Sparsity-Preserving Integration of Convex Curvature Information into Linear Relaxations for Quadratic Unconstrained Binary Optimization

Sven Mallach[1,2]

[1]Network & Data Science Management Group, University of Siegen, Germany
[2]University of Bonn, Germany

August 23, 2026

Abstract

We systematically investigate the potentials of improving the lower bound obtained with a linear relaxation of the Quadratic Unconstrained Binary Optimization problem by integrating curvature information from an accompanying quadratic convex underestimator via gradient inequalities. On the one hand, we exemplify to which extent this hybrid approach may provide a lower bound that is strictly stronger than the maximum of the respective linear and convex quadratic minima standalone. On the other, we prove that a strict improvement is impossible for the subclass of instances obtained from the Maximum Cut problem when choosing the underestimator and relaxation in a theoretically and practically relevant way. In addition, we address how the most common linear relaxation can be strengthened with the well-known triangle inequalities using only linearization variables matching the non-zero pattern of the original cost matrix.

1 Introduction

Quadratic Unconstrained Binary Optimization (QUBO) is the problem of determining a solution $x \in \{0,1\}^n$ of minimum value for the possibly non-convex function $\gamma(x) := x^\mathsf{T} A x + b^\mathsf{T} x$, where the symmetric matrix $A \in \mathbb{R}^{n \times n}$ and the vector $b \in \mathbb{R}^n$ are the given parameters. The problem is $\mathcal{NP}$-hard in general [11], and of overarching relevance with several immediate and indirect applications, see for instance the listing in the article [12] that is part of the recent book [26] on QUBO. Fundamental results regarding the problem have been established many decades ago, see e.g. [17, 18, 25, 8].

A key ingredient for solving QUBO instances to optimality is to have a tractable relaxation of the problem that provides a strong lower bound on the minimum of $\gamma(x)$ over all $x \in \{0,1\}^n$. Most prevalent approaches rely on bounds obtained from linear programming (LP) [6, 28], convex quadratic programming (QP) [2], or semidefinite programming (SDP) [23, 15, 30] relaxations. In these approaches, the domain of the variables x is then regularly relaxed to $[0,1]^n$, i.e., to the unit hypercube of dimension n, except for that SDP approaches frequently use a transformation to the domain $[-1,1]^n$. Moreover, standard LP and SDP approaches introduce additional linearization variables to substitute for the bilinear products in γ while it is straightforward to exploit a potential sparsity of A, i.e. to introduce a variable supposed to represent $x_i \cdot x_j$ only if $A_{ij} \neq 0$, in the LP case.

In this paper, we investigate the potentials of a hybrid approach that integrates curvature information, obtained from a convex quadratic underestimator f of γ, into a linear programming relaxation for QUBO. Our method is thereby conceptually similar to algorithms that outer approximate the feasible region of convex *constraints* by gradient inequalities such as e.g. Kelley's algorithm [22], but it employs f as an "alternative rating" of relaxation solutions instead, and as a means to improve the bound provided by the linear relaxation if it falls below this rating. It is also conceptually similar to methods that aim to outer approximate the positive semidefinite cone by eigenvector inequalities and it has further links to SDP as we address below in more detail.

Our main goal is to systematically explore, from a theoretical as well as practical perspective, to which extent such a hybrid approach may result in a bound that is strictly stronger than either of the LP and QP bounds standalone. Besides that, of course, it may already be of advantage to reach the QP bound within a linear relaxation if it is larger than the LP bound, and our algorithm provides a principle means to do so. On the one hand, we thus keep the description of our method as general as possible, allowing the relaxation to take the form of a linear minimization oracle satisfying very mild assumptions. Principally, any linear relaxation of the Boolean Quadric Polytope [25] could be employed. On the other, we have standard and efficiently computable representatives of both the underestimator and the relaxation in mind. Specifically, a straightforward choice is an underestimator obtained by means of the SDP used in the Quadratic Convex Reformulation (QCR) method [2], and the LP relaxation of the commonly used "standard linearization" [10, 14] with McCormick inequalities [24].

A scenario of special interest and practical relevance arises when this relaxation is further strengthened with the well-known triangle inequalities detailed in Section 2. For this particular setting, we show how to preserve the (otherwise implicitly given) adaptivity of our approach to the non-zero (or sparsity) pattern of A. Moreover, we show that the bound achieved with our method is equal to the maximum of the two bounds provided by the original relaxation and f if $A = -L(G)$, where $L(G)$ is the (weighted) Laplace matrix of an undirected graph, and $b = 0$. This case is particularly relevant as it appears precisely if an instance of the Maximum Cut problem is straightforwardly transformed into an instance of QUBO [8]. At the same time, a sometimes significant improvement over the maximum of the QP and LP bounds is sustained for other QUBO instances, and we reveal the actual quality of these bounds for established benchmark instances.

Another relevant aspect of this work is that it links and provides an alternative – with trade-offs – to methods that strengthen LP relaxations by eigenvector (or "psd") inequalities, like e.g. in [27, 7, 16]. In fact, one could substitute them for the gradient inequalities respectively the convex underestimator used in our presentation and so obtain an according approach for QUBO. As opposed to the gradient-based setting, the theoretically achievable bound is then easy to anticipate – it is precisely the one of the SDP that involves the constraints of the chosen linear relaxation. Practically solving, in this way, a relaxation that provides a stronger bound than the SDP used in the QCR method is however prevalently costly. At the same time, the gradient-based approach does not require any eigenvector or eigenvalue calculations, and it can be implemented so as to preserve the original non-zero pattern of the matrix A (or likewise, its pendant in the convex underestimator f) in a natural way.

Finally, it is worth emphasizing that our method is not restricted to QUBO in principle. Rather, we believe that the presentation is more accessible when restricting it to QUBO, and that this context already provides the aforementioned interesting aspects to focus on. Besides that, additional constraints might impede the (here to a certain extent assumed) fast derivation of a convex quadratic underestimator.

The paper is organized as follows. In Section 2, we provide preliminaries and a more formal discussion of the aforementioned motivating aspects, underestimators, and linear relaxations. Moreover, we review closely related work, in particular techniques that are used, referenced, or extended in this paper. We then present our hybrid relaxation approach in Section 3 along with our theoretical results, and evaluate its practical performance on a selection of established benchmark instances in Section 4. Finally, a brief conclusion is given in Section 5.

2 Preparations, Preliminaries, and Related Work

We consider the Quadratic Unconstrained Binary Optimization (QUBO) problem

$$\min_{x\in\{0,1\}^n} \gamma(x) := x^\mathsf{T}Ax + b^\mathsf{T}x$$

where $A \in \mathbb{R}^{n\times n}$ is symmetric but not necessarily positive semidefinite, and $b \in \mathbb{R}^n$. We denote by $\gamma^* := \min_{x\in\{0,1\}^n} \gamma(x)$ the value of an optimum solution to this problem.

2.1 Convex Quadratic Programming Relaxations

Since γ is a non-convex function in general, we assume that we are also given a convex underestimation ("convexification") of γ on the domain $[0,1]^n$ of the form

$$f(x) := x^\mathsf{T}Qx + c^\mathsf{T}x$$

where now Q is positive semidefinite, $f(x) \le \gamma(x)$ for all $x \in [0,1]^n$, and $f(x) = \gamma(x)$ whenever $x \in \{0,1\}^n$ whence in particular $\min_{x\in\{0,1\}^n} f(x) = \gamma^*$. Given a matrix A with negative eigenvalues, such a convexification can be constructed efficiently, for instance by choosing $Q = A + \mathrm{Diag}(u)$ and $c = b - u$ with $u = -\lambda_A^{\min} e$ where $\lambda_A^{\min}$ is the smallest eigenvalue of A, $\mathrm{Diag}(u)$ is the diagonal matrix built from u, and e is the all-ones vector of dimension n [2]. The vector $u \in \mathbb{R}^n$ that gives the tightest such "diagonal" convex underestimation can be obtained via the QCR method [2] that involves solving a semidefinite program. Naturally, if A is already positive semidefinite, then one may just set $Q = A$ and $c = b$.

