \hat{y}) \notin \mathcal{R}^{\text{y}}_{\text{BQO}}$. For every $v \in V$, let $\hat{x}_{v1} = \hat{x}_{v2} = 0.5$. For every vertex $v \in V$ and every partition $j \in \{3, 4, \ldots, k\}$, we define $\hat{x}_{vj} = 0$. Furthermore, for every edge $\{u, v\} \in E$, we define $\hat{y}_{uv} = 1$. It is simple to check that $(\hat{x}, \hat{y})$ in $\mathcal{R}_{\text{V-MILO}}$. So, point $(\hat{x}, \hat{y}) \in \mathcal{R}_{\text{V-MILO}} \setminus \mathcal{R}^{\text{y}}_{\text{BQO}}$ because $(\hat{x}, \hat{y})$ violates constraints (5). Thus, the proof is complete. □

The following remark shows the V-MILO formulation has a weak relaxation.

**Remark 2.** *The optimal objective of the continuous relaxation for the* V-MILO *formulation* (2) *is equal to* $\sum_{\{u,v\} \in E} \max\{w_{uv}, 0\}$.

To see this, note that an optimal solution $(x^*, y^*)$ for the continuous relaxation of the V-MILO formulation (2) is obtained by setting $x^*_{vj} = \frac{1}{k}$ for every $v \in V$ and $j \in P$. Also for every edge $\{u, v\} \in E$, we set $y^*_{uv}$ to 1 if $w_{uv} > 0$ and 0 otherwise.

To conduct a theoretical comparison between the continuous relaxations of the BQO and E-MILO formulations, we lift the dimensionality of the BQO by introducing new $z$ variables.

$$z_{uv} \coloneqq \sum_{j \in P} x_{uj}x_{vj}, \qquad \forall \{u, v\} \in \binom{V}{2}. \tag{11}$$

We define the set of lifted continuous relaxation of the BQO formulation in

$z$-space as follows.

$$\mathcal{R}^{\mathrm{z}}_{\mathrm{BQO}} \coloneqq \Big\{ (x,z) \in [0,1]^{n\times k} \times \mathbb{R}^{\binom{n}{2}} \;\Big|\; (x,z) \text{ satisfies constraints (1b) and (11)} \Big\}.$$

We also define the polytope of the E-MILO formulation as follows.

$$\mathcal{R}_{\text{E-MILO}} \coloneqq \Big\{ z \in [0,1]^{\binom{n}{2}} \;\Big|\; z \text{ satisfies constraints (3b)–(3c)} \Big\}.$$

We show that the continuous relaxation of a projection of the lifted BQO formulation on the $z$ space is stronger than that of the E-MILO formulation.

**Theorem 3.** $\operatorname{proj}_z \mathcal{R}^{\mathrm{z}}_{\mathrm{BQO}} \subset \mathcal{R}_{\text{E-MILO}}$ *for* $n > k$*.*

*Proof.* Consider a point $(\hat{x}, \hat{z}) \in \mathcal{R}^{\mathrm{z}}_{\mathrm{BQO}}$. We are to show that $\hat{z} \in \mathcal{R}_{\text{E-MILO}}$. For every set $\{u, v, w\} \subseteq V$, we show that point $\hat{z}$ satisfies constraints (3b).

$$\begin{aligned}
\hat{z}_{uv} + \hat{z}_{vw} &= \sum_{j\in P} \hat{x}_{uj}\hat{x}_{vj} + \sum_{j\in P} \hat{x}_{vj}\hat{x}_{wj} && \text{(12a)}\\
&= \sum_{j\in P} \hat{x}_{vj}(\hat{x}_{uj} + \hat{x}_{wj}) && \text{(12b)}\\
&\leq \sum_{j\in P} \hat{x}_{vj}(1 + \hat{x}_{uj}\hat{x}_{wj}) && \text{(12c)}\\
&= \sum_{j\in P} \hat{x}_{vj} + \sum_{j\in P} \hat{x}_{vj}(\hat{x}_{uj}\hat{x}_{wj}) && \text{(12d)}\\
&= 1 + \sum_{j\in P} \hat{x}_{vj}(\hat{x}_{uj}\hat{x}_{wj}) && \text{(12e)}\\
&\leq 1 + \sum_{j\in P} \hat{x}_{uj}\hat{x}_{wj} && \text{(12f)}\\
&= 1 + \hat{z}_{uw}. && \text{(12g)}
\end{aligned}$$

Equality (12a) holds by definition (11). Inequality (12c) holds by inequality (7) in Lemma 1. Equality (12e) holds by constraints (1b). Inequality (12f) holds by the fact that $\hat{x}_{vj} \leq 1$. Equality (12g) holds by definition (11).

