Polynomial Optimization via Penalized Conic Relaxation

Ramtin Madani, Mohsen Kheirandishfard, Javad Lavaei, and Alper Atamtürk

Abstract—This paper revisits conic programming relaxations for the class of quadratically-constrained quadratic programs (QCQPs). We present penalty terms, whose incorporation into the objective of convex relaxations enables the retrieval of feasible and near-optimal solutions for non-convex QCQPs. We introduce generalized linear independence constraint qualification (GLICQ) and prove that any GLICQ regular initial point that is sufficiently close to the original QCQP feasible set can be used to construct an appropriate penalty term. As a consequence, a sequential conic optimization method is developed that preserves feasibility and aims to improve the solution at every round. Numerical experiments on benchmark examples from the library of quadratic programming (QPLIB) instances demonstrate the ability of the proposed framework in finding feasible and near-globally optimal points.

I. INTRODUCTION

Polynomial optimization is the problem of minimizing a polynomial function within a feasible set that is characterized by polynomial functions. Physics laws and characteristics of dynamical systems are widely modeled using polynomials. As a result, polynomial optimization arises in various scientific and engineering applications, such as electric power systems [1]–[4], imaging science [5]–[8], signal processing [9]–[14], automatic control [4], [15]–[17], quantum mechanics [18]–[21], and cybersecurity [22]–[25]. The development of efficient optimization techniques and numerical algorithms for finding global minima of polynomial optimization problems has been an active area of research for decades. Due to the barriers imposed by NP-hardness, the focus of some research efforts has shifted from designing general-purpose algorithms to specialized methods that are robust and scalable for specific application domains. Notable examples for which methods with guaranteed performance have been offered in the literature include the problems of multisensor beamforming in communication theory [26], phase retrieval in signal processing [27], and matrix completion in machine learning [28], [29].

This paper advances a popular framework for the global optimization of polynomial optimization, which involves convex hull characterization by forming hierarchies of semidefinite programming (SDP) relaxations [30]–[36]. SDP relaxation enlarges a possibly non-convex feasible set into a convex set characterizable via convex functions, and then provide the exact or a lower bound on the optimal objective value associated with a global solution. The SDP relaxation technique provides a lower bound on the minimum cost of the original problem, which can be used for various purposes such as the branch-and-bound algorithm [37]–[39]. To understand the quality of the SDP relaxation, its optimal objective value is shown to be at most 14% different from the globally optimal cost for the MAXCUT problem [40]. The maximum possible gap between the solution of a graph optimization and its SDP relaxation is defined as the Grothendieck constant of the graph [41], [42]. This constant has been derived for some special cases in [43]. The paper [44] shows how a complex SDP relaxation may solve the max-3-cut problem. This approach has been generalized in several papers [45]–[52]. If the SDP relaxation provides the same optimal objective value as the original problem, the relaxation is said to be exact. The exactness of the SDP relaxation has been verified for a variety of problems [53]–[57]. One of the primary challenges for the application of SDP hierarchies beyond small-scale instances is the rapid growth of dimensionality. In response, some studies have exploited sparsity and structural patterns to boost efficiency [54], [58]–[61]. Another direction is pursued in [17], [62]–[66], which is based on lower-complexity relaxations as alternatives to computationally demanding semidefinite programming relaxations.

A. Contributions

This paper is concerned with non-convex QCQP for which standard convex relaxations are inexact and fail to produce feasible solutions. We incorporate linear penalty terms into the objective of conic relaxations and show that feasible and near-globally optimal points can be obtained for the original QCQP by solving the resulting penalized relaxation problem. Each penalty term is based on an arbitrary initial point for the original QCQP. Our first result states that if the initial point is feasible and satisfies the linear independence constraint qualification (LICQ) condition, then the penalized relaxation has a unique solution that is feasible for the original QCQP and its objective value is not worst than that of the initial point. As shown in Section III, this result can be readily applied to general polynomial optimization problems. Our second result states that if the initial point is not feasible but instead is sufficiently close to the feasible set of QCQP and satisfies a generalized LICQ condition, then the penalized relaxation produces feasible solutions for the QCQP. Lastly, we propose a sequential procedure for general QCQPs and demonstrate its performance on benchmark examples from the QPLIB library. Remarkably, the proposed sequential approach is able to find feasible and near-globally optimal points even for those instances where convex relaxations are unbounded. This framework has been investigated in [67], [68] for convex relaxation of bilinear matrix inequalities, as well as in papers [69], [70], which demonstrate the ability of sequential penal-
ized relaxation in solving computationally-hard power system optimization problems.

