Generating Cutting Inequalities Successively for Quadratic Optimization Problems in Binary Variables

Sunyoung Kim*, Masakazu Kojima†

July 19, 2021

Abstract

We propose a successive generation of cutting inequalities for binary quadratic optimization problems. Multiple cutting inequalities are successively generated for the convex hull of the set of the optimal solutions $\subset \{0,1\}^n$, while the standard cutting inequalities are used for the convex hull of the feasible region. An arbitrary linear inequality with integer coefficients and the right-hand side value in integer is considered as a candidate for a valid inequality. The validity of the linear inequality is determined by solving a conic relaxation of a subproblem such as the doubly nonnegative relaxation, under the assumption that an upper bound for the unknown optimal value of the problem is available. Moreover, the candidates generated for the multiple cutting inequalities are tested simultaneously for their validity in parallel. Preliminary numerical results on 60 quadratic unconstrained binary optimization problems with a simple implementation of the successive cutting inequalities using an 8- or 32-core machine show that the exact optimal values are obtained for 70% of the tested problems, demonstrating the strong potential of the proposed technique.

Key words. Quadratic optimization problems, Binary variables, Cutting inequalities, Cutting planes, Conic relaxations, DNN relaxations, Newton-bracketing method, Lower bounds.

AMS Classification. 90C10, 90C20, 90C25, 90C26.

1 Introduction

Cutting inequalities [6, 7, 8, 12, 21] and conic relaxations such as linear programming (LP), semidefinite programming (SDP) [3], doubly nonnegative (DNN) [14, 5] have been regarded as the two most basic tools for solving nonconvex and/or combinatorial optimization problems. They have been frequently incorporated in the brach-and-bound and branch-and-cut framework [10, 11, 17, 18] to solve the problems.

*Department of Mathematics, Ewha W. University, Seoul, 52 Ewhayeodae-gil, Sudaemoon-gu, Seoul 120-750, Korea (skim@ewha.ac.kr). The research was supported by NRF 2021-R1A2C1003810.

†Department of Industrial and Systems Engineering, Chuo University, Tokyo 192-0393, Japan (kojima@is.titech.ac.jp). This research was supported by Grant-in-Aid for Scientific Research (A) 19H00808.

The main purpose of this paper is to propose a *successive cutting inequality technique*, abbreviated by *SCIT*, for binary quadratic optimization problems (QOPs, *i.e.*, QOPs in binary variables), and to demonstrate its strong potential to become a very powerful tool for solving binary QOPs, through preliminary numerical results by an experimental method that implements the very basics of SCIT.

To describe the motivation and basic idea of SCIT, we consider a general nonconvex optimization problem:

$$\text{P: } \zeta = \min\{f(x) : \boldsymbol{x} \in S\},$$

where f denotes a real valued function on the n-dimensional Euclidean space $\mathbb{R}^n$ and S denotes a closed subset of $\mathbb{R}^n$. The assumption that f is a polynomial function in $\boldsymbol{x} \in \mathbb{R}^n$ and S is described by polynomial equalities and inequalities [20] is required at least for the discussion of a conic relaxation of problem P. While the case where those polynomials are linear or quadratic is mainly dealt with, such assumptions are not so relevant in the discussion below.

Except for LP relaxation, a conic relaxation problem with a linear objective function over a closed convex feasible region $\widehat{S}$ is embedded in a different space $\mathbb{V}$, often called a lifted space, with a higher dimension such as the linear space of symmetric matrices. $\widehat{S}$ is described by linear equalities and inequalities in $\mathbb{V}$ and a closed convex cone $\mathbb{K} \subset \mathbb{V}$. In short, when the cone $\mathbb{K}$ used is the nonnegative orthant of the Euclidean space, the positive semidefinite matrix cone or the doubly nonnegative matrix cone, the conic relaxation is called an LP relaxation, an SDP relaxation or a DNN relaxation, respectively. The lifted space $\mathbb{V}$ is identified with the original n-dimensional Euclidean space $\mathbb{R}^n$ where problem P is defined, and f itself is assumed to be linear, for simplicity of discussion below. Then, it is clear that $\zeta = \min\{f(\boldsymbol{x}) : \mathrm{co}(S)\}$, where $\mathrm{co}(S)$ denotes the convex hull of S.

Under the above setting and assumptions, the conic relaxation of P can be described as a convex optimization problem $\widehat{\mathrm{P}}$ with a linear objective function : $\min \hat{\zeta} = \{\hat{f}(\boldsymbol{x}) : \boldsymbol{x} \in \widehat{S}\}$, where $\widehat{S}$ is a closed convex subset of $\mathbb{V}$ (identified with $\mathbb{R}^n$) such that $S \subseteq \mathrm{co}(S) \subseteq \widehat{S}$ and $\hat{f}(\boldsymbol{x}) = f(\boldsymbol{x})$ for every $\boldsymbol{x} \in S$. In practice, $\widehat{\mathrm{P}}$ is constructed as a numerically tractable convex optimization problem. If $\mathrm{co}(S) = \widehat{S}$ held, we would have the optimal value ζ of problem P by solving the relaxation problem $\widehat{\mathrm{P}}$. In most cases, however, $\mathrm{co}(S)$ is a proper subset of $\widehat{S}$, $\hat{\zeta} < \zeta$, and an optimal solution $\bar{\boldsymbol{x}}$ of $\widehat{\mathrm{P}}$ is not a feasible solution of P. As a result, an inequality which cuts off $\bar{\boldsymbol{x}}$ from $\widehat{S}$ but does not remove any $\boldsymbol{x}$ from $\mathrm{co}(S)$ is desired to improve the lower bound $\hat{\zeta}$ of ζ and to compute a feasible approximate optimal solution of P. This is a standard role and usage of cutting inequalities. Note that a cutting inequality is chosen from the family of valid inequalities of S. In fact, $\mathrm{co}(S)$ can be described as the set of points $\boldsymbol{x}$ which satisfies all the valid inequalities of S. Well-known triangle inequalities, which forms a sub-family of the valid inequalities for the binary polytopes, are frequently used to strengthen the SDP relaxation of binary QOPs [17].

In this paper, a cutting inequality plays a more active role under the additional assumption that $S \subseteq \{0,1\}^n$. We propose to generate a cutting inequality for the convex hull $\mathrm{co}(S^*)$ of the set S^* of optimal solutions of P, an inequality which is aimed at cutting off $\boldsymbol{x} \notin \mathrm{co}(S^*)$ from $\mathrm{co}(S^*)$ (Recall that the standard cutting inequality is for the convex hull of the feasible region S). Generating such a cutting inequality is based on the following ideas: Assume that an upper bound η of the unknown optimal value ζ is available. Let $\boldsymbol{g}$ be an arbitrary integer column vector in $\mathbb{R}^n$. For every integer α, we consider a pair of subset $S(\alpha) = \{\boldsymbol{x} \in S : \boldsymbol{g}^T\boldsymbol{x} \leq \alpha\}$ and $S(\alpha)^+ = \{\boldsymbol{x} \in S : \boldsymbol{g}^T\boldsymbol{x} \geq \alpha + 1\}$. Since $\boldsymbol{g}^T\boldsymbol{x}^*$ is an integer for every $\boldsymbol{x}^* \in S^* \subset \{0,1\}^n$, it is obvious that S^* is included in the union of $S(\alpha)$ and $S(\alpha)^+$. Hence if $S^* \cap S(\alpha) = \emptyset$, then $S^* \subseteq S(\alpha)^+$, *i.e.*, $\boldsymbol{g}^T\boldsymbol{x} \geq \alpha + 1$ serves as a cutting inequality for S^*. To obtain a certificate of $S^* \cap S(\alpha) = \emptyset$, we solve a conic relaxation $\widehat{\mathrm{P}}(\alpha)$: $\hat{\zeta}(\alpha) = \min\{\hat{f}(\boldsymbol{x}) : \boldsymbol{x} \in \widehat{S}(\alpha)\}$ of a subproblem $\mathrm{P}(\alpha)$: $\zeta(\alpha) = \min\{f(\boldsymbol{x}) : \boldsymbol{x} \in S(\alpha)\}$, where $\widehat{S}(\alpha)$ denotes a closed convex subset of $\mathbb{V} = \mathbb{R}^n$ containing $S(\alpha)$. If $\eta < \hat{\zeta}(\alpha)$ holds, then $\zeta \leq \eta < \hat{\zeta}(\alpha)$; hence $S(\alpha) \cap S^* \subseteq \widehat{S}(\alpha) \cap S^* = \emptyset$. Therefore, the inequality $\eta < \hat{\zeta}(\alpha)$ is a certificate for $\boldsymbol{g}^T\boldsymbol{x} \geq \alpha + 1$ to be a cutting inequality for S^*. The largest α such that $\eta < \hat{\zeta}(\alpha)$ is most desirable to cut off a larger portion of $S \backslash S^*$. For such an α, a 1-dimensional search can be applied over the set of integers with starting $\alpha = -1$ since $\boldsymbol{g}^T\boldsymbol{x} \geq 0$ is a trivial cutting inequality for S^*.