Furthermore, we show that point $\hat{z}$ satisfies constraints (3c). For every vertex set $Q \subseteq V$ with $|Q| = k+1$, we have

$$
\begin{aligned}
\sum_{\{u,v\}\in\binom{Q}{2}} \hat{z}_{uv} &= \sum_{\{u,v\}\in\binom{Q}{2}} \sum_{j\in P} \hat{x}_{uj}\hat{x}_{vj} && \text{(13a)}\\
&= \sum_{j\in P}\left(\sum_{\{u,v\}\in\binom{Q}{2}} \hat{x}_{uj}\hat{x}_{vj}\right) && \text{(13b)}\\
&\geq \sum_{j\in P}\left(\sum_{u\in Q} \hat{x}_{uj} - 1\right) && \text{(13c)}\\
&= \sum_{u\in Q}\sum_{j\in P} \hat{x}_{uj} - k && \text{(13d)}\\
&= k+1-k = 1. && \text{(13e)}
\end{aligned}
$$

Equality (13a) holds by definition (11). Inequality (13c) holds by Lemma 1. Equality (13e) holds by constraints (1b) and because $|Q| = k+1$.

Finally, for every $\{u,v\} \in \binom{V}{2}$, we show that $0 \leq \hat{z}_{uv} \leq 1$. Because for every vertex $v \in V$ and every partition $j \in P$ we have $\hat{x}_{vj} \geq 0$, it follows that $\hat{z}_{uv} \geq 0$. For every $\{u,v\} \in \binom{V}{2}$, we show that $\hat{z}_{uv} \leq 1$.

$$\hat{z}_{uv} = \sum_{j\in P} \hat{x}_{uj}\hat{x}_{vj} \leq \sum_{j\in P} \hat{x}_{uj} = 1.$$

The first equality holds by definition (11). The inequality holds because $\hat{x}_{vj} \leq 1$ for every vertex $v \in V$ and every partition $j \in P$. The last equality holds by constraints (1b). This implies that $\operatorname{proj}_z \mathcal{R}^{\mathrm{z}}_{\mathrm{BQO}} \subseteq \mathcal{R}_{\text{E-MILO}}$.

Now we show that the inclusion is strict for any non-trivial instance of the max $k$-cut problem with $n > k$. For every $\{u,v\} \in \binom{V}{2}$, we define $\hat{z}_{uv}$ as a point that belongs to the polytope of the E-MILO formulation; that is, $\hat{z} \in \mathcal{R}_{\text{E-MILO}}$.

$$\hat{z}_{uv} := \frac{2}{k(k+1)}.$$

For every vertex $v \in V$, let $\mathbf{x}_v \in [0,1]^k$ be the assignment vector of vertex $v$. By definition (11), we have

$$\hat{z}_{uv} = \hat{\mathbf{x}}_u^T \hat{\mathbf{x}}_v = \|\hat{\mathbf{x}}_u\|_2 \|\hat{\mathbf{x}}_v\|_2 \cos\hat{\theta}_{uv}. \tag{14}$$

By constraints (1b), we have $\|\hat{\mathbf{x}}_v\|_1 = 1$. Then for every vertex $v \in V$, we have

$$\frac{1}{\sqrt{k}} \leq \|\hat{\mathbf{x}}_v\|_2 \leq 1. \tag{15}$$

The first inequality holds because $\|\hat{\mathbf{x}}_v\|_2$ reaches its minimum when $\hat{x}_{vj} = \frac{1}{k}$ for every partition $j \in P$.