B. Notations

Throughout the paper, scalars, vectors, and matrices are respectively shown by italic letters, lower-case italic bold letters, and upper-case italic bold letters. The symbols $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{R}^{n \times m}$ denote the sets of real scalars, real vectors of size $n$, and real matrices of size $n \times m$, respectively. The set of $n \times n$ real symmetric matrices is shown by $\mathbb{S}_n$. For a given vector $a$ and a matrix $A$, the symbols $a_i$ and $A_{ij}$ respectively indicate the $i$th element of $a$ and the $(i,j)$th element of $A$. The symbols $\langle \cdot, \cdot \rangle$ and $\| \cdot \|_F$ denote the Frobenius inner product and norm of matrices, respectively. The notation $\| \cdot \|$ represents either the absolute value operator or cardinality of a set, depending on the context. The notation $\| \cdot \|_2$ denotes the $\ell_2$ norm of vectors, matrices, and matrix pencils. The $n \times n$ identity matrix is denoted by $I_n$. The origin of $\mathbb{R}^n$ is denoted by $0_n$. The superscript $(\cdot)^\top$ and the symbol $\text{tr}\{ \cdot \}$ represent the transpose and trace operators, respectively. Given a matrix $A \in \mathbb{R}^{m \times n}$, the notation $\sigma_{\min}(A)$ represents the minimum singular value of $A$. The notation $\sigma(A) \geq 0$ means that $A$ is symmetric positive-semidefinite. For every pair of $n \times n$ symmetric matrices $(A, B)$ and every proper cone $C \subseteq \mathbb{S}_n$, the notation $A \succeq_C B$ means that $A - B \succeq 0$ is the cone of $n \times n$ symmetric matrices whose $r \times r$ principal submatrices are all positive semidefinite. Similarly, define $C_r^+$ as the dual cone of $C_r$, i.e., the cone of $n \times n$ symmetric matrices with factor-width bounded by $r$. Given a matrix $A \in \mathbb{R}^{m \times n}$ and two sets of positive integers $S_1$ and $S_2$, define $A\{S_1, S_2\}$ as the submatrix of $A$ obtained by removing all rows of $A$ whose indices do not belong to $S_1$, and all columns of $A$ whose indices do not belong to $S_2$. Moreover, define $A\{S_1\}$ as the submatrix of $A$ obtained by removing all rows of $A$ that do not belong to $S_1$. Given a vector $a \in \mathbb{R}^n$ and a set $F \subseteq \mathbb{R}^n$, define $d_F(a)$ as the minimum distance between $a$ and members of $F$. Given a pair of integers $(n, r)$, the binomial coefficient “$n$ choose $r$” is denoted by $C^n_r$. The notations $\nabla_a f(a)$ and $\nabla^2_a f(a)$, respectively, represent the gradient and Hessian of the function $f$, with respect to the vector $a$, at a point $a$.

II. Problem Formulation

Consider a general quadratically-constrained quadratic program (QCQP):

\begin{align}
\text{minimize } & \quad q_0(x) \\
\text{subject to } & \quad q_k(x) \leq 0 \quad k \in I, \quad (1a) \\
& \quad q_k(x) = 0 \quad k \in E, \quad (1c)
\end{align}

where $I$ and $E$ index the sets of inequality and equality constraints, respectively. For every $k \in \{0\} \cup I \cup E$, $q_k : \mathbb{R}^n \to \mathbb{R}$ is a quadratic function of the form $q_k(x) \equiv x^\top A_k x + b_k^\top x + c_k$, where $A_k \in \mathbb{S}_n$, $b_k \in \mathbb{R}^n$, and $c_k \in \mathbb{R}$. Denote $F$ as the feasible set of the QCQP $(1a) - (1c)$. To derive the optimality conditions for a given point, it is useful to define the Jacobian matrix of the constraint functions.