For ease of reference, we define the QP relaxation bound $f_{\mathrm{QP}}^* := \min_{x\in[0,1]^n} f(x)$. We emphasize that f_{QP}^* need not necessarily be known a priori in order to apply our approach presented in Section 3. Rather, we will employ the curvature information available from f by making use of the fact that, since f is convex on the unit hypercube, the following *gradient inequalities*, where here $\nabla f(y) = (2Qy + c)$ is the gradient of f at y, are satisfied:

$$f(x) \geq f(y) + \nabla f(y)^{\mathsf{T}}(x - y) \text{ for all } 0 \leq x, y \leq 1$$

Recall in this context that the true global minimum of f is not necessarily attained on the unit hypercube whence the gradient of f may be non-zero in the entire domain of interest. Moreover, minimizers $\bar{x}$ of f on $[0,1]^n$ frequently reside in the interior. Therefore, it is not straightforward to strengthen the bound f^*_{QP} in the original variable space, besides by performing a branching step. Here lies a potential in the hybrid approach considered in this paper which combines the (possible) strength of f^*_{QP} with strengthenings by linear inequalities in the extended space that results from the linearization.

2.2 Linear Programming Relaxations

Using linear relaxations to obtain a lower bound for γ^* has a long tradition, as is evident for instance by Padberg's work on the Boolean Quadric Polytope [25]. They are the fundament of state-of-the-art LP-based solvers such as [6, 28] which are especially successful for QUBO instances with a *sparse* matrix A.

Of course, working with linear relaxations first requires some form of *linearization* of the original problem. To formalize this, we thus first define the linear function g that serves as an underestimator for γ. Since this implicitly displays useful relations, we further formulate g in terms of a matrix variable $X \in S_n$ where S_n is the set of symmetric $n \times n$ matrices, as is common in the context of semidefinite programming relaxations. Straightforwardly, the linearized objective function thus takes the customary form

$$g(x, X) := \langle A, X \rangle + b^{\mathsf{T}} x.$$

As discussed in the context of our approach presented in Section 3, it will be however more meaningful to define g with respect to Q and c, as we will directly relate g to f.

To ensure that g matches γ (and f) for all $x \in \{0,1\}^n$, and that it provides an underestimation of γ^* on the unit hypercube otherwise, we further need to ensure that the pairs (x, X) satisfy certain box, diagonal, and linearization constraints commonly known as "standard" linearization [10, 14] or McCormick inequalities [24]. We denote these constraints and the separate polyhedra defined by them as follows:

$$P_{\text{Box}} := \{(x, X) : 0 \leq x \leq 1, 0 \leq X \leq 1\}$$

$$P_{\text{Eq}} := \{(x, X) : \text{diag}(X) = x\} := \{(x, X) : X_{ii} = x_i\}$$

$$P_{\text{McC}} := \{(x, X) : X_{ij} \leq x_i, X_{ij} \leq x_j, x_i + x_j - X_{ij} \leq 1 \text{ for } 1 \leq i < j \leq n\}$$

For convenience, we further define $P_{\text{Lin}} := P_{\text{Box}} \cap P_{\text{Eq}} \cap P_{\text{McC}}$. It may be instructive to observe that while the relations

$$g(x, X) \leq \gamma(x) \text{ for all } (x, X) \in P_{\text{Lin}}$$

and

$$g(x, X) = f(x) = \gamma(x) \text{ for all } (x, X) \in P_{\text{Lin}} \text{ where } x \in \{0,1\}^n$$

hold, it may well be, in general, that

$$g(x,X) > f(x) \text{ for some (other) } (x,X) \in P_{\text{Lin}}.$$

In other words, g is *not* (necessarily) a point-wise linear underestimator of f for every x in the feasible domain P_{Lin} of the relaxation of the "standard" linearization, even if g is derived based on f, and even if f is derived via the QCR method. Thus, in particular, the minimum f^*_{QP} of an according quadratic underestimation is not necessarily at least as large as the "basic" LP-based bound $g^*_{\text{Lin}} := \min\{g(x,X) : (x,X) \in P_{\text{Lin}}\}$. Rather, while $g^*_{\text{Lin}} \leq \gamma^*$ is guaranteed, g^*_{Lin} may be in any relation to f^*_{QP}, as is discussed in more detail in Section 3.2. Therefore, the possibility of providing a better lower bound than f^*_{QP} particularly exists for LP-based strengthenings of g^*_{Lin} as described in the next paragraph. At the same time, our method is partly designed to take potential advantage of points (minimizers) x which are "underrated" in the sense that, under the restrictions on the choice of X given by the considered relaxation, $g(x,X) < f(x)$.

We consider a possible further strengthening of g^*_{Lin}, at first in the general form of an arbitrary finite set of valid constraints which are linear w.r.t. the extended variable space,

$$P_{\text{Str}} := \{(x,X) : \langle D_i, X\rangle + d_i^{\mathsf{T}} x \leq r_i,\ i = 1,\ldots,m\}.$$

Then, the according lower bound is $g^*_{\text{Lin+Str}} := \min\{g(x,X) : (x,X) \in P_{\text{Lin}} \cap P_{\text{Str}}\}$. From a computational point of view, it is attractive to choose a strengthening by inequalities that do not have any non-zero entries in D_i that map to zeroes in A respectively Q, since variables X_{ij} with zero coefficients in the objective and constraints need not be introduced in the LP-based setting.

As a concrete example, we investigate the common strengthening by the well-known *triangle inequalities* [25]:

$$X_{ij} - X_{jk} + X_{ik} - x_i \leq 0 \qquad \text{for all } 1 \leq i < j < k \leq n \tag{1}$$
$$X_{ij} + X_{jk} - X_{ik} - x_j \leq 0 \qquad \text{for all } 1 \leq i < j < k \leq n \tag{2}$$
$$-X_{ij} + X_{jk} + X_{ik} - x_k \leq 0 \qquad \text{for all } 1 \leq i < j < k \leq n \tag{3}$$
$$x_i + x_j + x_k - X_{ij} - X_{jk} - X_{ki} \leq 1 \qquad \text{for all } 1 \leq i < j < k \leq n \tag{4}$$

We denote the according polytope by

$$P_{\text{Tr}} := \{(x,X) : (1),(2),(3),(4)\}$$

and the according lower bound by $g^*_{\text{Lin+Tr}} := \min\{g(x,X) : (x,X) \in P_{\text{Lin}} \cap P_{\text{Tr}}\}$. While the matrices associated with the triangle inequalities do not match the non-zero pattern of the matrix A respectively Q unless it has a fully populated strict upper triangle, we present in Section 3.4 best practices to retain an approach that acts only on the variables with a non-zero coefficient in the objective function.

2.3 Semidefinite Programming Relaxations

Another well-studied and closely related approach to obtain a relaxation for QUBO is via semidefinite programming [29]. The elementary SDP relaxation [30] consists of intersecting the diagonal equations defining P_{Eq} with the following non-polyhedral set induced by imposing positive semidefiniteness on an augmented matrix

$$S_{\mathrm{SDP}} \coloneqq \left\{ (x,X) : \begin{bmatrix} 1 & x^{\mathsf{T}} \\ x & X \end{bmatrix} \succeq 0 \right\}$$

where X is precisely the symmetric matrix also considered in Section 2.2.