By lines (14) and (15), we have

$$\frac{2}{k(k+1)} \leq \cos \hat{\theta}_{uv} \leq \frac{2}{k+1}.$$

For every $\{u, v\} \in \binom{V}{2}$, this implies that we have the following relations because $k \geq 2$.

$$\arccos\left(\frac{1}{\sqrt{k}}\right) < \arccos\left(\frac{2}{k+1}\right) \leq \hat{\theta}_{uv} \leq \arccos\left(\frac{2}{k(k+1)}\right). \tag{16}$$

Consider $k+1$ vectors in $\mathbb{R}^k_+$ and let $\theta_{\min}$ be the minimum angle between all vector pairs. It follows that the maximum value of $\theta_{\min}$ is $\arccos\left(\frac{1}{\sqrt{k}}\right)$. This case happens when $k$ vectors are located on the axes, and one vector is located at the center of the positive orthant. Without loss of generality, consider $k$ unit vectors on $k$ different axes in $\mathbb{R}^k_+$ and a vector with all entries equal to $\frac{1}{\sqrt{k}}$. For example, the maximum values of $\theta_{\min}$ are $45°$ and $\arccos\left(\frac{1}{\sqrt{3}}\right) \approx 54.7°$ for $k = 2$ and $k = 3$, respectively.

Since all vectors $\hat{\mathbf{x}}_v$ are in the positive orthant and $n > k$, there are vectors $\hat{\mathbf{x}}_a$ and $\hat{\mathbf{x}}_b$ with $\hat{\theta}_{ab} \leq \arccos\left(\frac{1}{\sqrt{k}}\right)$. However, this contradicts the first inequality of line (16). Thus, there is no feasible solution of the BQO formulation that satisfies definition (11). This completes the proof. □

Wang and Hijazi [12] prove that their reduced E-MILO (RE-MILO) formulation is as strong as the projection of the E-MILO formulation on the edges of an extended chordal graph. Further, their computational experiments show the superiority of the RE-MILO over the E-MILO for sparse chordal graphs.

Now we define the continuous relaxation of the MISDO formulation (4) as follows.

$$\mathcal{R}_{\text{MISDO}} \coloneqq \left\{ Z \in [0,1]^{n \times n} \;\middle|\; Z \text{ satisfies constraints (4b)–(4c)} \right\}.$$

To conduct a theoretical comparison between the continuous relaxations of the BQO and MISDO formulations, we lift the dimensionality of the BQO formulation by introducing a new symmetric matrix $Z \in \mathbb{R}^{n\times n}$ defined as follows.

$$Z \coloneqq D^x + \sum_{j\in P} \mathbf{x}_j \mathbf{x}_j^T, \tag{17}$$

with $D^x$ be a diagonal matrix and $D^x_{vv} = 1 - \sum_{j\in P} x^2_{vj}$. For every partition $j \in P$, we redefine vector $\mathbf{x}_j \in [0,1]^n$ such that $\mathbf{x}_{jv} = x_{vj}$. For comparison purposes, we also define $\mathcal{R}^{\mathrm{Z}}_{\mathrm{BQO}}$.

$$\mathcal{R}^{\mathrm{Z}}_{\mathrm{BQO}} \coloneqq \Big\{ (x, Z) \in [0,1]^{n\times k} \times \mathbb{R}^{n\times n} \;\Big|\; (x, Z) \text{ satisfies constraints (1b) and (17)} \Big\}.$$

Eisenblätter [10] developed a semidefinite formulation for the max $k$-cut problem and showed that its continuous relaxation is strong. However, they declare that continuous relaxations of the semidefinite optimization and E-MILO formulations have exclusive points. Furthermore, de Sousa et al. [18] propose MISDO-based constraints for the E-MILO formulation to strengthen its relaxation. The following theorem compares the relaxation strength of the MISDO formulation against an extension of the BQO formulation.

**Theorem 4.** $\mathrm{proj}_Z\, \mathcal{R}^{\mathrm{Z}}_{\mathrm{BQO}} \subseteq \mathcal{R}_{\mathrm{MISDO}}$.

*Proof.* For any fractional solution $x \in [0,1]^{n\times k}$, we have $\sum_{j\in P} x^2_{vj} \leq 1$ by constraints (1b). Thus, $kD^x \succeq 0$ and $Z \in [0,1]^{n\times n}$ by definition (17). Matrix $k\sum_{j\in P} \mathbf{x}_j\mathbf{x}_j^T - ee^T$ is positive-semidefinite if and only if for every vector $b \in \mathbb{R}^n$, we have

$$b^T\Bigg(k\sum_{j\in P} \mathbf{x}_j\mathbf{x}_j^T - ee^T\Bigg)b \geq 0.$$

We define

$$\beta_j \coloneqq b^T\mathbf{x}_j \;\; \forall j \in P, \text{ and } \alpha \coloneqq b^T e. \tag{18}$$