In the proposed SCIT, multiple candidates for cutting inequalities in the form $\boldsymbol{g}_j^T\boldsymbol{x} \geq \alpha_j + 1$ $(j = 1, \ldots, m)$ are arranged before the iteration starts and set $\alpha_j + 1 = 0$ so that $\boldsymbol{g}_j^T\boldsymbol{x} \geq 0$ becomes a trivial cutting inequality for S^* $(j = 1, \ldots, m)$. At each iteration of SCIT, it verifies whether $\boldsymbol{g}_j^T\boldsymbol{x} \geq \alpha'_j + 1$ remains a valid inequality for some $\alpha'_j > \alpha_j$ by solving a conic relaxation problem for all $j = 1, \ldots, m$. If it does, then α_j is updated to α'_j, otherwise α'_j is replaced by a smaller $\alpha'_j \in [\alpha_j, \alpha'_j)$ for the next iteration. Notice that these verifications and updates with $j = 1, \ldots, m$ can be simultaneously performed (within one iteration) in parallel.

The effectiveness and efficiency of the performance of SCIT on large scale binary QOPs is dependent on the followings:

(I) A tight upper bound η is available for the unknown optimal value ζ.

(II) A strong conic relaxation method that generates a tight lower bound $\hat{\zeta}(\alpha)$ for $\zeta(\alpha)$ can be utilized.

(III) A powerful computer system can be used for parallel computing.

As an application of SCIT, we consider quadratic unconstrained binary optimization problems (QUBOs) in Section 4. There exist many heuristic methods, which can be used for computing a tight upper bound η of the optimal value ζ of a QUBO, such as the tabu search [10] and the genetic algorithm [22]. For (II), we utilize the Lagrangian-DNN relaxation [14, 5, 4], which is known to be much stronger than the standard SDP relaxation, and NewtBracket [15] (the Newton-bracketing method [16]) as a numerical method to compute its optimal value. For (III), preliminary numerical results with a small scale computer are reported. Inequalities of the form $\sum_{i \in I} x_i \geq \alpha + 1$ for some $I \subseteq \{1, \ldots, n\}$ are considered as the candidates for the cutting inequalities. If the inequality is shown to be a cutting inequality with $I = \{j\}$ and $\alpha + 1 = 1$, then x_j can be fixed to $x_j = 1$ and the size of the QUBO to be solved can be reduced. This is an important feature of SCIT.

We investigate the numerical performance of SCIT through an experimental method on 60 QUBO instances with dimensions up to 250 from BIQMAC [23]. Although the method is just a simple implementation of SCIT, not a well-designed software for solving QUBOs, it attained the exact optimal value within 10 iterations for 70% cases of the 60 instances. This is a remarkable result, which could not be expected. It shows the promising potential of SCIT when it is incorporated into the branch-and-bound method [10, 11, 17, 18]. We mention that theoretical aspects of SCIT including the convergence to the convex hull of S^* are not dealt with here.

In Section 2, we present the fundamental facts which our construction of cutting inequalities build on after introducing notation and symbols. We present some details on SCIT in Section 3, and discuss its application to QUBOs in Section 4. The preliminary numerical results mentioned above are given in Section 4.4. We conclude in Section 5.

2 Preliminaries

2.1 Notation and symbols

Let $\mathbb{R}$ = the set of real numbers and $\mathbb{Z}$ = the set of integers. For $\mathbb{V} = \mathbb{R}$ or $\mathbb{Z}$, $\mathbb{V}^n$ denotes the set of n-dimensional column vectors $(v_1, \ldots, v_n)$ with elements $v_i \in \mathbb{V}$ $(i = 1, \ldots, n)$, and $\mathbb{V}^{\ell\times\ell}$ the set of $\ell \times \ell$ matrices $\boldsymbol{V} = [V_{ij}]$ with elements $V_{ij} \in \mathbb{V}$ $(1 \leq i, j \leq \ell)$. In particular, $\mathbb{R}^n$ denotes the n-dimensional Euclidean space of column vectors. $\boldsymbol{v}^T$ stands for the transposed row vector of $\boldsymbol{v}$ for every $\boldsymbol{v} \in \mathbb{V}^n$, and $\boldsymbol{u}^T\boldsymbol{v}$ the inner product $\sum_{i=1}^n u_i v_i$ of $\boldsymbol{u},\ \boldsymbol{v} \in \mathbb{V}^n$. For $\boldsymbol{U}, \boldsymbol{V} \in \mathbb{V}^{\ell\times\ell}$, their inner product is written as $\langle \boldsymbol{U},\ \boldsymbol{V} \rangle = \sum_{i=1}^\ell \sum_{j=1}^\ell U_{ij}V_{ij}$. Let

$$
\begin{aligned}
\mathbb{S}^\ell &= \text{the linear space of } \ell \times \ell \text{ symmetric matrices } \boldsymbol{X} = [X_{ij}]\ (1 \leq i, j \leq \ell), \\
\mathbb{S}^\ell_+ &= \text{the cone of positive semidefinite matrices in } \mathbb{S}^\ell.
\end{aligned}
$$

Throughout the paper, $\widehat{\ }$ and $\hat{\ }$ (also $\widetilde{\ }$ and $\tilde{\ }$) are used for conic relaxation problems such that a conic relaxation problem $\widehat{\mathrm{P}}$ (2) of the optimization problem P (1) and their optimal values $\hat{\zeta}$ and ζ, respectively. We use the subscripts $_\mathrm{a}$ and $_\mathrm{b}$ for a pair of subproblems obtained from their common parent problem by adding a cutting inequality; for example, a pair of subproblems P_a (3) and P_a (4) of P (1) and their optimal values ζ_a and ζ_b, respectively.

2.2 Basic ideas to generate cutting inequalities

We begin with the following simple facts on which our cutting inequalities are constructed.

Lemma 2.1. *Let ζ, ζ_a, ζ_b, η, $\hat{\zeta}_\mathrm{a}$, and $\hat{\zeta}_\mathrm{b}$ be real numbers satisfying $\zeta = \min\{\zeta_\mathrm{a}, \zeta_\mathrm{b}\} \leq \eta$, $\hat{\zeta}_\mathrm{a} \leq \zeta_\mathrm{a}$ and $\hat{\zeta}_\mathrm{b} \leq \zeta_\mathrm{b}$. Assume that $\eta < \hat{\zeta}_\mathrm{a}$. Then $\hat{\zeta}_\mathrm{b} \leq \zeta_\mathrm{b} = \zeta$.*

Proof. It follows from $\zeta = \min\{\zeta_\mathrm{a}, \zeta_\mathrm{b}\}$ that at least one of $\zeta = \zeta_\mathrm{a}$ and $\zeta = \zeta_\mathrm{b}$ holds. If $\zeta_\mathrm{a} = \zeta$ held, then we would have $\hat{\zeta}_\mathrm{a} \leq \zeta_\mathrm{a} = \zeta \leq \eta$. This contradicts to the assumption that $\eta < \hat{\zeta}_\mathrm{a}$. □

We note that the cutting inequalities in Section 3.3 are constructed by Lemma 2.1. More precisely,

- ζ corresponds to the unknown optimal (minimum) value of the optimization problem P (1), our target problem to solve, and η to a known upper bound of ζ.
- ζ_a and ζ_b correspond to optimal values of a pair of subproblems P_a (3) and P_b (4), which are generated by adding cut inequalities to the feasible region of P (1); hence $\zeta \leq \zeta_\mathrm{a}$ and $\zeta \leq \zeta_\mathrm{b}$. The identity $\min\{\zeta_\mathrm{a}, \zeta_\mathrm{b}\} = \zeta$ means at least one of P_a and P_b attains the same objective value as P (1).

- $\hat{\zeta}_{\rm a}$ and $\hat{\zeta}_{\rm b}$ correspond to the optimal values of $\widehat{\rm P}_a$ (5) and $\widehat{\rm P}_{\rm b}$ (6), which are conic relaxations of $\rm P_a$ (3) and $\rm P_b$ (4), respectively.

Lemma 2.2. *Let ζ, η, $\hat{\zeta}$, $\tilde{\zeta}_{\rm a}$ and $\tilde{\zeta}_{\rm b}$ be real numbers satisfying $\hat{\zeta} \leq \min\{\tilde{\zeta}_{\rm a}$ and $\tilde{\zeta}_{\rm b}\} \leq \zeta \leq \eta$. Assume that $\eta < \tilde{\zeta}_{\rm a}$. Then, $\hat{\zeta} \leq \tilde{\zeta}_{\rm b} \leq \zeta$.*

Proof. Obvious.