Let $g^*_{\mathrm{SDP+Eq}} \coloneqq \min\{g(x,X) : (x,X) \in P_{\mathrm{Eq}} \cap S_{\mathrm{SDP}}\}$. In the present context, it is worth to mention that the according SDP is precisely the one that gives the tightest diagonal convexification of γ via the QCR method [2]. Specifically, the optimal dual solutions regarding the constraints $\mathrm{diag}(X) = x$ provide the according vector u addressed in Section 2.1. Therefore, if f is determined this way, one has $f^*_{\mathrm{QP}} = g^*_{\mathrm{SDP+Eq}}$ [2].

Enriching the elementary SDP relaxation with Box, McCormick and triangle inequalities gives a stronger bound which is however also comparably expensive to compute. Except for small instances, this usually requires sophisticated solution methods like in [30, 19, 5]. We denote this bound as $g^*_{\mathrm{SDP+Lin+Tr}} \coloneqq \min\{g(x,X) : (x,X) \in S_{\mathrm{SDP}} \cap P_{\mathrm{Lin}} \cap P_{\mathrm{Tr}}\}$.

Principally, another way to compute $g^*_{\mathrm{SDP+Eq}}$ or $g^*_{\mathrm{SDP+Lin+Tr}}$ via linear programming is to iteratively employ eigenvector (or "psd") inequalities

$$\begin{pmatrix} s & v^{\mathsf{T}} \end{pmatrix} \begin{bmatrix} 1 & x^{\mathsf{T}} \\ x & X \end{bmatrix} \begin{pmatrix} s \\ v \end{pmatrix} \geq 0 \qquad \text{for all } s \in \mathbb{R}, v \in \mathbb{R}^n,$$

where $\begin{pmatrix} s & v^{\mathsf{T}} \end{pmatrix}$ is an eigenvector corresponding to a strictly negative eigenvalue of the respective current solution matrix, as cutting planes [27, 7]. Approximating S_{SDP} this way has a natural connection to our approach as presented in Section 3, especially if the convex underestimator is determined by the QCR method. However, the repeated computation of eigenvectors and eigenvalues is expensive and numerically ill-posed, and eigenvector inequalities are dense in general, which counteracts the potentials of exploiting sparsity in the linearized problem. Deriving sparse violated eigenvector inequalities and maintaining sparsity in general has attracted a lot of research interest in recent years [4, 9, 16]. Despite significant advancements, approaching $g^*_{\mathrm{SDP+Lin+Tr}}$ with linear relaxations remains costly in general.

3 Gradient-based Integration of Convex Curvature Information

In this section, we present our approach to integrate convex quadratic curvature information into a linear relaxation for QUBO. Thereby, our goal is to compute a lower bound on γ^* by means of linear programming techniques that is potentially strictly stronger than, or at least as strong as, the maximum of the bound f^*_{QP} and the one provided by the chosen linear relaxation standalone. We will start with the basic methodology and the intuition behind it, providing a general background. Then, we discuss the different possible scenarios in terms of the proportionality of the lower bounds provided by the convex quadratic and the (pure) linear relaxation. Afterwards, we present an algorithmic description of our hybrid method, and address the special case of employing triangle inequalities without introducing variables with a zero coefficient in the objective. Finally, we focus on a customary setting where the bound achieved with our method is provably equal to the maximum of the two bounds provided by the original relaxation and f if an instance of the Maximum Cut problem is considered.

3.1 A Methodological Baseline

Suppose that we have determined a convex underestimator f for our original objective function γ with associated (known or unknown) lower bound f^*_{QP}. Moreover, for the basic methodology discussed in this subsection, we simply work with the most simple linear relaxation whose feasible set is P_{Box}. Since we will directly relate it to f, we further assume that the linear objective is (re)defined in terms of f (rather than γ) as:

$$g(x,X) := \langle Q,X\rangle + c^{\mathsf{T}}x$$

Using for instance one of the methods discussed in Section 2.1 to derive f, the non-zero off-diagonal patterns of Q and A may thereby be kept consciously coherent. We emphasize that our approach can be implemented to act only on those variables X_{ij}, $i \neq j$, where $Q_{ij} \neq 0$, see also Section 3.4 for more details and an extension to the case of using triangle inequalities.

As an instructive conceptual starting point, consider that we may principally compute (or recover) f^*_{QP} by an outer approximation of the epigraph of f in the variable space of the linearized problem. That is, we (approximately) obtain f^*_{QP} as the optimum solution value when solving

$$\min g(x,X) : (x,X) \in S_{\nabla} \cap P_{\mathrm{Box}}$$

where the set

$$S_{\nabla} := \{(x,X) : g(x,X) \geq f(y) + \nabla f(y)^{\mathsf{T}}(x-y) \text{ for all } 0 \leq y \leq 1\}$$

is induced by all gradient inequalities w.r.t. f, and where we replace the variable non-linear part $f(x)$ on the left-hand side with our linear approximation $g(x,X)$. This has a strong analogy to Kelley's algorithm [22], see also [31], and also strongly relates to (approximately) recovering f^*_{QP} via eigenvector inequalities as addressed in Section 2.3.

Even though we do not propose to do this outer approximation as such, this perspective leads straightforwardly to a helpful understanding regarding the proportionalities between f and g that we subsume in the following statement.

Proposition 1. *Let $(\bar{x},\bar{X}) \in P_{Box}$. Then the gradient inequality for $y = \bar{x}$ in terms of S_{∇} cuts off $(\bar{x},\bar{X})$ if and only if $g(\bar{x},\bar{X}) < f(\bar{x})$.*

Proof. For the reference vector $y = \bar{x}$, the gradient inequality $\langle Q,X\rangle + c^{\mathsf{T}}x \geq y^{\mathsf{T}}Qy + c^{\mathsf{T}}y + (2Qy+c)^{\mathsf{T}}(x-y)$ rewrites to $\langle Q,X\rangle + c^{\mathsf{T}}x \geq \bar{x}^{\mathsf{T}}Q\bar{x} + c^{\mathsf{T}}\bar{x} + (2Qy+c)^{\mathsf{T}}(x-\bar{x})$. Evaluating it then for $X = \bar{X}$ and $x = \bar{x}$, the right-most summand vanishes and thus the inequality is violated if and only if $g(\bar{x},\bar{X}) < f(\bar{x})$. □

The intuition behind this simple proposition that we will actually use is that if $(\bar{x},\bar{X})$ is our current relaxation solution, then it can be cut off in this fashion if and only if $(\bar{x},\bar{X})$ is "underrated" by g compared to the objective value of $\bar{x}$ regarding f. Indeed, we do not even need to know f^*_{QP} in order to exploit this. Gradient inequalities in terms of S_{∇} simply cut off (x,X) combinations that give a g-rating that is lower than the f-rating, concomitant with the potential to improve the current lower bound of the relaxation. Since a minimizer of f is feasible also for the linear relaxation, and its accompanying X is only restricted by gradient inequalities besides the box constraints, an iterative application therefore also guarantees that the final solution $(\bar{x},\bar{X})$ obtained this way (approximately) satisfies $g(\bar{x},\bar{X}) = f^*_{\text{QP}}$.