It suffices to show $k\sum_{j\in P} \mathbf{x}_j\mathbf{x}_j^T - ee^T \succeq 0$. We can rewrite constraints (1b) as $\sum_{j\in P} \mathbf{x}_j = e$. Thus, we have

$$\sum_{j\in P} \beta_j = \sum_{j\in P} b^T\mathbf{x}_j = b^T \sum_{j\in P} \mathbf{x}_j = b^T e = \alpha. \tag{19}$$

So, we have that

$$
\begin{aligned}
b^T \left( k \sum_{j \in P} \mathbf{x}_j \mathbf{x}_j^T - ee^T \right) b &= k \sum_{j \in P} \beta_j^2 - \alpha^2 \\
&= k \sum_{j \in P} \beta_j^2 - \left( \sum_{j \in P} \beta_j \right)^2 \\
&= \sum_{\{i,j\} \in \binom{P}{2}} \left( \beta_i - \beta_j \right)^2 \geq 0.
\end{aligned}
$$

The first equality holds by definitions (18). The second equality holds by line (19). This completes the proof. □

## 4. Computational Experiments

In this section, we conduct a set of experiments to evaluate our theoretical results in practice. In better words, we compare the relaxations of the discussed formulations using state-of-the-art solvers. We run the computational experiments on a machine with Dual Intel Xeon® CPU E5-2630 @ 2.20 GHz (20 cores) and 64 GB of RAM. We have developed the Python package `MaxKcut` [19] to conduct the computational experiments. We employ Gurobi 10.0.0 [13] to solve our mixed integer optimization formulations. We also use MOSEK 10.0.27 [20] to run our experiments for the MISDO formulation. In both Gurobi and MOSEK solvers, we set the number of threads and time limit to 10 and 3,600 seconds, respectively.

We run our experiments on the following sets of instances: (i) band [12], (ii) spinglass [12], (iii) Color02 [21], and (iv) Steiner-160 [22] (in total 97 different instances). We classify these instances into different classes based on the number of vertices (from 50 to 250) and their density (from 5 percent to 100 percent). Thanks to the spatial branch and bound algorithm, Gurobi can solve a nonconvex formulation to global-$\epsilon$ optimality; however, we stop the solving process whenever the solver reaches the one-hour time limit. We scale the upper bounds by the best obtained upper bound to report the results. Furthermore, we use the geometric mean of the scaled upper bound for every batch of instances. The code, instances, and results are available on GitHub [19].

Because of the large number of clique constraints in the RE-MILO formulation (i.e., a sparse variant of the E-MILO model (3) proposed by Wang and Hijazi [12]) for large-scale instances, it is not practical to add all of them upfront. Thus, we initially relax the clique constraints in the RE-MILO formulation to avoid memory shortage. After solving the relaxed formulation to optimality, we iteratively add the most violated constraints using the dual Simplex method and solve the problem. We stop whenever there is not enough free memory. For small instances with $n \leq 100$, all the clique constraints are added upfront. We note that the solving process of all instances with $n > 100$ reaches the time limit.

Figure 1 provides a summary of our results with instances on the horizontal axis and the geometric mean of the scaled relaxed objective value on the vertical axis. For $k \in \{3, 4\}$, we solve the following models.

(i) The BQO formulation (1);

(ii) the continuous relaxation of the BQO formulation (1);

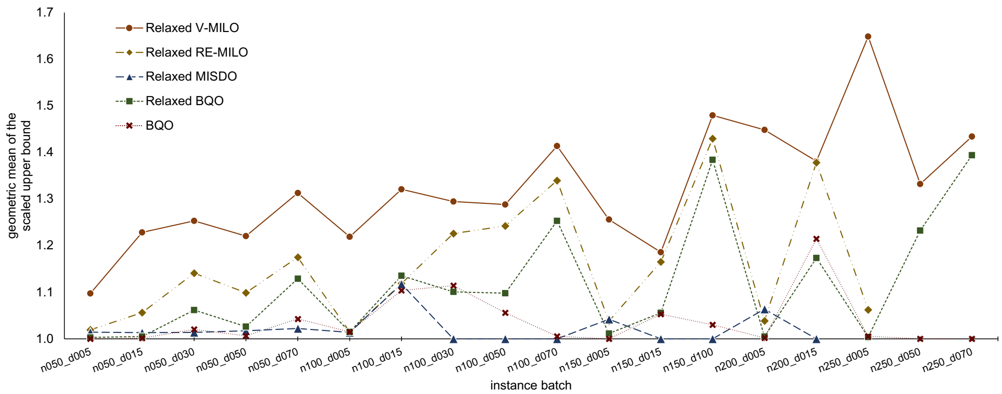
(iii) the continuous relaxation of the V-MILO formulation (2);