Definition 1 (Jacobian Matrix). For every $\hat{x} \in \mathbb{R}^n$, the Jacobian matrix $J(\hat{x})$ for the constraint functions $\{q_k\}_{k \in I \cup E}$ is

\begin{align}
J(\hat{x}) & \triangleq [\nabla_a q_1(\hat{x}), \ldots, \nabla_a q_{|I \cup E|}(\hat{x})]^\top. \quad (2a)
\end{align}

For every $Q \subseteq I \cup E$, define $J_Q(\hat{x})$ as the submatrix of $J(\hat{x})$ resulting from the rows that belong to $Q$.

Given a feasible point for the QCQP $(1a) - (1c)$, the well-known linear independence constraint qualification (LICQ) condition can be used as a regularity criterion.

Definition 2 (LICQ Condition). A feasible point $\hat{x} \in F$ is LICQ regular if the rows of $J_{\hat{x}}(\hat{x})$ are linearly independent, i.e., the cone of $n \times n$ symmetric matrices with factor-width bounded by $r$. Given a matrix $A \in \mathbb{R}^{m \times n}$ and two sets of positive integers $S_1$ and $S_2$, define $A\{S_1, S_2\}$ as the submatrix of $A$ obtained by removing all rows of $A$ whose indices do not belong to $S_1$, and all columns of $A$ whose indices do not belong to $S_2$. Moreover, define $A\{S_1\}$ as the submatrix of $A$ obtained by removing all rows of $A$ that do not belong to $S_1$. Given a vector $a \in \mathbb{R}^n$ and a set $F \subseteq \mathbb{R}^n$, define $d_F(a)$ as the minimum distance between $a$ and members of $F$. Given a pair of integers $(n, r)$, the binomial coefficient “$n$ choose $r$” is denoted by $C^n_r$. The notations $\nabla_a f(a)$ and $\nabla^2_a f(a)$, respectively, represent the gradient and Hessian of the function $f$, with respect to the vector $a$, at a point $a$.

The feasibility distance $d_F(\hat{x}) \equiv \min\{\|x - \hat{x}\|_2 \mid x \in F\}$. (3)

Definition 3 (Feasibility Distance). The feasibility distance function $d_F : \mathbb{R}^n \to \mathbb{R}$ is defined as

\begin{align}
B & \triangleq \mathcal{E} \cup \left\{k \in I \mid q_k(\hat{x}) + \|\nabla q_k(\hat{x})\|_2d_F(\hat{x}) \geq 0 \right\}. \quad (4)
\end{align}

The point $\hat{x}$ is said to satisfy the GLICQ condition if the rows of $J_{\hat{x}}(\hat{x})$ are linearly independent. Moreover, the singularity function $s : \mathbb{R}^n \to \mathbb{R}$ is defined as

\begin{align}
s(\hat{x}) & \triangleq \begin{cases} \sigma_{\min}(J(\hat{x})) & s(\hat{x}) \text{ satisfies GLICQ} \\ 0 & \text{otherwise} \end{cases}. \quad (5)
\end{align}

where $\sigma_{\min}(J(\hat{x}))$ denotes the smallest singular value of $J(\hat{x})$.

Observe that if $\hat{x}$ is feasible, then $d_F(\hat{x}) = 0$, and LICQ and GLICQ conditions are equivalent. Moreover, GLICQ is satisfied if and only if $s(\hat{x}) > 0$. The next definition introduces the notion of matrix pencil corresponding to the QCQP $(1a) - (1c)$, which will be used later as a sensitivity measure.

Definition 5 (Pencil Norm). For the QCQP $(1a) - (1c)$, define the corresponding matrix pencil $P : \mathbb{R}^{|I|} \times \mathbb{R}^{|E|} \to \mathbb{S}_n$ as follows:

\begin{align}
P(\gamma, \mu) & \triangleq \sum_{k \in I} \gamma_k A_k + \sum_{k \in E} \mu_k A_k. \quad (6)
\end{align}
Moreover, define the pencil norm $\|P\|_2$ as

$$\|P\|_2 \triangleq \max \left\{ \|P(\gamma, \mu)\|_2 \mid \|\gamma\|_2 + \|\mu\|_2 = 1 \right\}, \tag{7}$$

which is upperbounded by $\sqrt{\sum_{k \in \mathcal{U} \cup \mathcal{E}} \|A_k\|_2^2}$.