Observing in addition that one has $\bar{x}_i^2 < \bar{x}_i$ in an optimum solution $\bar{x}$ of the convex quadratic program $\min_{x\in[0,1]^n} f(x)$ unless $\bar{x}_i \in \{0,1\}$, it further becomes apparent that $S_\nabla \cap P_{\text{Box}} \cap P_{\text{Eq}}$ may already be a stronger relaxation than $S_\nabla \cap P_{\text{Box}}$. Moreover, adding McCormick and further inequalities may clearly provide an even stronger one.

3.2 Possible Scenarios when Relating the LP and QP Bounds

As mentioned in Section 2.2, the optimum value g^* of some chosen linear relaxation (without any added gradient inequalities) may be in either possible relation to f^*_{QP}. We now consider more in detail the scenarios that one may encounter, focusing on the question whether an improvement of g^* is possible via gradient inequalities based on f. Thereby, we generally consider an arbitrary linear relaxation whose feasible set P satisfies $P \subseteq P_{\text{Box}}$ and discuss potential further more specific cases during the course.

The case $g^* > f^*_{\text{QP}}$: By the discussion in Section 3.1, if the lower bound g^* obtained from the linear relaxation already supersedes f^*_{QP}, we must have chosen $P \subsetneq P_{\text{Box}}$. Moreover, the curvature information on f available via gradient inequalities may still be of help in improving on g^* in this case. To this end, let $(\bar{x},\bar{X})$ be an optimal solution to the linear relaxation. Then, by Proposition 1, a necessary condition to increase the bound in this fashion is that $\bar{x}$ is underrated with respect to f. This condition is however not sufficient in general, since $(\bar{x},\bar{X})$ need not be a unique minimizer of the linear relaxation. Thus, after cutting off $(\bar{x},\bar{X})$ by a gradient inequality, we might obtain another point $(\hat{x},\hat{X})$ with $g(\hat{x},\hat{X}) = g(\bar{x},\bar{X})$ and $\hat{x} \neq \bar{x}$, which may or may not be underrated, or more fortunately, a point $(\hat{x},\hat{X})$ with $g(\hat{x},\hat{X}) > g(\bar{x},\bar{X})$ which may or may not be underrated as well. In our experiments, we found several cases where the finally obtained minimizer (x^*,X^*) has $g(x^*,X^*) > g^*$, see Section 4.

The case $g^* = f^*_{\text{QP}}$: In this scenario, it is worth differentiating whether $P = P_{\text{Box}}$ or $P \subsetneq P_{\text{Box}}$. For if the linear relaxation does not have additional restrictions beyond the box constraints, then f^*_{QP} is clearly the best achievable bound by means of gradient inequalities in terms of f because $f^*_{\text{QP}} = \min g(x,X) : (x,X) \in S_\nabla \cap P_{\text{Box}}$ as discussed in Section 3.1. Otherwise, i.e. if $g^* = f^*_{\text{QP}}$ happens to appear "accidentally", then (the bound provided with only the box constraints might be lower than f^*_{QP} and) a minimizer $(\bar{x},\bar{X})$ of g satisfying the additional constraints may be underrated. Thus, it may possibly be cut off by a gradient inequality and the bound may or may not improve then just as in the previous case.

The case $g^* < f^*_{\text{QP}}$: If $P \subseteq P_{\text{Box}}$ and the lower bound obtained from the chosen linear relaxation is strictly smaller than f^*_{QP}, then the optimal point $(\bar{x},\bar{X})$ computed is clearly underrated. Therefore, an improvement is possible by gradient inequalities at least until f^*_{QP} is reached, and possibly beyond if the respective optima are then still underrated. For instance, we will see instances in Section 4 where $g^*_{\text{Lin+Tr}} < f^*_{\text{QP}}$ but the final bound, i.e. then one after adding gradient inequalities while keeping $(x,X) \in P_{\text{Lin}} \cap P_{\text{Tr}}$ at the same time, is larger than f^*_{QP}.

3.3 The Hybrid Algorithm

To keep the presentation as general as possible, we will preserve the assumption from the previous subsection that the feasible set P of the linear relaxation just satisfies $P \subseteq P_{\text{Box}}$. Moreover, we assume that we may efficiently minimize $g(x,X)$ over P, while the constraints defining P need not be explicitly given, an efficient separation algorithm is sufficient. Effectively, we thus assume to be given a linear minimization oracle (LMO) for minimizing $g(x,X)$ over P, respectively over $S_\nabla \cap P$ since gradient inequalities can be separated efficiently. Algorithm 1 displays our according hybrid cutting plane procedure.

Algorithm 1 Cutting Plane Algor. for the Linear Relaxation enriched by Gradient Ineq.

Input: A convex underestimator f of γ, an LMO for minimizing $g(x,X)$ over $P \cap S_\nabla$, $P \subseteq P_{\text{Box}}$.
Output: A point $(\bar{x},\bar{X})$ with $g_{\nabla+P} := g(\bar{x},\bar{X}) = \min\{g(x,X) : (x,X) \in P \cap S_\nabla\}$.

1: Set $R = P$.
2: Solve the LP $\min\{g(x,X) : (x,X) \in R\}$ via the LMO. Let $(\bar{x},\bar{X})$ be the solution.
3: **while** $(\bar{x},\bar{X})$ violates its associated gradient inequality regarding f (is underrated) **do**
4: Set $R = R \cap \{\langle Q,X\rangle + c^\top x \geq \bar{x}^\top Q\bar{x} + c^\top \bar{x} + (2Q\bar{x}+c)^\top (x-\bar{x})\}$.
5: Solve the LP $\min\{g(x,X) : (x,X) \in R\}$ via the LMO. Let $(\bar{x},\bar{X})$ be the solution.
6: **return** $(\bar{x},\bar{X})$

Given a convex quadratic function f, Algorithm 1 iteratively solves linear programs as long as the current optimum solution $(\bar{x},\bar{X})$ obtained is underrated by g compared to f. The procedure as such is similar to Kelley's algorithm, though it does not approximate the feasible set by gradient inequalities but employs them to cut off points where the underestimation of γ by g is weaker than the one by f.

Proposition 2. *Let $(\bar{x},\bar{X})$ be the solution returned by Algorithm 1. Then $g(\bar{x},\bar{X}) = \min\{g(x,X) : (x,X) \in P \cap S_\nabla\}$ and thus in particular $g(\bar{x},\bar{X}) \geq f^*_{QP}$.*

Proof. When Algorithm 1 terminates, $(\bar{x},\bar{X})$ has been the solution computed either in line 2 or in line 5 during the final iteration of the inner while loop. Thus $(\bar{x},\bar{X})$ is a minimizer of g on R. Moreover, $(\bar{x},\bar{X})$ did not violate its associated gradient inequality, whence $(\bar{x},\bar{X})$ also minimizes $g(x,X)$ on $R \cap S_\nabla$. Finally, since $f^*_{\text{QP}} = \min\{g(x,X) : (x,X) \in P_{Box} \cap S_\nabla\}$ and the latter polytope is a relaxation of $R \cap S_\nabla$, it follows that $g(\bar{x},\bar{X}) \geq f^*_{\text{QP}}$. □

3.4 Sparsity-Preserving Realization with Triangle Inequalities

Besides the order of computations as emphasized in the previous subsection and the choice of an LMO that gives a strong bound as such, an efficient realization of the LMO is another key aspect of using Algorithm 1 in practice. This remains simple when choosing $P = P_{\text{Box}} \cap P_{\text{Eq}}$ or $P = P_{\text{Lin}}$. In these cases, simplicity is retained in particular because the McCormick inequalities may simply be restricted to the pairs i,j, $i<j$, such that $Q_{ij}\neq 0$, see also below. Also the gradient inequalities act precisely on the set of variables with non-zero objective coefficients, with the only exception of the diagonal X_{ii} variables, and as soon as the constraints $\mathrm{diag}(X)=x$ are present, the gradient inequality can be simply rewritten to

$$\sum_{i<j:Q_{ij}\neq 0} 2Q_{ij}X_{ij}+\sum_i (Q_{ii}-2(Qy)_i)x_i \geq \sum_{i<j:Q_{ij}\neq 0} 2Q_{ij}y_iy_j+\sum_i Q_{ii}y_i^2.$$