(iv) the continuous relaxation of the RE-MILO formulation;

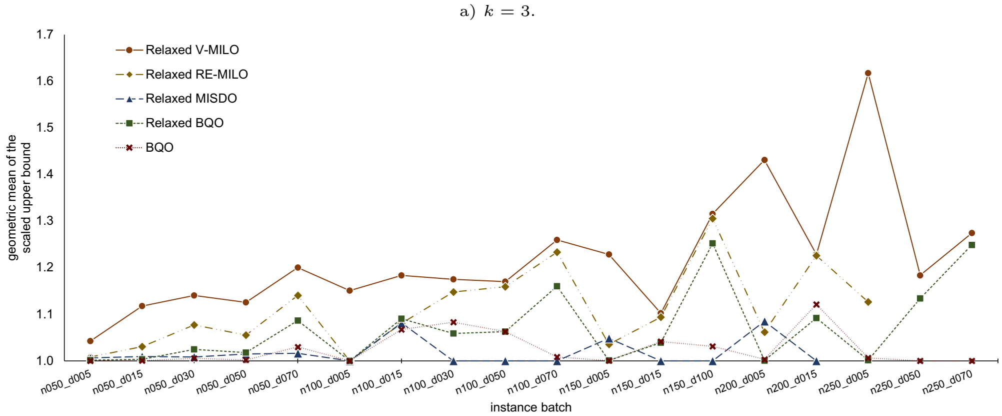
(v) the continuous relaxation of the MISDO formulation (4).

Figure 1 shows the superiority of the BQO formulation and its relaxed variant over both V-MILO and RE-MILO formulations when $k \in \{3, 4\}$. This observation matches the results of Theorems 2 and 3. We also observe that the upper bound obtained by solving the BQO formulation (1) outperforms its continuous relaxation for most sets of instances while they both have the same optimal objective value by Theorem 1.

Figure 1 illustrates the inferiority of the relaxation of the V-MILO formulation (2) over all other formulations for all sets of instances when $k \in \{3, 4\}$. This behavior is justifiable by Remark 2. We also observe that the relaxed MISDO formulation (4) provides a tight convex relaxation. For sparse instances, it performs similarly to the relaxed RE-MILO formulation. MOSEK employs the *interior point method* (IPM) to solve the semidefinite optimization formulations [20]. For a given instance with $n$ vertices, the IPM requires the solution of a linear system in $\mathbb{R}^{\mathcal{O}(n^2) \times \mathcal{O}(n^2)}$ at every iteration. These extremely large linear systems exhaust all the memory as the algorithm converges to an optimal solution, and their condition numbers grow [23]. Regardless of the decent performance of the MISDO formulation (4) on small and medium-sized instances, it is not practical for solving instances with more than 200 vertices.




a) $k = 3$.




b) $k = 4$.

Figure 1: The geometric mean of the scaled upper bound obtained by different methods in an hour time limit ($n$ represents the number of vertices, and $d$ denotes the graph density in percentage).

## 5. Conclusion

Motivated by the importance of continuous relaxation in the solution process of the mixed integer optimization formulations of NP-hard problems, we compared the continuous relaxations of four well-known formulations of a fundamental combinatorial optimization problem, that is the max $k$-cut problem. We proved that the continuous relaxation of a binary quadratic optimization formulation is tighter than other existing formulations; specifically, vertex-based and edge-based mixed integer linear optimization formulations. Interestingly, we observed that our numerical experiments support most of the theoretical ones. As a future work, one might be interested in comparing the nonlinear optimization formulations of other combinatorial optimization problems with convex counterparts. While many believe that convex formulations may outperform in practice, our results show that this might not be true any more thanks to recent advances in non-convex optimization solvers.

### Acknowledgments

This work is supported by the Defense Advanced Research Projects Agency (DARPA), ONISQ grant W911NF2010022 titled *The Quantum Computing Revolution and Optimization: Challenges and Opportunities*. This research also used resources from the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

### References

[1] B. Müller, G. Muñoz, M. Gasse, A. Gleixner, A. Lodi, F. Serrano, On generalized surrogate duality in mixed-integer nonlinear programming, Mathematical Programming 192 (2022) 89–118.