By Lemma 2.2, the cutting inequalities in Section 3.4 are constructed, where

- ζ corresponds to the unknown optimal (minimum) value of the optimization problem P (1), the target problem, and η to a known upper bound of ζ.
- $\hat{\zeta}$ corresponds to the optimal values of a conic relaxation problem $\hat{\rm P}$ (2) of P (1).
- $\tilde{\zeta}_{\rm a}$ and $\tilde{\zeta}_{\rm b}$ correspond the optimal values of a pair of subproblems $\widetilde{\rm P}_{\rm a}$ (7) and $\widetilde{\rm P}_{\rm b}$ (8) of $\widehat{\rm P}$ (2), which are generated by adding cutting inequalities to the feasible region of $\widehat{\rm P}$ (2). The inequality $\min\{\tilde{\zeta}_{\rm a}, \tilde{\zeta}_{\rm b}\} \leq \zeta$ means at least one of $\widetilde{\rm P}_{\rm a}$ (7) and $\widetilde{\rm P}_{\rm b}$ (8) acts as a conic relaxation of P (1).

3 Conic relaxations of optimization problems in binary variable with cutting inequalities

3.1 An optimization problem in binary variables

Throughout this section, we consider the following nonconvex optimization problem in binary variables $x_i \in \{0,1\}$ $(i = 1,\ldots,n)$:

$$\text{P:}\ \zeta \ = \ \min\left\{f(\boldsymbol{x}) : \boldsymbol{x} \in S\right\}, \tag{1}$$

where

(a) $\emptyset \neq S \subset \{0,1\}^n$,

(b) $f(\boldsymbol{x}) \in \mathbb{Z}$ for every $\boldsymbol{x} \in \{0,1\}^n$.

Under these conditions, problem P has an optimal solution $\boldsymbol{x}^*$. The optimal value ζ and solution $\boldsymbol{x}^*$ are unknown. Moreover, we assume that

(c) An upper bound $\eta \in \mathbb{Z}$ for ζ is available.

3.2 A conic relaxation of problem P in the lifted symmetric matrix space

We first introduce a finite dimensional vector space into which the conic relaxation is embedded. For simplicity of discussion and convenience of presenting an application in Section 4, we focus on the case where the linear space is $\mathbb{S}^\ell$ of $\ell \times \ell$ symmetric matrices.

Consider

$$\widehat{\text{P}}\text{:}\ \hat{\zeta} \ = \ \min\left\{\langle \boldsymbol{Q},\, \boldsymbol{X}\rangle : \boldsymbol{X} \in \widehat{S}\right\}. \tag{2}$$

Here

(d) $\boldsymbol{Q}$ is a matrix in $\mathbb{S}^\ell \cap \mathbb{Z}^{\ell\times\ell}$ such that $\langle \boldsymbol{Q},\, \boldsymbol{\Phi}(\boldsymbol{x})\rangle = f(\boldsymbol{x})$ for every $\boldsymbol{x} \in \{0,1\}^n$.

(e) $\boldsymbol{\Phi}$ is a mapping from $\mathbb{R}^n$ into $\mathbb{S}^\ell$ such that $\boldsymbol{\Phi}(\boldsymbol{x}) \in \mathbb{S}^\ell \cap \mathbb{Z}^{\ell\times\ell}$ if $\boldsymbol{x} \in \{0,1\}^n$.

(f) $\widehat{S}$ is a closed convex subset of $\mathbb{S}^\ell$ such that $\boldsymbol{\Phi}(S) \subset \widehat{S}$, *i.e.*, $\boldsymbol{\Phi}(\boldsymbol{x}) \in \widehat{S}$ for every $\boldsymbol{x} \in S$.

These three conditions characterize problem $\widehat{\mathrm{P}}$ (2) as a conic (SDP and DNN) relaxation problem of P (1) in the space $\mathbb{S}^\ell$. In particular, $\hat{\zeta} \leq \zeta \leq \eta$ and $\boldsymbol{\Phi}(\boldsymbol{x}^*) \in \widehat{S}$.

Example 3.1. (A simple illustrative example). Let

$$\begin{aligned}
\boldsymbol{x} &= (x_1,\ldots,x_n) \in \mathbb{R}^n,\ \boldsymbol{Q} \in \mathbb{S}^n \cap \mathbb{Z}^{n\times n},\\
S &= \left\{\boldsymbol{x} \in \mathbb{R}^n :\ x_1 = 1,\ x_1x_i = x_i^2\ (i=2,\ldots,n)\right\} \subset \{0,1\}^n,\\
\boldsymbol{\Phi}(\boldsymbol{x}) &= \boldsymbol{x}\boldsymbol{x}^T \in \mathbb{S}^n,\ f(\boldsymbol{x}) \ =\ \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} = \langle \boldsymbol{Q},\, \boldsymbol{x}\boldsymbol{x}^T\rangle = \langle \boldsymbol{Q},\, \boldsymbol{\Phi}(\boldsymbol{x})\rangle,\\
\widehat{S} &= \left\{\boldsymbol{X} \in \mathbb{S}^n : \boldsymbol{X} \in \mathbb{S}^n_+,\ X_{11} = 1,\ X_{1i} = X_{ii}\ (i=2,\ldots,n)\right\}.
\end{aligned}$$

Note that $f(\boldsymbol{x})$ with $x_1 = 1$ can be rewritten as

$$f(\boldsymbol{x}) \ =\ \sum_{i=2}^n\sum_{j=2}^n Q_{ij}x_ix_j + 2\sum_{j=2}^n Q_{1j}x_j + Q_{11}.$$

Thus, $\min\left\{\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} : \boldsymbol{x} \in S\right\}$ corresponds to a QUBO, and $\min\left\{\langle \boldsymbol{Q},\, \boldsymbol{X}\rangle : \boldsymbol{X} \in \widehat{S}\right\}$ to the standard SDP relaxation. Conditions (a), (b) and (e) are obviously satisfied with $\ell = n$. It is also straightforward to see that Conditions (d) and (f) are satisfied.

Remark 3.2. Under Condition (d), problem P (1) can be reformulated as

$$\zeta \ =\ \min\left\{\langle \boldsymbol{Q},\, \boldsymbol{\Phi}(\boldsymbol{x})\rangle : \boldsymbol{x} \in S\right\}.$$

It is known that if f is a polynomial function with integer coefficients in $\boldsymbol{x} \in \{0,1\}^n$, then we can take a matrix $\boldsymbol{Q} \in \mathbb{S}^\ell \cap \mathbb{Z}^\ell$ and a mapping $\boldsymbol{\Phi} : \mathbb{R}^n \to \mathbb{S}^\ell$ for some ℓ such that Conditions (d) and (e) hold. Therefore, all the discussions in this section are valid for polynomial optimization problems in binary variables (See [20, 19]). We mention that the authors' main interest is to develop a practical numerical method for solving large scale linearly constrained quadratic optimization problems in binary variables by effectively utilizing *SCIT, which will be presented in Section 3.5.*

3.3 A cutting inequality for the feasible region S of problem P

Let $\boldsymbol{g} \in \mathbb{Z}^n$ and $\alpha \in \mathbb{Z}$. We consider the following pair of subproblems of P:

$$\mathrm{P_a} : \zeta_\mathrm{a} \ =\ \min\left\{f(\boldsymbol{x}) : \boldsymbol{x} \in S_\mathrm{a}\right\},\ \text{where } S_\mathrm{a} = \left\{\boldsymbol{x} \in S : \boldsymbol{g}^T\boldsymbol{x} \leq \alpha\right\}, \tag{3}$$

$$\mathrm{P_b} : \zeta_\mathrm{b} \ =\ \min\left\{f(\boldsymbol{x}) : \boldsymbol{x} \in S_\mathrm{b}\right\},\ \text{where } S_\mathrm{b} = \left\{\boldsymbol{x} \in S : \boldsymbol{g}^T\boldsymbol{x} \geq \alpha+1\right\}. \tag{4}$$

Let $\hat{\zeta}_\mathrm{a}$ and $\hat{\zeta}_\mathrm{b}$ denote the optimal values of conic relaxations of $\mathrm{P_a}$ and $\mathrm{P_b}$ such that $\hat{\zeta}_\mathrm{a} \leq \zeta_\mathrm{a}$ and $\hat{\zeta}_\mathrm{b} \leq \zeta_\mathrm{b}$ hold, respectively. More precisely, the conic relaxations of $\mathrm{P_a}$ and $\mathrm{P_b}$ are written as

$$\widehat{\mathrm{P}}_\mathrm{a} : \hat{\zeta}_\mathrm{a} \ =\ \min\left\{\langle \boldsymbol{Q},\, \boldsymbol{X}\rangle : \boldsymbol{X} \in \widehat{S}_\mathrm{a}\right\}, \tag{5}$$

$$\widehat{\mathrm{P}}_\mathrm{b} : \hat{\zeta}_\mathrm{b} \ =\ \min\left\{\langle \boldsymbol{Q},\, \boldsymbol{X}\rangle : \boldsymbol{X} \in \widehat{S}_\mathrm{b}\right\}, \tag{6}$$