As already explained in Section 2.2, sparsity is however not straightforwardly preserved when adding triangle inequalities. The expedient is that solving the linear program $\min\{g(x,X):(x,X)\in P_{\mathrm{Lin}}\cap P_{\mathrm{Tr}}\}$ provides exactly the same lower bound as solving the LP relaxation of the "standard" linearization enriched by odd-cycle inequalities:

$$\begin{array}{lrl}
\min & \displaystyle\sum_{i<j:Q_{ij}\neq 0} 2Q_{ij}X_{ij}+\sum_i (Q_{ii}+c_i)x_i & \\
\text{s.t.} & x(V_1)-x(V_2)-X(C_1)+X(C_2)\leq \left\lfloor \frac{1}{2}|C_1|\right\rfloor & \text{for all Q-cycles } C=C_1\cup C_2, |C_1| \text{ odd}\\
& X_{ij}-x_i\leq 0 & \text{for all } 1\leq i<j\leq n: Q_{ij}\neq 0\\
& X_{ij}-x_j\leq 0 & \text{for all } 1\leq i<j\leq n: Q_{ij}\neq 0\\
& x_i+x_j-X_{ij}\leq 1 & \text{for all } 1\leq i<j\leq n: Q_{ij}\neq 0\\
& X_{ij}\geq 0 & \text{for all } 1\leq i<j\leq n: Q_{ij}\neq 0
\end{array}$$

Here, the odd-cycle inequalities are written in a compact way that is to be interpreted as follows. Q-cycles are ordinary cycles in an undirected graph $G=(V,E)$ having the set of vertices $V=\{1,\dots,n\}$ and the edge set $E=\{\{i,j\}:Q_{ij}\neq 0\}$. In other words, the above relaxation (and thus its extension by gradient inequalities) refers precisely to the variables present according to the non-zero pattern of Q. Given a Q-cycle $C\subseteq E$ such that $C=C_1\cup C_2$ where C_1 has odd cardinality, the set V_1 respectively V_2 refers to the vertices having two C_1- respectively two C_2-adjacent edges, see [20, 21] also for details on how to separate only facet-defining odd-cycle inequalities. In addition, we use the notation $x(V_i):=\sum_{v\in V_i}x_v$ and $X(C_i):=\sum_{\{u,v\}\in C_i}X_{uv}$ for $i\in\{1,2\}$. The number of these inequalities is larger than the number of triangle inequalities in general, and it may even be exponential, but their separation is possible in polynomial time [1]. Finally, the bound obtained is equal to $g^*_{\mathrm{Lin+Tr}}$ because the triangle inequalities are the only irredundant odd-cycle inequalities when adding all non-edges to G, and these non-edges do not contribute to the objective function. Thereby, the irredundancy follows directly from the fact that an odd-cycle inequality is facet-defining for the Boolean Quadric Polytope if and only if the cycle C is chordless [1, 25].

A particularly interesting setting is when $f^*_{\mathrm{QP}}=g^*_{\mathrm{SDP+Eq}}$ holds by construction, and f is used with $P=P_{\mathrm{Lin}}\cap P_{\mathrm{Tr}}$ in Algorithm 1. In this case, we know that the solution $(\bar{x},\bar{X})$ returned satisfies $f^*_{\mathrm{QP}}=g^*_{\mathrm{SDP+Eq}}\leq g(\bar{x},\bar{X})=g^*_{\nabla+\mathrm{Lin+Tr}}\leq g^*_{\mathrm{SDP+Lin+Tr}}$. At the same time, it is generally unclear how well the latter (and comparably strong) bound can be approached by the here presented method. For a special case of practical relevance, we will show in the upcoming Section 3.5 that $g^*_{\nabla+\mathrm{Lin+Tr}}$ cannot exceed $\max\{f^*_{\mathrm{QP}},g^*_{\mathrm{Lin+Tr}}\}$. In general, however, it may well do so as we show experimentally in Section 4.

3.5 A Special Case Concerning the Maximum Cut Problem

In the following, we will show for the particular case where A is a symmetric matrix satisfying $Ae = 0$, and where we have in addition $b = 0$, $Q = A + \mathrm{Diag}(u)$ and $c = b - u$ such that $Q \succeq 0$, that the bound $g^*_{\nabla+\mathrm{Lin}+\mathrm{Tr}}$ given by the relaxation $P_{\mathrm{Lin}} \cap P_{\mathrm{Tr}}$ enriched with gradient inequalities cannot exceed $\max\{g^*_{\mathrm{Lin}+\mathrm{Tr}}, f^*_{\mathrm{QP}}\}$.

The mentioned properties are given in the important special case of transforming a Maximum Cut instance into a QUBO instance. More precisely, let $L_G \in \mathbb{R}^{n\times n}$ be the weighted Laplace matrix of an undirected graph $G = (V,E)$. Then solving the Maximum Cut problem for G is equivalent to solving the QUBO instance with $A = -L_G$ and $b = 0$ [8], so that, in particular, $Ae = 0$ because $L_G e = 0$ [32].

Lemma 1. *Let the symmetric matrix A satisfy $Ae = 0$. Moreover, let $b = 0$, and let f be defined as in Section 3.1 for $Q = A + \mathrm{Diag}(u)$ and $c = b - u$ such that $Q \succeq 0$. Then $\frac{1}{2}e \in \arg\min_{x\in\mathbb{R}^n} f(x)$ and thus $\min_{x\in[0,1]^n} f(x) = -\frac{1}{4}e^\top u$.*

Proof. Since $Ae = 0$, it follows that $Qe = (A + \mathrm{Diag}(u))\, e = u$, hence $\nabla f(\frac{1}{2}e) = 2Q\left(\frac{1}{2}e\right) - u = 0$. Finally, $f\left(\frac{1}{2}e\right) = (\frac{1}{2}e)^\top Q(\frac{1}{2}e) - (\frac{1}{2}e)^\top u = (\frac{1}{2}e)^\top(\frac{1}{2}u) - (\frac{1}{2}e)^\top u = -\frac{1}{4}e^\top u.$ □

We now state and prove the main result of this subsection, which holds in particular when u is derived using the QCR method, i.e., when $f^*_{\mathrm{QP}} = g^*_{\mathrm{SDP+Eq}}$.