A. Lifting and Reformulation Linearization

A common practice for tackling the non-convex QCQP (1a)-(1c) involves introducing an auxiliary variable $X \in \mathbb{S}_n$ accounting for $xx^T$. Using $X$, the objective function (1a) and constraints (1b)-(1c) can be recast in a linear way. For every $k \in \{0\} \cup \mathcal{U} \cup \mathcal{E}$, define $\bar{q}_k : \mathbb{R}^n \times \mathbb{S}_n \to \mathbb{R}$ as

$$\bar{q}_k(x, X) \triangleq \langle A_k, X \rangle + 2b_k^T x + c_k. \tag{8}$$

Moreover, in the presence of affine constraints, the well-known reformulation-linearization technique (RLT) or Sherali and Adams [71] can be used to produce additional inequalities with respect to $x$ and $X$ to strengthen convex relaxations. Define $\mathcal{L}$ as the set of affine constrains in the QCQP (1a)-(1c), i.e., $\mathcal{L} \triangleq \{k \in \mathcal{U} \cup \mathcal{E} \mid A_k = 0_n \}$. Define also

$$H \triangleq [B(L \cap I)^T, B(L \cap \mathcal{E})^T, -B(L \cap \mathcal{E})^T]^T, \quad h \triangleq [c(L \cap I)^T, c(L \cap \mathcal{E})^T, -c(L \cap \mathcal{E})^T]^T, \tag{9a}$$

where $B \triangleq [b_1, \ldots, b_{|L \cap \mathcal{E}|}]^T$ and $c \triangleq [c_1, \ldots, c_{|L \cap \mathcal{E}|}]^T$.

Every $x \in \mathcal{F}$ satisfies

$$Hx + h \leq 0, \tag{10}$$

and, as a result, all elements of the matrix

$$Hxx^T H^T + hx^T H^T + Hxh^T + hh^T \tag{11}$$

are positive if $x$ is feasible. Hence, the inequality

$$e_j^T V(x, xx^T) e_j \geq 0 \tag{12}$$

holds true for every $x \in \mathcal{F}$ and $(i, j) \in \mathcal{H} \times \mathcal{H}$, where $V : \mathbb{R}^n \times \mathbb{S}_n \to \mathbb{R}^{|\mathcal{H}|}$ is defined as

$$V(x, X) \triangleq HXH^T + hx^T H^T + Hxh^T + hh^T, \tag{13}$$

$\mathcal{H} \triangleq \{1, 2, \ldots, |L \cap I| + 2|L \cap \mathcal{E}|\}$, and $e_1, \ldots, e_{|\mathcal{H}|}$ denote the standard bases in $\mathbb{R}^{|\mathcal{H}|}$.

B. Convex Relaxation

Consider the following reformulation of QCQP (1a)-(1c):

$$\begin{align*}
\text{minimize}_{x \in \mathbb{R}^n} & \quad \bar{q}_0(x, X) \tag{14a} \\
\text{subject to} & \quad \bar{q}_k(x, X) \leq 0, \quad k \in \mathcal{I} \tag{14b} \\
& \quad \bar{q}_k(x, X) = 0, \quad k \in \mathcal{E} \tag{14c} \\
& \quad X - xx^T \succeq C_r, \quad 0 \tag{14d} \\
& \quad e_j^T V(x, X) e_j \geq 0, \quad (i, j) \in \mathcal{V} \tag{14e}
\end{align*}$$

where $\mathcal{V} \subseteq \mathcal{H} \times \mathcal{H}$ is a selection of RLT inequalities, the additional conic constraint (14d) is a convex relaxation of the equation $X = xx^T$, and

$$C_r \triangleq \{ Y \mid Y \{K, K\} \succeq 0, \quad \forall K \subseteq \{1, \ldots, n\} \land |K| = r \}. \tag{15}$$

If $\mathcal{V} \neq \emptyset$, we refer to the convex problem (14a)-(14e) as the $r$-order conic programming relaxation of the QCQP (1a)-(1c) with RLT inequalities from $\mathcal{V}$, and yield the well-known semidefinite programming (SDP) and second-order conic programming (SOCP) relaxations, respectively.