respectively. Here we assume that condition

(f_{ab}) $\widehat{S}_{\mathrm{a}}$ and $\widehat{S}_{\mathrm{b}}$ are closed convex subsets of $\mathbb{S}^{\ell}$ such that $\boldsymbol{\Phi}(S_{\mathrm{a}}) \subset \widehat{S}_{\mathrm{a}}$ and $\boldsymbol{\Phi}(S_{\mathrm{b}}) \subset \widehat{S}_{\mathrm{b}}$

holds in addition to Condition (d) and (e). Hence $\hat{\zeta}_a \leq \zeta_a$ and $\hat{\zeta}_b \leq \zeta_b$. Since $\boldsymbol{g}^T\boldsymbol{x}^* \in \mathbb{Z}$, $\boldsymbol{x}^* \in S$ lies in either S_a or S_b. This implies $\zeta = \min\{\zeta_{\mathrm{a}}, \zeta_{\mathrm{b}}\}$. Therefore, we can conclude that if $\eta < \hat{\zeta}_{\mathrm{a}}$, where η denotes a known upper bound of ζ (See Condition (c)), then $\hat{\zeta}_{\mathrm{b}} \leq \zeta_{\mathrm{b}} = \zeta$. (Recall Lemma 2.1). In other words, problem P_b (4) with the cutting inequality $\boldsymbol{g}^T\boldsymbol{x} \geq \alpha+1$ shares the same optimal value $\zeta_b = \zeta$ as the original problem P (1), and its conic relaxation $\widehat{\mathrm{P}}_b$ (6) provides a lower bound $\hat{\zeta}_b$, which is at least as tight as the original lower bound $\hat{\zeta}$ for P.

3.4 A cutting inequality for the feasible region $\widehat{S}$ of the conic relaxation problem $\widehat{\mathrm{P}}$

Let $\boldsymbol{G} \in \mathbb{S}^{\ell} \cap \mathbb{Z}^{\ell\times\ell}$ and $\alpha \in \mathbb{Z}$. We consider the following pair of subproblems of $\widehat{\mathrm{P}}$ (2):

$$\widetilde{\mathrm{P}}_{\mathrm{a}} : \tilde{\zeta}_{\mathrm{a}} \;=\; \min\left\{\langle \boldsymbol{Q},\, \boldsymbol{X}\rangle : \boldsymbol{X} \in \widehat{S},\; \langle \boldsymbol{G},\, \boldsymbol{X}\rangle \leq \alpha\right\}. \tag{7}$$

$$\widetilde{\mathrm{P}}_{\mathrm{b}} : \tilde{\zeta}_{\mathrm{b}} \;=\; \min\left\{\langle \boldsymbol{Q},\, \boldsymbol{X}\rangle : \boldsymbol{X} \in \widehat{S},\; \langle \boldsymbol{G},\, \boldsymbol{X}\rangle \geq \alpha+1\right\}. \tag{8}$$

Obviously, $\hat{\zeta} \leq \min\{\tilde{\zeta}_{\mathrm{a}}, \tilde{\zeta}_{\mathrm{b}}\}$. Since $\boldsymbol{\Phi}(\boldsymbol{x}^*) \in \widehat{S} \cap \mathbb{Z}^{\ell\times\ell}$ and $\langle \boldsymbol{G},\, \boldsymbol{\Phi}(\boldsymbol{x}^*)\rangle \in \mathbb{Z}$, we have either $\langle \boldsymbol{G},\, \boldsymbol{\Phi}(\boldsymbol{x}^*)\rangle \leq \alpha$ or $\langle \boldsymbol{G},\, \boldsymbol{\Phi}(\boldsymbol{x}^*)\rangle \geq \alpha+1$, which implies

$$\tilde{\zeta}_{\mathrm{a}} \leq \langle \boldsymbol{Q},\, \boldsymbol{\Phi}(\boldsymbol{x}^*)\rangle = \zeta \text{ or } \tilde{\zeta}_{\mathrm{b}} \leq \langle \boldsymbol{Q},\, \boldsymbol{\Phi}(\boldsymbol{x}^*)\rangle = \zeta.$$

Thus, $\min\{\tilde{\zeta}_{\mathrm{a}}, \tilde{\zeta}_{\mathrm{b}}\} \leq \zeta$. Consequently, we can conclude that if $\eta < \tilde{\zeta}_{\mathrm{a}}$, then $\hat{\zeta} \leq \tilde{\zeta}_{\mathrm{b}} \leq \zeta$. (See Lemma 2.2). In other words, $\widetilde{\mathrm{P}}_b$ (8) with the cutting inequality $\langle \boldsymbol{G},\, \boldsymbol{X}\rangle \geq \alpha+1$ provides a lower bound $\tilde{\zeta}_b$, at least as tight as the original lower bound $\hat{\zeta}$, for the unknown optimal value ζ of problem P (1).

3.5 Successive cutting inequality technique (SCIT)

The generation of a single cutting inequality for S presented in Section 3.3 can be extended in a straightforward fashion to simultaneous generation of multiple cutting inequalities. Let $(\boldsymbol{g}_j, \alpha_j) \in \mathbb{Z}^{n+1}$ $(j = 1,\ldots,m)$. For each j, we consider the following subproblem of P (1):

$$\mathrm{P}_{\mathrm{a}j} : \zeta_{\mathrm{a}j} \;=\; \min\left\{f(\boldsymbol{x}) : \boldsymbol{x} \in S_{\mathrm{a}j}\right\}, \text{ where } S_{\mathrm{a}j} = \left\{\boldsymbol{x} \in S : \boldsymbol{g}_j^T\boldsymbol{x} \leq \alpha_j\right\},$$

and a conic relaxation of $\mathrm{P}_{\mathrm{a}j}$

$$\widehat{\mathrm{P}}_{\mathrm{a}j} : \widehat{\zeta}_{\mathrm{a}j} \;=\; \min\left\{\langle \boldsymbol{Q},\, \boldsymbol{X}\rangle : \boldsymbol{X} \in \widehat{S}_{\mathrm{a}j}\right\}. \tag{9}$$

Let $J = \{j : \eta < \hat{\zeta}_{\mathrm{a}j}\}$. Then, by adding the cutting inequalities $\boldsymbol{g}_j^T\boldsymbol{x} \geq \alpha_j + 1$ $(j \in J)$ to problem P, we obtain

$$\mathrm{P}^1\!: \zeta \;=\; \min\left\{f(\boldsymbol{x}) : \boldsymbol{x} \in S^1\right\}, \tag{10}$$

where $S^1 = \left\{\boldsymbol{x} \in S : \boldsymbol{g}_j^T\boldsymbol{x} \geq \alpha_j + 1 \; (j \in J)\right\}$, and its conic relaxation:

$$\widehat{\mathrm{P}}^1 : \hat{\zeta}^1 \;=\; \min\left\{\langle \boldsymbol{Q},\, \boldsymbol{X}\rangle : \boldsymbol{X} \in \widehat{S}^1\right\}. \tag{11}$$

Problem P^1 (10) shares the same optimal value ζ and optimal solution $\boldsymbol{x}^*$ with problem P, and $\hat{\zeta} \leq \hat{\zeta}^1 \leq \zeta$. Note that problems $\widehat{\mathrm{P}}_{\mathrm{a}j}$ $(j = 1, \ldots, m)$ can be solved independently in parallel.

The simultaneous generation of multiple cutting inequalities above can be applied now to the feasible region S^1 of problem P^1, and a new problem can be constructed as

$$\mathrm{P}^2 : \zeta \;=\; \left\{ f(\boldsymbol{x}) : \boldsymbol{x} \in S^2 \right\},$$

which is equivalent to P and P^1, and its conic relaxation problem

$$\widehat{\mathrm{P}}^2 : \hat{\zeta}^2 \;=\; \left\{ \langle \boldsymbol{Q},\, \boldsymbol{X} \rangle : \boldsymbol{X} \in \widehat{S}^2 \right\}.$$

We continue this process successively to generate a sequence of conic relaxation problems $\{\widehat{\mathrm{P}}^k : k = 1, 2, \ldots\}$ (of $\{\mathrm{P}^k : k = 1, 2, \ldots\}$) and a sequence $\{\hat{\zeta}^k \ (k = 1, \ldots, \ )\}$ of their optimal values, which serve as lower bounds for ζ such that $\hat{\zeta} \leq \hat{\zeta}^k \leq \hat{\zeta}^{k+1} \leq \zeta$ $(k = 0, 1, \ldots)$. This entire process constitutes *SCIT*.