Theorem 3. *Let A satisfy $Ae = 0$. Moreover, let $b = 0$, and let f be defined as in Section 3.1 for $Q = A + \mathrm{Diag}(u)$ and $c = b - u$ such that $Q \succeq 0$. Consider the linear relaxation with feasible set $P_{Lin} \cap P_{Tr}$ and g as defined in Section 3.1 with $g^*_{Lin+Tr} := \min\{g(x,X) : (x,X) \in P_{Lin} \cap P_{Tr}\}$ and $g^*_{\nabla+Lin+Tr} := \min\{g(x,X) : (x,X) \in P_{Lin} \cap P_{Tr} \cap S_\nabla\}$. Then $g^*_{\nabla+Lin+Tr} = \max\{g^*_{Lin+Tr}, f^*_{QP}\}$.*

Proof. We start by defining, for a point $(x,X) \in P_{\mathrm{Lin}} \cap P_{\mathrm{Tr}}$, its *complement* point by

$$\widehat{x}_i := 1 - x_i,\ \widehat{X}_{ii} := \widehat{x}_i,\ \text{and}\ \widehat{X}_{ij} := 1 - x_i - x_j + X_{ij}\ \text{for all}\ i,j \in \{1,\dots,n\}, i \neq j. \quad (5)$$

The McCormick and triangle inequalities are invariant under this transformation, so we have $(\widehat{x},\widehat{X}) \in P_{\mathrm{Lin}} \cap P_{\mathrm{Tr}}$. In addition, we have $g(\widehat{x},\widehat{X}) = g(x,X)$ because, by (5), we can write $\widehat{X} = X + ee^\top - ex^\top - xe^\top$, and consequently,

$$\begin{aligned} g(\widehat{x},\widehat{X}) &= \langle Q,\widehat{X}\rangle + c^\top(e - x) \\ &= \langle Q,X\rangle + \langle Q,ee^\top\rangle - \langle Q,ex^\top\rangle - \langle Q,xe^\top\rangle + c^\top e - c^\top x \\ &= \langle Q,X\rangle + e^\top Qe - 2x^\top Qe + c^\top e - c^\top x. \end{aligned}$$

Using $Qe = u = -c$, we obtain

$$e^\top Qe + c^\top e = 0 \text{ and } -2x^\top Qe - c^\top x = 2x^\top c - c^\top x = c^\top x,$$

and thus finally

$$g(\widehat{x},\widehat{X}) = \langle Q,X\rangle + c^\top x = g(x,X).$$

Let now (x^*,X^*) be an optimal solution w.r.t. to $P_{\mathrm{Lin}} \cap P_{\mathrm{Tr}}$, i.e. $g(x^*,X^*) = g^*_{\mathrm{Lin+Tr}}$. The midpoint $(\bar{x},\bar{X})$ of (x^*,X^*) and its (also optimal) complement $(\widehat{x}^*,\widehat{X}^*)$ has

$$\bar{x} := \frac{1}{2}(x^* + \widehat{x}^*) = \frac{1}{2}e \text{ and } \bar{X} := \frac{1}{2}(X^* + \widehat{X}^*),$$

and is feasible for $P_{\text{Lin}} \cap P_{\text{Tr}}$ since the latter is a convex set. Moreover, $g(\bar{x},\bar{X}) = g^*_{\text{Lin+Tr}}$ because g is linear. It follows that there is an $\bar{X}$ such that $\left(\frac{1}{2}e,\bar{X}\right) \in P_{\text{Lin}} \cap P_{\text{Tr}}$ and

$$g\left(\frac{1}{2}e,\bar{X}\right) = g^*_{\text{Lin+Tr}}. \tag{6}$$

Consider further the point (x^0,X^0) with $x^0 = \frac{1}{2}e$, and $X^0 = \frac{1}{2}ee^\mathsf{T}$. It satisfies the McCormick and triangle inequalities as well, i.e., $(x^0,X^0) \in P_{\text{Lin}} \cap P_{\text{Tr}}$, and moreover, with $Qe = u$, we have

$$g(x^0,X^0) = \left\langle Q, \frac{1}{2}ee^\mathsf{T} \right\rangle - \frac{1}{2}e^\mathsf{T}u = \frac{1}{2}e^\mathsf{T}Qe - \frac{1}{2}e^\mathsf{T}u = 0. \tag{7}$$

We now use the defined points in order to prove the main statement, thereby distinguishing the two possible cases regarding the proportionalities of $g^*_{\text{Lin+Tr}}$ and f^*_{QP}.

Case $g^*_{\text{Lin+Tr}} \geq f^*_{\text{QP}}$. By Lemma 1, we know that $f(x^0) = f^*_{\text{QP}}$ and by (6), we have under this assumption that

$$g\left(\frac{1}{2}e,\bar{X}\right) = g^*_{\text{Lin+Tr}} \geq f^*_{\text{QP}} = f\left(\frac{1}{2}e\right).$$

Hence, $\left(\frac{1}{2}e,\bar{X}\right)$ satisfies all gradient cuts (it is not underrated), and thus $g^*_{\nabla+\text{Lin+Tr}} \leq g^*_{\text{Lin+Tr}}$. On the other hand, we have $g^*_{\nabla+\text{Lin+Tr}} \geq g^*_{\text{Lin+Tr}}$ because adding valid inequalities can only restrict the feasible region. Therefore, the statement of the theorem is shown for the case $g^*_{\text{Lin+Tr}} \geq f^*_{\text{QP}}$.

Case $g^*_{\text{Lin+Tr}} < f^*_{\text{QP}}$. This assumption and the fact that $f(0) = 0$ imply that $g^*_{\text{Lin+Tr}} < f^*_{\text{QP}} \leq 0$.

For $\lambda \in [0,1]$, define $X(\lambda) := (1-\lambda)\bar{X} + \lambda X^0$, such that $\left(x^0, X(\lambda)\right) \in P_{\text{Lin}} \cap P_{\text{Tr}}$ by convexity of $P_{\text{Lin}} \cap P_{\text{Tr}}$. Moreover, given the objective values of $(\frac{1}{2}e,\bar{X})$ and (x^0,X^0) as of (6) respectively (7), and since g is linear, $g(x^0,X(\lambda)) = (1-\lambda)g^*_{\text{Lin+Tr}}$. Since $g^*_{\text{Lin+Tr}} < f^*_{\text{QP}} \leq 0$, there exists

$$\lambda^* = 1 - \frac{f^*_{\text{QP}}}{g^*_{\text{Lin+Tr}}} \in [0,1)$$

such that

$$g(x^0,X(\lambda^*)) = f^*_{\text{QP}} = f(x^0).$$

Hence, this point satisfies all gradient cuts (is not underrated) as well. It follows that $g^*_{\nabla+\text{Lin+Tr}} \leq f^*_{\text{QP}}$ while we have $g^*_{\nabla+\text{Lin+Tr}} \geq f^*_{\text{QP}}$ from Proposition 2, and so the statement of the theorem is shown for the case $g^*_{\text{Lin+Tr}} < f^*_{\text{QP}}$. □

4 Computational Experiments

To exemplify the practical application of Algorithm 1, we construct precisely the setting of special interest addressed at the end of Section 3.4. To this end, we derive the convex underestimator f using the QCR method such that $f^*_{\text{QP}} = g^*_{\text{SDP+Eq}}$ holds. Further, we choose $P = P_{\text{Lin}} \cap P_{\text{Tr}}$, so that the obtained lower bound is $g^*_{\nabla+\text{Lin+Tr}}$ and satisfies $\max\{g^*_{\text{Lin+Tr}}, g^*_{\text{SDP+Eq}}\} \leq g^*_{\nabla+\text{Lin+Tr}} \leq g^*_{\text{SDP+Lin+Tr}}$. One the one hand, this choice allows us to provide an impression on the quality of the lower bounds that can be obtained with this approach for some established QUBO benchmark instances. On the other, it allows us to demonstrate practical results for native Maximum Cut instances that align with Theorem 3 from Section 3.5.

To this end, we ran Algorithm 1 on established instance sets, so as to cover cases where $g^*_{\text{Lin+Tr}} < g^*_{\text{SDP+Eq}}$ as well as where $g^*_{\text{SDP+Eq}} < g^*_{\text{Lin+Tr}}$. We then selected the following results in terms of the possible scenarios discussed in Section 3.2.