If the relaxed problem (14a)-(14e) has an optimal solution $(\hat{x}, \hat{X})$ that satisfies $X = \hat{x}\hat{x}^T$, then the relaxation is said to be exact and $\hat{x}$ is a globally optimal solution for the QCQP (1a)-(1c).

The next section offers a penalization method for addressing the case where the relaxation is not exact.

C. Penalization

If the conic relaxation problem (14a)-(14e) is not exact, the resulting solution is not necessarily feasible for the original QCQP (1a)-(1c). In this case, an initial point $\bar{x} \in \mathbb{R}^n$ (either feasible or infeasible) can be used to revise the objective function, resulting in a penalized conic programming relaxation problem of the form:

$$\begin{align*}
\text{minimize}_{x \in \mathbb{R}^n} & \quad \bar{q}_0(x, X) + \eta(\text{tr}(X) - 2x^T x + \hat{x}^T \hat{x}) \tag{16a} \\
\text{subject to} & \quad \bar{q}_k(x, X) \leq 0, \quad k \in \mathcal{I} \tag{16b} \\
& \quad \bar{q}_k(x, X) = 0, \quad k \in \mathcal{E} \tag{16c} \\
& \quad X - xx^T \succeq C_r, \quad 0 \tag{16d} \\
& \quad e_j^T V(x, X) e_j \geq 0, \quad (i, j) \in \mathcal{V} \tag{16e}
\end{align*}$$

where $\eta > 0$ is a fixed penalty parameter. The penalization is said to be tight if the problem (16a)-(16e) has a unique optimal solution $(\hat{x}, \hat{X})$ that satisfies $X = \hat{x}\hat{x}^T$.

The following theorem guarantees that if $\hat{x}$ is feasible and satisfies the LICQ regularity condition, then the solution of (16a)-(16e) is guaranteed to be feasible for the QCQP (1a)-(1c) for an appropriate choice of $\eta$.

**Theorem 1.** Let $\hat{x}$ be a feasible point for the QCQP (1a)-(1b) that satisfies the LICQ condition. If $\eta$ is sufficiently large, then the convex program (16a)-(16e) has a unique optimal solution $(\hat{x}, \hat{X})$ such that $X = \hat{x}\hat{x}^T$. Moreover, $\hat{x}$ is feasible for (1a)-(1c) and satisfies $\bar{q}_0(\hat{x}) \leq q_0(\hat{x})$.

**Proof.** Please refer to Section IV for the proof.

If $\hat{x}$ is not feasible but satisfies the generalized LICQ regularity condition and is close enough to the feasible set $\mathcal{F}$, then the penalization is still tight if $\eta$ is large. This will be explained next.

**Theorem 2.** Let $\hat{x} \in \mathbb{R}^n$ satisfy the GLICQ condition for the QCQP (1a)-(1b), and assume that

$$d_x(\hat{x}) < \frac{s(\hat{x})}{2(1 + C_{n-1, r-1}) \|P\|_2}. \tag{17}$$

If $\eta$ is sufficiently large, then the convex program (16a)-(16e) has a unique optimal solution $(\hat{x}, \hat{X})$ such that $X = \hat{x}\hat{x}^T$ and $\hat{x}$ is feasible for (1a)-(1c).

**Proof.** Please refer to Section IV for the proof.
Algorithm 1 Sequential Penalized Conic Relaxation.

Input: \( r \geq 2, \eta > 0, \hat{x} \in \mathbb{R}^n, \{q_k\}_{k \in \{0\} \cup \mathcal{I} \cup \mathcal{E}} \)
1: repeat
2: \ solve the convex problem \( \text{(16a)-(16c)} \) with the initial point \( \hat{x} \) to obtain \( (\hat{x}, \hat{X}) \)
3: \ set : \( \hat{x} := \hat{X} \)
4: until stopping criterion is met.
Output: \( \hat{x} \)

D. Sequential Penalization

If the initial point \( \hat{x} \) does not result in a tight penalization, the convex problem \( \text{(16a)-(16c)} \) can be solved sequentially by updating the initial point until a feasible and near-globally optimal point is obtained. This procedure is delineated in Algorithm 1. According to Theorem 2, once we are close to the feasible set \( \mathcal{F} \), the relaxation becomes tight. Then, according to Theorem 1, feasibility is preserved and the objective value does not increase.