For a similar extension of generating a single cutting inequality for $\widehat{S}$ presented in Section 3.4 to SCIT, let $\boldsymbol{G}_j \in \mathbb{S}^\ell \cap \mathbb{Z}^{\ell \times \ell}$ and $\alpha_j \in \mathbb{Z}$ $(j = 1, \ldots, m)$. We consider

$$\widetilde{\mathrm{P}}_{\mathrm{a}j} : \tilde{\zeta}_{\mathrm{a}j} \;=\; \min \left\{ \langle \boldsymbol{Q},\, \boldsymbol{X} \rangle : \boldsymbol{X} \in \widehat{S},\ \langle \boldsymbol{G}_j,\, \boldsymbol{X} \rangle \leq \alpha_j \right\}$$

for each $j = 1, \ldots, m$. Let $J = \{j : \eta < \tilde{\zeta}_{\mathrm{a}j}\}$. Then, the cutting inequalities $\langle \boldsymbol{G}_j,\, \boldsymbol{X} \rangle \geq \alpha_j + 1$ $(j \in J)$ can be applied to problem $\widehat{\mathrm{P}}$ (1):

$$\widetilde{\mathrm{P}}^1 : \tilde{\zeta}^1 \;=\; \min \left\{ \langle \boldsymbol{Q},\, \boldsymbol{X} \rangle : \boldsymbol{X} \in \widetilde{S}^1 \right\},$$

where $\widetilde{S}^1 = \left\{ \boldsymbol{X} \in \widehat{S} : \langle \boldsymbol{G}_j,\, \boldsymbol{X} \rangle \geq \alpha_j + 1 \ (j \in J) \right\}$. As a result, we obtain that $\hat{\zeta} \leq \tilde{\zeta}^1 \leq \zeta$. We note that conic relaxation problems $\widetilde{\mathrm{P}}_{\mathrm{a}j}$ $(j = 1, \ldots, m)$ can be solved independently in parallel.

Now, the discussion above is applied to the feasible region $\widetilde{S}^1$ of problem $\widetilde{\mathrm{P}}^1$, and a new conic relaxation problem $\widetilde{\mathrm{P}}^2$ of problem P is constructed as:

$$\widetilde{\mathrm{P}}^2 : \tilde{\zeta}^2 \;=\; \min \left\{ \langle \boldsymbol{Q},\, \boldsymbol{X} \rangle : \boldsymbol{X} \in \widetilde{S}^2 \right\}.$$

Continuing this process, a sequence of conic relaxation problems $\left\{ \widetilde{\mathrm{P}}^1, \widetilde{\mathrm{P}}^2, \cdots \right\}$ and a sequence $\{\widetilde{\zeta}^1, \widetilde{\zeta}^2, \ldots\}$ of their optimal values are generated such that $\hat{\zeta} \leq \widetilde{\zeta}^k \leq \widetilde{\zeta}^{k+1} \leq \zeta$ $(k = 1, 2, \ldots)$. Thus generating a single cutting inequality for $\widehat{S}$ has been extended to SCIT for $\widehat{S}$.

4 An application to quadratic unconstrained binary optimization problem (QUBO)

We demonstrate in this section that SCIT presented in Section 3.5 has promising prospects for solving QUBOs. More precisely, we show that the exact optimal values of some QUBO instances can be obtained by simply applying SCIT, without a well-designed numerical method for solving QUBOs.

It should be mentioned that SCIT needs to be eventually combined with other practical numerical methods, such as the branch-and-bound method [10, 11, 17, 18], heuristic methods including the tabu search [10] and the genetic algorithm [22] to solve QUBOs and other binary QOPs. Before designing a specific numerical method using SCIT and conducting extensive numerical experiments on a parallel machine with a large number of cores, we investigate the performance of SCIT on 60 QUBO instances with dimensions 100 - 250 from BIQMAC [23]. Obviously, a lot of flexibility exists in implementing SCIT and many details should be determined. In the subsequent discussion, we choose some specific values for SCIT to just carry out numerical experiments. Those settings are not to propose a numerical method for solving QUBOs. Nevertheless, the exact optimal values could be attained for 70% of the 60 QUBO instances in 10 iterations (see Section 4.4).

4.1 A QUBO

We consider a QUBO:

$$\zeta \quad = \quad \min\left\{\boldsymbol{u}^T\boldsymbol{R}\boldsymbol{u} : \boldsymbol{u} \in \{0,1\}^m\right\}, \tag{12}$$

where $\boldsymbol{R} \in \mathbb{S}^m$. Introducing a slack variable vector $\boldsymbol{v} \in \{0,1\}^m$, we transform the QUBO to

$$\text{P: } \zeta \quad = \quad \min\left\{\boldsymbol{u}^T\boldsymbol{R}\boldsymbol{u} : \boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S\right\}, \tag{13}$$

where $S = \{\boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in \{0,1\}^{2m} : \boldsymbol{u}+\boldsymbol{v} = \boldsymbol{e}\}$ and $\boldsymbol{e}$ denotes the m-dimensional column vector of 1's. It is known that introducing the slack variabl vector $\boldsymbol{v} \in \{0,1\}^m$ is crucial to strengthen the conic relaxation of QUBO (13) (see, for example, [13, Section 6.1]). We note that $(x_1,\ldots,x_m)$ corresponds to $\boldsymbol{u} \in \mathbb{R}^m$ and $(x_{m+1},\ldots,x_{2m})$ to $\boldsymbol{v} \in \mathbb{R}^m$, and that $\boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S$ implies the complementarity $x_i x_{m+i} = u_i v_i = 0$ between $x_i = u_i$ and $x_{m+i} = v_i$ $(i = 1,\ldots,m)$; hence $\sum_{i=1}^{2m} x_i = m$ holds for every $\boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S$. Let $\boldsymbol{x}^* = (\boldsymbol{u}^*,\boldsymbol{v}^*)$ be an unknown optimal solution of QUBO (13), and η a known upper bound of the optimal value ζ.

4.2 Cutting inequalities in the number of 1's in $(\boldsymbol{u},\boldsymbol{v}) \in \{0,1\}^{2m}$

For each $I \subset \{1,\ldots,2m\}$ and $\alpha \in \{0,1,\ldots,|I|-1\}$ where $|I|$ denotes the number of elements of I, we consider the following type of cutting inequality for the feasible region S of QUBO (13):

$$\sum_{i\in I} x_i \geq \alpha + 1,$$

which together with $\boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S$ requires that the number of 1's among x_i $(i \in I)$ is at least $\alpha+1$. In particular, if we take $I = \{i\}$ with $i \in \{1,\ldots,m\}$ (or $i \in \{m+1,\ldots,2m\}$) and $\alpha+1 = 1$, the cutting inequality $\sum_{i\in I} x_i \geq \alpha+1$ requires $u_i = 1$ and $v_i = 0$ (or $u_i = 0$ and $v_i = 1$). Thus, it is possible to fix u_i to 1 (or 0) and reduce the size of QUBO if the inequality is shown to be valid for $\boldsymbol{x} = \boldsymbol{x}^*$.

4.3 An experimental method using SCIT

To initialize the sequence

$$\left\{\{(I,\alpha_I^k,\beta_I^k) : I \in \mathcal{I}^k\} : k = 0,1,\ldots,\right\}, \tag{14}$$

which is to be generated, set

$$\begin{array}{rcl} \mathcal{I}^0 & = & \text{a family of nonempty subsets of } \{1,\ldots,2m\}, \\ \alpha_I^0 & = & 0 \ \text{ and } \beta_I^0 \ = \ \lfloor \gamma |I| \rfloor \text{ for every } I \in \mathcal{I}^0, \end{array}$$

where $\gamma = 1/3$ is used in the preliminary numerical experiment reported in Section 4.4. For each $k = 0,1,\ldots$, we consider

$$\mathrm{P}^k : \zeta^k \ = \ \min\left\{ \boldsymbol{u}^T\boldsymbol{R}\boldsymbol{u} : \boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S, \ \sum_{i\in I} x_i \geq \alpha_I^k \ (I \in \mathcal{I}^k) \right\}, \tag{15}$$

and its conic relaxation problem $\widehat{\mathrm{P}}^k$ with the optimal value $\hat{\zeta}^k$.

Let $k = 0$. Since $\alpha_I^k = 0$ for every $I \in \mathcal{I}^k$, the inequalities $\sum_{i\in I} x_i \geq \alpha_I^k \ (I \in \mathcal{I}^k)$ obviously hold for every $\boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S$ and problem P^k (15) is equivalent to problem P (13). Hence,

$$\boldsymbol{x}^* \text{ remains an optimal solution of } \mathrm{P}^k \ \text{ and } \hat{\zeta}^k \leq \zeta^k = \zeta. \tag{16}$$

Assuming that (16) holds for some iteration $k \in \{0,1,\ldots\}$, we show how to update $\{(I,\alpha_I^k,\beta_I^k) : I \in \mathcal{I}^k\}$ to $\{(I,\alpha_I^{k+1},\beta_I^{k+1}) : I \in \mathcal{I}^{k+1}\}$ so that

$$\boldsymbol{x}^* \text{ remains an optimal solution of } \mathrm{P}^{k+1} \ \text{ and } \hat{\zeta}^k \leq \hat{\zeta}^{k+1} \leq \zeta^{k+1} = \zeta. \tag{17}$$

In the numerical experiment whose results are reported in Section 4.4, the Lagrangian-DNN relaxation [14] (see also [5, 4]) for $\widehat{\mathrm{P}}^k$ and $\widehat{\mathrm{P}}_a^k(I')$ described below was employed, and NewtBracket [15] (the Newton-bracketing method [16]) was applied to them for their optimal values $\hat{\zeta}^k$ and $\hat{\zeta}^k(I')$, respectively.