Case 1: $g^*_{\textbf{SDP+Eq}} < g^*_{\textbf{Lin+Tr}} < g^*_{\nabla+\textbf{Lin+Tr}} < g^*_{\textbf{SDP+Lin+Tr}}$ An example QUBO instance class whose members satisfy $g^*_{\text{SDP+Eq}} < g^*_{\text{Lin+Tr}}$ is the set **gka-b** by Glover, Kochenberger and Alidaee [13] which has matrix dimensions between 20 and 125, density 1, diagonal coefficients from the interval $[-63, 0]$, and off-diagonal coefficients from $[0, 100]$. For these instances, we end up with a significantly better lower bound $g^*_{\nabla+\text{Lin+Tr}} > g^*_{\text{Lin+Tr}}$. The bound is sometimes closer to $g^*_{\text{Lin+Tr}}$ and sometimes closer to $g^*_{\text{SDP+Lin+Tr}}$ which is optimal for these instances, see Table 1. So as it turns out, for all of these instances, all minimizers of g in $P_{\text{Lin}} \cap P_{\text{Tr}}$ are underrated compared to their value in terms of the here chosen underestimator, such that the gradient inequalities lead to a strict improvement of the lower bound. For these instances, we further observed that already using $P_{\text{Box}} \cap P_{\text{Eq}}$ as the linear relaxation provides an improvement over $g^*_{\text{SDP+Eq}}$ when combined with gradient inequalities.

Instance	OPT	$g^*_{\text{Lin+Tr}}$	$g^*_{\text{SDP+Eq}}$	$g^*_{\nabla+\text{Lin+Tr}}$	$g^*_{\text{SDP+Lin+Tr}}$
gka1b	-133	-260.67	-362.92	-195.81	-133.00
gka2b	-121	-345.67	-505.65	-242.51	-121.00
gka3b	-118	-416.33	-717.95	-266.04	-118.00
gka4b	-129	-548.33	-809.87	-310.11	-129.00
gka5b	-150	-632.67	-1034.79	-362.91	-150.00
gka6b	-146	-758.00	-1279.01	-389.96	-146.00
gka7b	-160	-840.00	-1362.49	-404.64	-160.00
gka8b	-145	-957.00	-1479.12	-443.87	-145.00
gka9b	-137	-1104.33	-1663.61	-472.76	-137.00
gka10b	-154	-1310.00	-2073.14	-535.66	-154.00

Table 1: Approximate results for gka-b instances exemplifying case 1.

Case 2: $g^*_{\textbf{Lin+Tr}} < g^*_{\textbf{SDP+Eq}} < g^*_{\nabla+\textbf{Lin+Tr}} < g^*_{\textbf{SDP+Lin+Tr}}$ The instance set **be150_3** with $n = 150$ and density 0.3 is by Billionnet and Elloumi [2], and here A takes integer values from $[-100, 100]$ on the diagonal and from $[-50, 50]$ on the off-diagonal entries. For each of these instances, Table 2 shows that $g^*_{\text{Lin+Tr}} < g^*_{\text{SDP+Eq}}$ except for be150.3.4, and that $g^*_{\nabla+\text{Lin+Tr}}$ gives a slight further improvement. Therefore, the minimizers in $P_{\text{Box}} \cap S_{\nabla}$ having objective value $f^*_{\text{QP}} = g^*_{\text{SDP+Eq}}$ did not satisfy the triangle inequalities, and the minimizers that do so and are not underrated by g compared to f have a strictly larger objective value.

Instance	OPT	$g^*_{\text{Lin+Tr}}$	$g^*_{\text{SDP+Eq}}$	$g^*_{\nabla+\text{Lin+Tr}}$	$g^*_{\text{SDP+Lin+Tr}}$
be150.3.1	-18889	-20362.22	-20175.17	-20014.39	-18999.84
be150.3.2	-17816	-19543.94	-19203.02	-19085.42	-17944.63
be150.3.3	-17314	-18585.56	-18416.48	-18226.70	-17333.54
be150.3.4	-19884	-20947.20	-21050.34	-20785.50	-19935.03
be150.3.5	-16817	-18655.60	-18210.59	-18095.31	-16972.26
be150.3.6	-16780	-19092.64	-18350.19	-18301.82	-17046.26
be150.3.7	-18001	-19790.60	-19502.58	-19355.80	-18138.41
be150.3.8	-18303	-20413.86	-19956.39	-19866.95	-18613.07
be150.3.9	-12838	-16024.02	-14559.92	-14556.52	-13193.52
be150.3.10	-17963	-19802.08	-19529.05	-19395.41	-18260.58

Table 2: Approximate results for be150.3 instances exemplifying case 2.

Case 3: $g^*_{\textbf{SDP+Eq}} < g^*_{\textbf{Lin+Tr}} = g^*_{\nabla+\textbf{Lin+Tr}} < g^*_{\textbf{SDP+Lin+Tr}}$ For the (Erdös-Renyi-Gilbert) random graphs **pm1s_100** from the Biq Mac Library [3] with $|V| = 100$, edge weights ± 1 and density 0.1, we set $A = -L(G)$ where $L(G)$ is the Laplace matrix of the respective graph, and $b = 0$, thus effectively solving a Maximum Cut problem. Here, we observe that $g^*_{\text{SDP+Eq}} < g^*_{\text{Lin+Tr}}$ and, in line with Theorem 3 from Section 3.5, that no improvement of $g^*_{\text{Lin+Tr}}$ via gradient inequalities is possible, see Table 3.

Instance	OPT	$g^*_{\text{Lin+Tr}}$	$g^*_{\text{SDP+Eq}}$	$g^*_{\nabla+\text{Lin+Tr}}$	$g^*_{\text{SDP+Lin+Tr}}$
pm1s_80.0	-79	-79.00	-90.29	-79.00	-79.00
pm1s_80.1	-85	-86.22	-96.18	-86.22	-85.00
pm1s_80.2	-82	-87.80	-94.02	-87.80	-84.04
pm1s_80.3	-81	-83.20	-92.14	-83.20	-81.30
pm1s_80.4	-70	-71.61	-82.06	-71.61	-70.00
pm1s_80.5	-87	-88.96	-98.69	-88.96	-87.05
pm1s_80.6	-73	-76.65	-85.69	-76.65	-74.04
pm1s_80.7	-83	-85.46	-95.45	-85.46	-83.22
pm1s_80.8	-81	-83.74	-95.47	-83.74	-81.68
pm1s_80.9	-70	-72.02	-82.00	-72.02	-70.08

Table 3: Approximate results for pm1s instances exemplifying case 3.

Case 4: $g^*_{\textbf{Lin+Tr}} < g^*_{\textbf{SDP+Eq}} = g^*_{\nabla+\textbf{Lin+Tr}} < g^*_{\textbf{SDP+Lin+Tr}}$ Another class of (Erdös-Renyi-Gilbert) random graphs from the Biq Mac Library is **g05_60** with unit edge weights, $|V| = 60$ and density 0.5. For these graphs, we proceed as in Case 3 and the roles of $g^*_{\text{Lin+Tr}} < g^*_{\text{SDP+Eq}}$ switch, see Table 4. As expected from Theorem 3 still no further improvement via gradient inequalities is possible.