[2] A. Frieze, M. Jerrum, Improved approximation algorithms for MAX $k$-CUT and MAX BISECTION, Algorithmica 18 (1997) 67–81. doi:10.1007/BF02523688.

[3] C. H. Papadimitriou, M. Yannakakis, Optimization, approximation, and complexity classes, Journal of Computer and System Sciences 43 (1991) 425–440. doi:10.1016/0022-0000(91)90023-X.

[4] F. Barahona, M. Grötschel, M. Jünger, G. Reinelt, An application of combinatorial optimization to statistical physics and circuit layout design, Operations Research 36 (1988) 493–513. doi:10.1287/opre.36.3.493.

[5] C. De Simone, M. Diehl, M. Jünger, P. Mutzel, G. Reinelt, G. Rinaldi, Exact ground states of ising spin glasses: New experimental results with a branch-and-cut algorithm, Journal of Statistical Physics 80 (1995) 487–496. doi:10.1007/BF02178370.

[6] C. Hojny, I. Joormann, H. Lüthen, M. Schmidt, Mixed-integer programming techniques for the connected max-$k$-cut problem, Mathematical Programming Computation (2020) 1–58. doi:10.1007/s12532-020-00186-3.

[7] J. Poland, T. Zeugmann, Clustering pairwise distances with missing data: Maximum cuts versus normalized cuts, in: International Conference on Discovery Science, Springer, 2006, pp. 197–208. doi:10.1007/11893318_21.

[8] R. C. Carlson, G. L. Nemhauser, Scheduling to minimize interaction cost, Operations Research 14 (1966) 52–58. doi:10.1287/opre.14.1.52.

[9] S. Chopra, M. R. Rao, The partition problem, Mathematical Programming 59 (1993) 87–115. doi:10.1007/BF01581239.

[10] A. Eisenblätter, The semidefinite relaxation of the $k$-partition polytope is strong, in: W. J. Cook, A. S. Schulz (Eds.), Integer Programming and Combinatorial Optimization, Springer, Berlin, Heidelberg, 2002, pp. 273–290. doi:10.1007/3-540-47867-1_20.

[11] S. Chopra, M. R. Rao, Facets of the $k$-partition polytope, Discrete Applied Mathematics 61 (1995) 27–48. doi:10.1016/0166-218X(93)E0175-X.

[12] G. Wang, H. Hijazi, Exploiting sparsity for the min $k$-partition problem, Mathematical Programming Computation 12 (2020) 109–130. doi:10.1007/s12532-019-00165-3.

[13] Gurobi Optimization, LLC, Gurobi Optimizer Reference Manual, 2023. URL: https://gurobi.com/.

[14] A. Frieze, M. Jerrum, Improved approximation algorithms for max $k$-cut and max bisection, Algorithmica 18 (1997) 67–81. doi:10.1007/BF02523688.

[15] R. Sotirov, An efficient semidefinite programming relaxation for the graph partition problem, INFORMS Journal on Computing 26 (2014) 16–30. doi:10.1287/ijoc.1120.0542.

[16] E. R. van Dam, R. Sotirov, New bounds for the max-$k$-cut and chromatic number of a graph, Linear Algebra and its Applications 488 (2016) 216–234. doi:10.1016/j.laa.2015.09.043.

[17] C. Lu, Z. Deng, A branch-and-bound algorithm for solving max-$k$-cut problem, Journal of Global Optimization (2021) 1–23. doi:10.1007/s10898-021-00999-z.

[18] V. R. de Sousa, M. Anjos, S. Le Digabel, Improving the linear relaxation of maximum k-cut with semidefinite-based constraints, EURO Journal on Computational Optimization 7 (2019) 123–151. doi:10.1007/s13675-019-00110-y.

[19] R. Fakhimi, H. Validi, Max $k$-cut, 2022. URL: https://github.com/qcol-lu/maxkcut.

[20] MOSEK, MOSEK Documentation, 2022. URL: https://mosek.com/.

[21] I. S. on Mathematical Programming, Graph coloring and its generalizations, 2002. https://mat.gsia.cmu.edu/COLOR02/.

[22] T. Koch, A. Martin, S. Voß, Steinlib: An updated library on steiner tree problems in graphs, in: Steiner trees in industry, Springer, 2001, pp. 285–325. doi:10.1007/978-1-4613-0255-1_9.

[23] C. Roos, T. Terlaky, J.-P. Vial, Interior Point Methods for Linear Optimization, Springer Science & Business Media, 2005. doi:10.1007/b100325.