For simplicity of discussion, we first deal with the case where $\mathcal{I}^{k+1} = \mathcal{I}^k$. For each $I' \in \mathcal{I}^k$, we consider the following problem:

$$\mathrm{P}_a^k(I') : \zeta^k(I') \ = \ \min\left\{ \boldsymbol{u}^T\boldsymbol{R}\boldsymbol{u} : \begin{array}{l} \boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S, \ \sum_{i\in I} x_i \geq \alpha_I^k \ (I \in \mathcal{I}^k \backslash I'), \\ \sum_{i\in I'} x_i \leq \alpha_{I'}^k + \beta_{I'}^k \end{array} \right\},$$

and solve its conic relaxation $\widehat{\mathrm{P}}_a^k(I')$ to compute its optimal value $\hat{\zeta}_a^k(I')$. If $\eta < \hat{\zeta}_a^k(I')$, let

$$\alpha_{I'}^{k+1} = \alpha_{I'}^k + \beta_{I'}^k + 1 \ \text{ and } \beta_{I'}^{k+1} = \min\left\{\beta_{I'}^k, |I'| - \alpha_{I'}^{k+1} - 1\right\}.$$

Otherwise, let

$$\alpha_{I'}^{k+1} = \alpha_{I'}^k \text{ and } \beta_{I'}^{k+1} = \lfloor \beta_{I'}^k/2 \rfloor.$$

Thus $\{(I,\alpha_I^k,\beta_I^k) : I \in \mathcal{I}^k\}$ has been updated to $\{(I,\alpha_I^{k+1},\beta_I^{k+1}) : I \in \mathcal{I}^{k+1}\}$. From the discussion in Sections 3.3 and 3.5, we see that (17) holds.

4.3.1 An algorithm for generating $\mathcal{I}^0$

If $\boldsymbol{x}^*$ were known in advance, it would be easy to construct an ideal cutting inequality of the form $\sum_{i\in I} x_i \geq \alpha$ such that $\left\{\boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S : \sum_{i\in I} x_i \geq \alpha\right\} = \{\boldsymbol{x}^*\}$.In fact, we could take $I = \{i : x_i^* = 1\}$ and $\alpha = m$, which is impossible. If a branch-and-bound method, for instance, is applied to solve QUBO (13), then more accurate information on the location of optimal solutions becomes available as it proceeds. In such a case, it is reasonable to incorporate such information into $\{(I, \alpha_I^k, \beta_i^k) : I \in \mathcal{I}^k\}$. This will be discussed in Section 4.3.2.

For the case where no information on the location of the optimal solutions of QUBO (13) is available, we propose 'to distribute the cutting inequalities uniformly'. There still remains a great deal of flexibility in choosing $\mathcal{I}^0$ to initialize the sequence (14). In general, as the members of $\mathcal{I}^0$ increase, a tighter lower bound $\hat{\zeta}^k$ for the optimal value ζ of P (13) at each kth iteration can be expected.

Let us show a simple example of $\mathcal{I}^0$ below, which may provide an idea for a general choice of $\mathcal{I}^0$.

Step 0: Let $r = 0$. $\mathcal{J}^0 = \{\{1,\dots,m\}\}$.

Step 1: If $|J| = 1$ for all $J \in \mathcal{J}^r$ then let

$$\mathcal{J} \;=\; \bigcup_{p=0}^{r} \mathcal{J}^p,\ \mathcal{I}^0 \;=\; \mathcal{J} \bigcup \{\{m+j : j \in J\} : J \in \mathcal{J}\},$$

and stop.

Step 2: Let $\mathcal{J}^{r+1} = \emptyset$. For every $J \in \mathcal{J}^r$ with $|J| \geq 2$, choose two subsets J_1 and J_2 of J (randomly) such that $J_1 \cup J_2 = J$ and $|J_1| = |J_2| = \lceil |J|/2 \rceil$ (then $|J_1 \cap J_2| \leq 1$), and add them to $\mathcal{J}^{r+1}$.

Step 3: Let $r = r + 1$ and go to Step 1.

If $n = 4$, the above algorithm generates

$$\begin{aligned}
&\mathcal{J}^0 = \{\{1,2,3,4\}\},\ \mathcal{J}^1 = \{\{1,3\},\{2,4\}\},\ \mathcal{J}^2 = \{\{1\},\{2\},\{3\},\{4\}\}, \\
&\mathcal{J} = \bigcup_{p=1}^{2} \mathcal{J}^p = \{\{1,2,3,4\},\{1,3\},\{2,4\},\{1\},\{2\},\{3\},\{4\}\}, \\
&\mathcal{I}^0 = \mathcal{J} \bigcup \{\{5,6,7,8\},\{5,7\},\{6,8\},\{5\},\{6\},\{7\},\{8\}\}.
\end{aligned}$$

4.3.2 Adding more cutting inequalities at each iteration

We now consider the case where some information of the location of the optimal solutions of P (13) is available, and discuss how it can be used in the construction of $\mathcal{I}^{k+1}$. Suppose that we have generated $\{(I, \alpha_I^k, \beta_I^k) : I \in \mathcal{I}^k\}$ at which (16) holds. Assume that the information is given as an $\bar{\boldsymbol{x}} = (\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}) \in [0,1]^{2m}$ but not necessary $\bar{\boldsymbol{x}} \in \{0,1\}^{2m}$, which is obtained from an optimal solution of a conic (SDP and DNN) relaxation of P^k (15). We note that $\bar{\boldsymbol{x}} = (\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}})$ satisfies $\bar{\boldsymbol{u}} + \bar{\boldsymbol{v}} = \boldsymbol{e}$ approximately, but may not satisfy the complementarity

$\bar{u}_i\bar{v}_i = 0$ $(i = 1,\dots,m)$. In this case, for the computation of an approximate solution $\hat{\boldsymbol{u}} \in \{0,1\}^m$ of QUBO (12), rounding is frequently applied to $\bar{\boldsymbol{u}} \in [0,1]^m$ and/or a heuristic method such as the tabu search [10] and the genetic algorithm [22] to QUBO (12) with the initial solution $\bar{\boldsymbol{u}} \in [0,1]^m$.

For the construction of a family $\mathcal{I}^+$ of subsets of $\{1,\dots,2m\}$ to be added to $\mathcal{I}^k$, each $\bar{u}_i \in [0,1]$ and $\bar{v}_i \in [0,1]$ are regarded to represent the probability $\Pr\{u_i^* = 1\}$ and $\Pr\{v_i^* = 1\})$, respectively, for the unknown optimal solution $\boldsymbol{x}^* = (\boldsymbol{u}^*,\boldsymbol{v}^*)$ of P (13), and $2q$ points $\boldsymbol{u}^p \in \{0,1\}^s$ $(p = 1,\dots,q))$ and $\boldsymbol{v}^p \in \{0,1\}^s$ $(p = 1,\dots,q))$ are generated randomly using the probability. Then, let

$$\begin{aligned} \mathcal{I}^+ &= \{\{i : u_i^p = 1\} : p = 1,\dots,q\} \bigcup \{\{i+m : v_i^p = 1\} : p = 1,\dots,q\}, \\ \mathcal{I}^{k+1} &= \mathcal{I}^k \bigcup \mathcal{I}^+. \end{aligned}$$

We took $q = 10$ in the numerical experiment presented in Section 4.4.

Now, we consider the case where an approximate optimal solution $\hat{\boldsymbol{x}} = (\hat{\boldsymbol{u}},\hat{\boldsymbol{v}}) \in S$ of P (13), which is likely to be optimal but has not been proved to be optimal, is known with the objective value $\eta = \bar{\boldsymbol{u}}^T\boldsymbol{R}\bar{\boldsymbol{u}}$. Note that $\zeta \leq \eta$ is guaranteed. Such a case frequently occurs when we try to solve P (13) by a high performance heuristic method. Let $\widehat{I} = \{i : \hat{x}_i = 1\}$. Then $|\widehat{I}| = m$. For every nonempty subset J of $\widehat{I}$ and $\alpha_J \in \{0,\dots,|J|-1\}$, consider the following problem:

$$\mathrm{P}_a^k(J) : \zeta_{\mathrm{a}}^k(J) \quad = \quad \min \left\{ \boldsymbol{u}^T\boldsymbol{R}\boldsymbol{u} : \begin{array}{l} \boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S,\ \sum_{i \in I} x_i \geq \alpha_I^k\ (I \in \mathcal{I}^k), \\ \sum_{i \in J} x_i \leq \alpha_J \end{array} \right\},$$

and solve its conic relaxation $\widehat{\mathrm{P}}_a^k(J)$ to compute its optimal value $\hat{\zeta}_a^k(J)$. If $\eta < \hat{\zeta}_a^k(J)$ holds, then we know that $\sum_{i\in J} x_i \geq \alpha_J + 1$ is a cutting inequality for the set of optimal solutions of P (13). Moreover, if $\alpha_J + 1 = |J|$, then x_i can be fixed to $x_i = 1$ for all $i \in J$. If we take $J = \widehat{I}$ and $\alpha_J + 1 = |\widehat{I}| = m$, then $\eta < \hat{\zeta}_a^k(\widehat{I})$ provides a certificate for $\hat{\boldsymbol{x}} = (\hat{\boldsymbol{u}},\hat{\boldsymbol{v}}) \in S$ to be the unique optimal solution of P (13). Therefore, it is reasonable to include $\widehat{I}$ and/or some of its subsets J in $\mathcal{I}^+$.