III. APPLICATIONS TO POLYNOMIAL OPTIMIZATION

In this section, we show that the proposed penalized conic relaxation approach can be used for polynomial optimization as well. A polynomial optimization problem is formulated as

\[
\begin{aligned}
\text{minimize} & \quad u_0(x) \\
\text{subject to} & \quad u_k(x) \leq 0, \quad k \in \mathcal{I} \\
& \quad u_k(x) = 0, \quad k \in \mathcal{E}, \quad (18a)
\end{aligned}
\]

for every \( k \in \{0\} \cup \mathcal{I} \cup \mathcal{E} \), where each function \( u_k : \mathbb{R}^{m} \rightarrow \mathbb{R} \) is a polynomial of arbitrary degree. Problem \( \text{(18a)-(18c)} \) can be reformulated as a QCQP of the form:

\[
\begin{aligned}
\text{minimize} & \quad w_0(x, y) \\
\text{subject to} & \quad w_k(x, y) \leq 0, \quad k \in \mathcal{I} \\
& \quad w_k(x, y) = 0, \quad k \in \mathcal{E} \\
& \quad v_i(x, y) = 0, \quad i \in \mathcal{O}, \quad (19a)
\end{aligned}
\]

where \( y \in \mathbb{R}^{|\mathcal{O}|} \) is an auxiliary variable, and \( v_1, \ldots, v_{|\mathcal{O}|} \) and \( w_0, w_1, \ldots, w_{|\mathcal{O}|} \cup \mathcal{I} \cup \mathcal{E} \) are quadratic functions with the following properties:

- For every \( x \in \mathbb{R}^n \), the function \( v(x, \cdot) : \mathbb{R}^{|\mathcal{O}|} \rightarrow \mathbb{R}^{|\mathcal{O}|} \) is invertible.
- If \( v(x, y) = 0_n \), then \( w_k(x, y) = u_k(x) \) for every \( k \in \{0\} \cup \mathcal{I} \cup \mathcal{E} \).

Based on the above properties, there is a one-to-one correspondence between the feasible sets of \( \text{(18a)-(18c)} \) and \( \text{(19a)-(19d)} \). Moreover, a feasible point \( (\hat{x}, \hat{y}) \) is an optimal solution to the QCQP \( \text{(19a)-(19d)} \) if and only if \( \hat{x} \) is an optimal solution to the polynomial optimization problem \( \text{(18a)-(18c)} \).

Theorem 3 \([72]\). Suppose that \( \{u_k\}_{k \in \{0\} \cup \mathcal{I} \cup \mathcal{E}} \) are polynomials of degree at most \( d \), consisting of \( m \) monomials in total. There exists a QCQP reformulation of the polynomial optimization \( \text{(18a)-(18c)} \) in the form of \( \text{(19a)-(19d)} \), where \( |\mathcal{O}| \leq mn \left( |\log_2(d)| + 1 \right) \).

The next theorem shows that the LICQ regularity of a point \( \hat{x} \in \mathbb{R}^n \) is inherited by the corresponding point \( (\hat{x}, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^m \) of the QCQP \( \text{(19a)-(19d)} \).

Theorem 4. Consider a pair of vectors \( \hat{x} \in \mathbb{R}^n \) and \( \hat{y} \in \mathbb{R}^{|\mathcal{O}|} \) satisfying \( v(\hat{x}, \hat{y}) = 0_n \). The following two statements are equivalent:

1) \( \hat{x} \) is feasible and satisfies the LICQ condition for the polynomial optimization problem \( \text{(18a)-(18c)} \).

2) \( (\hat{x}, \hat{y}) \) is feasible and satisfies the LICQ condition for the QCQP \( \text{(19a)-(19d)} \).

Proof. Please refer to Section IV for the proof.