5 Conclusion

We have designed a hybrid algorithmic approach to systematically investigate the potential improvement of the lower bound obtained by a linear QUBO relaxation when enriching it with gradient inequalities derived from an accompanying quadratic convex underestimator. Naturally, this potential heavily depends on the choice of both the linear relaxation as well as the convex underestimator. Here, we focused the discussion on

Instance	OPT	$g^*_{\text{Lin+Tr}}$	$g^*_{\text{SDP+Eq}}$	$g^*_{\nabla\text{+Lin+Tr}}$	$g^*_{\text{SDP+Lin+Tr}}$
g05_60.0	-536	-590.00	-550.05	-550.05	-537.24
g05_60.1	-532	-590.00	-543.11	-543.11	-532.63
g05_60.2	-529	-590.00	-543.18	-543.18	-531.53
g05_60.3	-538	-590.00	-548.65	-548.65	-538.00
g05_60.4	-527	-590.00	-541.38	-541.38	-530.70
g05_60.5	-533	-590.00	-542.59	-542.59	-533.25
g05_60.6	-531	-590.00	-544.72	-544.72	-533.55
g05_60.7	-535	-590.00	-550.42	-550.42	-536.75
g05_60.8	-530	-590.00	-543.98	-543.98	-532.66
g05_60.9	-533	-590.00	-549.89	-549.89	-536.07

Table 4: Approximate results for g05 instances exemplifying case 4.

the customary case where the underestimator is derived by the QCR method and where the linear relaxation is a "standard" linearization strengthened with triangle inequalities. For this particular setting, we found that our hybrid approach provides lower bounds which can be significantly stronger than the according linear and quadratic programming bounds standalone. Our results also show that the actual improvement further depends strongly on the instance class considered, and may also be small or even provably impossible. In particular, for the relevant subset of QUBO instances that result from transforming an instance of the Maximum Cut problem, we have shown that no improvement is possible in this particular setting.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foun- dation) – 537033028.

References

[1] F. Barahona and A. R. Mahjoub. On the cut polytope. *Mathematical Programming*, 36(2):157–173, 1986.

[2] A. Billionnet and S. Elloumi. Using a mixed integer quadratic programming solver for the unconstrained quadratic 0-1 problem. *Mathematical Programming*, 109(1):55–68, Jan 2007.

[3] Biq Mac Library – a collection of Max-Cut and quadratic 0-1 programming instances of medium size. `http://biqmac.aau.at/biqmaclib.html`, 2009.

[4] G. Blekherman, S. S. Dey, M. Molinaro, and S. Sun. Sparse PSD approximation of the PSD cone. *Mathematical Programming*, 191(2):981–1004, Feb 2022.

[5] D. Brosch, J. Schwiddessen, and A. Wiegele. The augmented mixing method: computing high-accuracy primal-dual solutions to large-scale SDPs via column updates. *Optimization Methods and Software*, 41(3):754–796, May 2026.

[6] J. Charfreitag, M. Jünger, S. Mallach, and P. Mutzel. McSparse: Exact solutions of sparse maximum cut and sparse unconstrained binary quadratic optimization problems. In C. A. Phillips and B. Speckmann, editors, *2022 Proceedings of the Symposium on Algorithm Engineering and Experiments (ALENEX)*, pages 54–66, 2022.

[7] D. de Roux, R. Carr, and R. Ravi. Instance-specific linear relaxations of semidefinite optimization problems. *Mathematical Programming Computation*, 17(2):385–435, Jun 2025.

[8] C. De Simone. The cut polytope and the boolean quadric polytope. *Discrete Mathematics*, 79(1):71–75, 1990.

[9] S. S. Dey, A. Kazachkov, A. Lodi, and G. Muñoz. Cutting plane generation through sparse principal component analysis. *SIAM Journal on Optimization*, 32(2):1319–1343, 2022.

[10] R. Fortet. L'algèbre de boole et ses applications en recherche opérationnelle. *Cahiers du Centre d'Etudes de Recherche Opérationnelle*, 1:5–36, 1959.

[11] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1 edition, 1979.

[12] F. Glover, G. Kochenberger, and Y. Du. *Applications and Computational Advances for Solving the QUBO Model*, pages 39–56. Springer International Publishing, Cham, 2022.

[13] F. Glover, G. A. Kochenberger, and B. Alidaee. Adaptive memory tabu search for binary quadratic programs. *Management Science*, 44(3):336–345, 1998.

[14] F. Glover and E. Woolsey. Converting the 0–1 polynomial programming problem to a 0–1 linear program. *Oper. Res.*, 22(1):180–182, Feb. 1974.

[15] N. Gusmeroli, T. Hrga, B. Lužar, J. Povh, M. Siebenhofer, and A. Wiegele. BiqBin: A parallel branch-and-bound solver for binary quadratic problems with linear constraints. *ACM Trans. Math. Softw.*, 48(2), jul 2022.

[16] O. Günlük, P. Jünger, J. Linderoth, A. Lodi, and J. Luedtke. Sparse cuts for the positive semidefinite cone. arXiv Preprint 2603.09864, 2026.

[17] P. L. Hammer. Some network flow problems solved with pseudo-boolean programming. *Operations Research*, 13(3):388–399, 1965.

[18] P. L. Hammer, P. Hansen, and B. Simeone. Roof duality, complementation and persistency in quadratic 0–1 optimization. *Mathematical Programming*, 28(2):121–155, Feb 1984.

[19] Q. Han, C. Li, Z. Lin, C. Chen, Q. Deng, D. Ge, H. Liu, and Y. Ye. A low-rank ADMM splitting approach for semidefinite programming. *INFORMS Journal on Computing*, 2025.

[20] M. Jünger and S. Mallach. Odd-cycle separation for maximum cut and binary quadratic optimization. In M. A. Bender, O. Svensson, and G. Herman, editors, *27th Annual European Symposium on Algorithms (ESA 2019)*, volume 144 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 63:1–63:13, Dagstuhl, Germany, 2019.

[21] M. Jünger and S. Mallach. Exact facetial odd-cycle separation for maximum cut and binary quadratic optimization. *INFORMS Journal on Computing*, 33(4):1419–1430, 2021.

[22] J. E. Kelley. The cutting-plane method for solving convex programs. *Journal of the Society for Industrial and Applied Mathematics*, 8(4):703–712, 1960.

[23] N. Krislock, J. Malick, and F. Roupin. BiqCrunch: A semidefinite branch-and-bound method for solving binary quadratic problems. *ACM Trans. Math. Softw.*, 43(4), 3 2017.

[24] G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I — convex underestimating problems. *Mathematical Programming*, 10(1):147–175, Dec 1976.

[25] M. Padberg. The boolean quadric polytope: Some characteristics, facets and relatives. *Mathematical Programming*, 45(1):139–172, Aug 1989.

[26] A. P. Punnen, editor. *The Quadratic Unconstrained Binary Optimization Problem*. Springer, 2022.

[27] A. Qualizza, P. Belotti, and F. Margot. Linear programming relaxations of quadratically constrained quadratic programs. In J. Lee and S. Leyffer, editors, *Mixed Integer Nonlinear Programming*, pages 407–426, New York, NY, 2012. Springer New York.

[28] D. Rehfeldt, T. Koch, and Y. Shinano. Faster exact solution of sparse maxcut and qubo problems. *Mathematical Programming Computation*, Apr 2023.

[29] F. Rendl. *Semidefinite Relaxations for Integer Programming*, pages 687–726. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[30] F. Rendl, G. Rinaldi, and A. Wiegele. Solving Max-Cut to optimality by intersecting semidefinite and polyhedral relaxations. *Mathematical Programming*, 121(2):307–335, 2010.

[31] F. Serrano, R. Schwarz, and A. Gleixner. On the relation between the extended supporting hyperplane algorithm and Kelley's cutting plane algorithm. *Journal of Global Optimization*, 78(1):161–179, Sep 2020.

[32] D. Spielman. *Spectral Graph Theory*, chapter 16. 2019.