Remark 4.1. For the case above, a branching can be used instead of cutting inequalities to efficiently solve P (13) to optimality. More precisely, for each $\alpha = 0,1,\dots,m$, let

$$S(\alpha) = \left\{ \boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S : \sum_{i\in\widehat{I}} x_i = \alpha \right\}.$$

Then, P (13) is branched into $1+m$ subproblems

$$\mathrm{P}(\alpha) : \zeta(\alpha) = \{\boldsymbol{u}^T\boldsymbol{R}\boldsymbol{u} : \boldsymbol{x} = (\boldsymbol{u},\boldsymbol{v}) \in S(\alpha)\}\ (\alpha = 0,1,\dots,m).$$

The important features of this branching are:

- As α increases from 0 to m, $|S(\alpha)|$ deceases to $|S(m)| = 1$, so that subproblem $\mathrm{P}(\alpha)$ with a larger α is easier to solve.

- As α decreases from m to 0, the optimal value $\zeta(\alpha)$ of subproblem $\mathrm{P}(\alpha)$ is expected to increase, larger than η. As a result, the possibility that $\mathrm{P}(\alpha)$ is pruned by the lower bounding procedure using its conic relaxation is increased.

Further investigation of this branching is beyond the scope of the paper. It will be investigated in our future work.

4.4 Preliminary numerical results

An experimental method on QUBOs has been described in Section 4.3 for evaluating the performance of SCIT presented in Section 3.5. We applied the method to 60 QUBO instances from BIQMAC [23]:

bqp100-1,...,bqp100-10,be120.3.1,...,be120.3.10,be120.8.1,...,be120.8.10,
be150.3.1,...,be150.3.10,be150.8.1,...,be150.8.10,bqp250-1,...,bqp250-10.

The experiments were performed on iMac Pro with Intel Xeon W CPU (3.2 GHZ), 8 cores and 128 GB memory for the instances with dimensions 100, 120 and 150, and Intel Xeon 4216 2 CPUs with 32 cores and 128 GB memory for the instances with dimension 250.

As the optimal value ζ of each instance above is known, its upper bound η was set to ζ to ensure the best performance of SCIT. Recall that η is used in the certificate $\eta < \hat{\zeta}_{\mathrm{a}j}$ for $\boldsymbol{g}_j^T\boldsymbol{x} \geq \alpha_j + 1$ to be a cutting plane, where $\hat{\zeta}_{\mathrm{a}j}$ denotes the optimal value of $\widehat{\mathrm{P}}_{\mathrm{a}j}$ (9). As a smaller $\eta \geq \zeta$ is chosen (or η closer to ζ), more inequalities can become valid cutting inequalities. Thus, ζ is the best choice in our experiment. In the case where ζ is not known, η is usually obtained by a heuristic method as it needs to be the best known upper bound of the optimal value ζ.

For the conic relaxation $\widehat{\mathrm{P}}^k$ of P^k and $\widehat{\mathrm{P}}^k_{\mathrm{a}}(I')$ of $\mathrm{P}^k_{\mathrm{a}}(I')$ $(I' \in \mathcal{I}^k)$, the Lagrangian-DNN relaxation [14, 5, 4] of P^k and $\mathrm{P}^k_{\mathrm{a}}(I')$ $(I' \in \mathcal{I}^k)$ was employed, respectively, and NewtBracket [15] (the Newton-bracketing method [16]) as a numerical method to compute their optimal values $\hat{\zeta}^k$ and $\hat{\zeta}^k_{\mathrm{a}}(I')$ $(I' \in \mathcal{I}^k)$. $\mathcal{I}^0$ was constructed as described in Section 4.3.1. An approximate optimal solution $\overline{\boldsymbol{X}}$ of $\widehat{\mathrm{P}}^k$ was also computed, and an $\bar{\boldsymbol{x}} \in [0,1]^{2m}$ described in Section 4.3.2 from $\overline{\boldsymbol{X}}$ was obtained for the information on the location of the optimal solutions P^k. We used $q = 10$ for the additional cut inequalities associated with $\mathcal{I}^+$. $\mathcal{I}^0$ contains about $4m$ subsets of $\{1,\ldots,2m\}$, so that approximately $4m$ valid inequalities of the form $\sum_{i \in I} x_i \geq \alpha$ were prepared prior to the 0th iteration, and $|I^+| = 10$ inequalities were added prior to the kth iteration $(k = 1, 2, \ldots)$.

Each kth iteration consists of two phases: the first one for solving $\widehat{\mathrm{P}}^k$ and the second one for solving $\widehat{\mathrm{P}}^k_a(I')$ $(I' \in \mathcal{I}^k)$. After solving the second problems, $\{(I, \alpha^k_I, \beta^k_I) : I \in \mathcal{I}^k\}$ was updated to $\{(I, \alpha^{k+1}_I, \beta^{k+1}_I) : I \in \mathcal{I}^{k+1}\}$ as described in Sections 4.3. Since $\mathcal{I}^0$ contains a singleton $I = \{i\}$ $(i = 1, \ldots, 2m)$, the variable x_i was fixed to 1 when the inequality $x_i \geq 1$ became valid cutting inequality, and the number of free variables was reduced among $u_1 = x_1, \ldots, u_m = x_m$ (= the dimension of subQUBO denoted by d^k in Table 1) as well as the redundant cutting inequalities were removed.

The iteration terminated when $\hat{\zeta}^k$ attained the optimal value ζ in 10 iterations $k = 0, 1, \ldots, 9$ or k reached 9. Among 60 QUBO instances, 42 cases attained the optimal value within 10 iterations, *i.e.*, $\hat{\zeta}^k = \zeta$ for some $k <= 9$. The other 18 cases failed to obtain the exact optimal value. Table 1 shows the numerical results on 42 successful instances.

The following two aspects are crucial for evaluating the performance of SCIT:

(i) How many variables $u_1 = x_1, \ldots, u_m = x_m$ are fixed to either 0 or 1, which can be measured by the decrease of d_k $(k = 0, 1, \ldots)$.

(ii) Improvement in the lower bound of optimal value ζ, which can be observed by the increase of $\hat{\zeta}^k$ $(k = 0, 1, \ldots)$.

Overall, the method worked effectively in terms of the two aspects, but less effectively for larger dimensional cases; it took more iterations to attain a smaller d^k and a tighter $\hat{\zeta}^k$ to ζ. In practice, (i) is an important aspect of SCIT when it is combined with a numerical method for solving QUBOs. For the QUBO instance bqp250-2 in our numerical experiment, $237 (= 250 - 13)$ variables among $u_1, \ldots, u_{250}$ of QUBO (13) were fixed to 0 or 1 after 10 iterations, so the resulting subQUBO with 13 variables was easy to solve.

The 18 instances where $\hat{\zeta}^k$ could not attain ζ in 10 iterations (or $k \leq 9$) are not included in Table 2. As mentioned before, SCIT alone cannot be a numerical method for solving QUBOs. To successfully solve the 18 instances, a numerical method combining SCIT with other methods should be implemented.

We want to highlight that the exact optimal values of the 42 QUBO instances, 70 % of the 60 instances to which the method was applied, could be obtained in the numerical experiments. The numerical results reported here, though limited, present the promising potential of SCIT, especially when it is combined with the branch-and-bound method [10, 11, 17, 18] and heuristic methods [10, 22] for solving QUBOs.

5 Concluding Remarks

We have presented SCIT, a very flexible framework, to generate effective cutting inequalities for strengthening conic relaxations for computing lower bounds of the optimal value of a binary QOP. To be able to combine the experimental method with the branch-and-bound method, there remain many issues to be studied. In particular, the initial setting of the family of valid inequalities of the form $\sum_{i \in I} x_i \geq \alpha^0_I$ $(I \in \mathcal{I}^0)$ (Section 4.3.1) should be designed more carefully. Another important issue is to investigate how to effectively utilize the optimal solution information of the conic relaxation problem (Section 4.3.2). In addition, extensive numerical experiment is necessary.

The authors' future interests include applying SCIT to the quadratic assignment problem (QAP), which is known to be one of the most difficult combinatorial problems. They have participated in the joint project for solving large scale QAPs by the branch-and-bound method. See [9] for an intermediate report on the project. For the lower bounding procedure, the Lagrangian doubly nonnegative (DNN) relaxation [14, 5, 4] and the Newton-bracketing method [16, 15], which were used in the numerical results reported in Section 4.5, have been employed in the project. For the first time, tai30a and sko42 from QAPLIB [1, 2] were solved using their method. Although there still remain many unsolved instances in QAPLIB, approximate optimal solutions which are likely to be optimal are known in all of those instances. The additional cutting inequalities discussed in Section 4.3.2 are expected to work effectively to prove that they are truly optimal. The branching rules mentioned in Remark 4.1 can be also used to prove their optimality.

Table 1: Numerical results on SCIT applied to 42 QUBO instances from BIQMAC [23].

		d^k (Dim. of subQUBO P^k whose L-DNN relaxation $\widehat{\mathrm{P}}^k$ to be solved), $\hat{\zeta}^k$ (the optimal value of $\widehat{\mathrm{P}}^k$)									
QUBO	Opt.Val	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$
bqp100-1	-7970	100, -8036	42, -7970								
bqp100-2	-11036	100, -11036									
bqp100-3	-12723	100, -12723									
bqp100-4	-10368	100, -10368									
bqp100-5	-9083	100, -9083									
bqp100-6	-10210	100, -10341	56, -10291	38, -10270	27, -10248	17, -10220	12, -10210				
bqp100-7	-10125	100, -10159	36, -10125								
bqp100-8	-11435	100, -11435									
bqp100-9	-11455	100, -11455									
bqp100-10	-12565	100, -12565									
be120.3.1	-13067	120, -13343	93, -13268	78, -13204	58, -13135	35, -13075	9, -13067				
be120.3.2	-13046	120, -13163	44, -13046								
be120.3.3	-12418	120, -12609	75, -12477	30, -12418							
be120.3.4	-13867	120, -14039	71, -13939	31, -13868	2, -13867						
be120.3.5	-11403	120, -11558	59, -11407	12, -11403							
be120.3.6	-12915	120, -13022	46, -12915								
be120.3.7	-14068	120, -14128	27, -14068								
be120.3.8	-14701	120, -14812	40, -14701								
be120.3.10	-12201	120, -12413	83, -12298	46, -12202	4, -12201						
be120.8.2	-18827	120, -19351	102, -19271	96, -19167	83, -19065	61, -18903	29, -18827				
be120.8.3	-19302	120, -19791	102, -19653	80, -19509	50, -19396	19, -19302					
be120.8.4	-20765	120, -21063	65, -20824	12, -20765							
be120.8.5	-20417	120, -20677	46, -20457	21, -20417							
be120.8.6	-18482	120, -18954	98, -18804	74, -18615	35, -18482						
be120.8.9	-18195	120, -18685	101, -18539	80, -18384	49, -18231	23, -18195					
be120.8.10	-19049	120, -19380	66, -19157	28, -19055	6, -19049						
be150.3.1	-18889	150, -19202	117, -19098	89, -18978	50, -18889						
be150.3.2	-17816	150, -18200	129, -18123	110, -18069	102, -17982	76, -17861	36, -17816				
be150.3.3	-17314	150, -17510	79, -17315	13, -17314							
be150.3.4	-19884	150, -20080	82, -19917	27, -19884							
be150.3.5	-16817	150, -17216	-139, 17159	123, -17092	104, -16998	85, -16930	63, -16846	46, -16817			
be150.3.7	-18001	150, -18385	132, -18331	117, -18248	93, -18141	62, -18052	27, -18001				
be150.8.4	-26911	150, -27685	142, -27611	136, -27561	133, -27476	130, -27455	128, -27423	122, -27371	113, -27257	92, -27108	58, -26911
be150.8.5	-28017	150, -28634	116, -28470	96, -28343	68, -28198	44, -28076	26, -28022	6, -28017			
be150.8.10	-28374	150, -29125	139, -29071	132, -28950	118, -28894	112, -28810	102, -28691	84, -28607	69, -28492	43, -28374	
bqp250-1	-45607	250, -46244	214, -46102	181, -45927	134, -45719	52, -45607					
bqp250-2	-44810	250, -45585	241, -45551	230, -45469	219, -45413	211, -45346	199, -45281	178, -45186	156, -45085	109, -44841	13, -44810
bqp250-3	-49037	250, -49457	163, -49144	56, -49037							
bqp250-4	-41274	250, -42009	237, -41909	223, -41805	198, -41668	166, -41478	92, -41274				
bqp250-5	-47961	250, -48431	164, -48153	82, -48040	45, -47961						
bqp250-7	-46757	250, -47378	215, -47236	182, -47086	130, -46904	59, -46757					

References

[1] QAPLIB – A Quadratic Assignment Problem Library, Computtonal Optimization Research at Lehigh. https://coral.ise.lehigh.edu/data-sets/qaplib/, August 2011.

[2] M. F. Anjos. "QAPLIB is a a Quadratic Assignment Problem Library" in Miguel Anjos' Homepage. https://www.miguelanjos.com/qaplib.

[3] M. F. Anjos and J. B. Lasserre. *Handbook on Semidefinite, Conic and Polynomial Optimization*, volume 166 of *Internstional Series in Operations Research & Management Science*. Sprignger, 2012.

[4] N. Arima, S. Kim, M. Kojima, and K. C. Toh. Lagrangian-conic relaxations, Part I: A unified framework and its applications to quadratic optimization problems. *Pacific J. Optim.*, 14(1):161–192, 2018.

[5] N. Arima, S. Kim, M. Kojima, and K.C. Toh. A robust Lagrangian-DNN method for a class of quadratic optimization problems. *Comput. Optim. Appl.*, 66:453–479, 2017.

[6] E. Bals, S. Ceria, and G. Cornuejols. A lift-and-project cutting plane algorithm for mixed 0–1 programs. *Math. Program.*, 58:295–324, 1993.

[7] P. Bonami, A. Lod, J. Scheiger, and A. Tramontani. Solving quadratic programming by cutting plane. *SIAM J. Optimization*, 29:1076–1105, 2019.

[8] A. Engau, M. F. Anjos, and Vannelli. A. An improved interior-point cutting-plane method for binary quadratic optimization. *Electron. Notes in Discret. Math.*, 36:743–750, 2010.

[9] K. Fujii, N. Itoh, N. Kim, M. Kojima, Y. Shinano, and K. C. Toh. Solving challenging scale QAPs. Technical Report ZIB-Report-21-02, Zuse Institute Berlin, 14195 Berlin, Germany, January 2021.

[10] F. Glover and M. Laguna. Tabu search. In D. Du and P. M. Pardalos, editors, *Handbook of Combinatorial Optimization*, volume 3, pages 2093–2229. Springer, 1998.

[11] D. Guimaraes, A., A. S. da Cunha, and d. L. Perera. Semidefinite programming lower bounds and branch-and-bound algorithms for the quadratic minimum spanning tree problem. *Eur. J. Oper. Res.*, 280:46–58, 2020.

[12] H. Helmberg and F. Rendl. Solving quadratic (0,1)-problems by semidefinite programs and cutting planes. *Math. Program.*, 82:291–315, 1998.

[13] N. Ito, S. Kim, M. Kojima, A. Takeda, and K.C. Toh. Equivalences and differences in conic relaxations of combinatorial quadratic optimization problems. *J. Global Optim.*, 72(4):619–653, 2018.

[14] S. Kim, M. Kojima, and K. C. Toh. A Lagrangian-DNN relaxation: A fast method for computing tight lower bounds for a class of quadratic optimization problems. *Math. Program.*, 156:161–187, 2016.

[15] S. Kim, M. Kojima, and K.C. Toh. User manual of newtbracket: "A Newton-Bracketing method for a simple conic optimization problem" with aopplications to QOPs in binary variables. https://sites.google.com/site/masakazukojima1/softwares-developed/newtbracket, November 2020.

[16] S. Kim, M. Kojima, and K.C. Toh. A Newton-bracketing method for a simple conic optimization problem. *To appear in Optim. Methods and Softw.*, 36(1):371–388, 2021.

[17] N. Krislock, Malick J., and F. Roupin. Improved semidefinite bounding procedure for solving max-cut problems to optimality. *Math. Program.*, 143:61–86, 2014.

[18] N. Krislock, J. Malik, and F. Roupin. BiqCrunch: A semidefinite branch-and-bound method for solving binary quadratic problems. *ACM Trans. Math. Soft.*, 43(4, Article 32), 2017.

[19] J. B. Lasserre. An explicit exact SDP relaxation for nonlinear 0-1 programs. In *Integer Programming and Combinatorial Optimization*, pages 293–303. Springer, 2001.

[20] J. B. Lasserre. Global optimization with polynomials and the problems of moments. *SIAM J. Optim.*, 11:796–817, 2001.

[21] H. Marchand, A. Martin, R. Weismantel, and L. Waosey. Cutting planes in integer and mixed integer programming. *Discret. Appl. Math.*, 123:397–446, 2002.

[22] M. Mitchell. *An Introductioin to Genetic Algorithms.* The MIT Press, 1998.

[23] A. Wiegele. Biq mac library. http://www.biqmac.uni-klu.ac.at/biqmaclib.html, 2007.