A simple illustrative example is given to further elaborate on the process of constructing quadratic reformulations.

Example 1. Consider the following three-dimensional polynomial optimization problem:

\[
\begin{aligned}
\text{minimize} & \quad x_1^2 + x_2^2 + 3x_1x_2x_3 \\
\text{subject to} & \quad x_1^2 - 1 \leq 0 \\
& \quad x_3^2 - 1 = 0.
\end{aligned}
\]

To derive a QCQP reformulation of the problem \( \text{(20a)-(20c)} \), we introduce the auxiliary variables \( y_1 \) and \( y_2 \), accounting for the monomials \( x_1^2 \) and \( x_1x_2 \), respectively. Define

\[
\begin{aligned}
w_0(x, y) & \triangleq x_1y_1 + x_2^3 + 3y_2x_3 \\
w_1(x, y) & \triangleq y_1 - 1, \\
w_2(x, y) & \triangleq y_2 - x_1x_2.
\end{aligned}
\]

A QCQP reformulation \( \text{(19a)-(19d)} \) is as follows:

\[
\begin{aligned}
\text{minimize} & \quad x_1y_1 + x_2^3 + 3y_2x_3 \\
\text{subject to} & \quad y_1 - 1 \leq 0 \\
& \quad x_3^2 - 1 = 0 \\
& \quad y_1 - x_1^2 = 0 \\
& \quad y_2 - x_1x_2 = 0.
\end{aligned}
\]

Observe that every feasible point \( (x_1, x_2, x_3) \in \mathbb{R}^3 \) of the problem \( \text{(20a)-(20c)} \) can be mapped into a feasible point \( (x_1, x_2, x_3, x_1x_2) \in \mathbb{R}^5 \) of \( \text{(22a)-(22e)} \) and vice versa. Additionally, the constraint Jacobian matrices for the two problems are given as:

\[
\begin{aligned}
J_{\mathcal{O}}(x) &= \begin{bmatrix} 2x_1 & 0 & 0 \\ 0 & 0 & 2x_3 \end{bmatrix} \\
J_{\mathcal{QCQP}}(x, y) &= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2x_3 & 0 & 0 & 0 \\ -2x_1 & 0 & 0 & 1 & 0 \\ -x_1 & -x_2 & 0 & 0 & 1 \end{bmatrix}.
\end{aligned}
\]

It is straightforward to verify that \( J_{\mathcal{O}} \) has full row rank if and only if \( J_{\mathcal{QCQP}}(x, y) \) has full row rank.
In light of Theorems 3 and 4 with no loss of generality, the proposed penalization method can be applied to the class of polynomial optimization problems as well.

IV. PROOFS

**Proof of Theorem 4.** Due to the equality \( u(\hat{x}) = w(\hat{x}, y) \) as well as the invertibility assumption for \( v(\hat{x}, \cdot) \), we have

\[
\frac{\partial u(\hat{x})}{\partial x} \left[ \frac{\partial v(\hat{x}, y)}{\partial y} \right] = \left[ \frac{\partial w(\hat{x}, \hat{y})}{\partial y} \right]^{-1} \frac{\partial w(\hat{x}, \hat{y})}{\partial x}. \tag{24}
\]

Therefore, \( J_{\text{PO}}(\hat{x}) = \frac{\partial u(\hat{x})}{\partial x} \) is equal to the Schur complement of

\[
J_{\text{QCQP}}(\hat{x}, \hat{y}) = \left[ \frac{\partial w(\hat{x}, \hat{y})}{\partial y} \right]^{-1} \frac{\partial w(\hat{x}, \hat{y})}{\partial x}, \tag{25}
\]

which is the Jacobian matrix of the QCQP (19a)-19d at the point \((\hat{x}, \hat{y})\). As a result, the matrix \( J_{\text{PO}}(\hat{x}) \) is singular if and only if \( J_{\text{QCQP}}(\hat{x}, \hat{y}) \) is singular. □

In order to prove Theorems 1 and 2 it is useful to consider the following optimization problem:

**Lemma 1.** Given an arbitrary \( \hat{x} \in \mathbb{R}^n \) and \( \varepsilon > 0 \), every optimal solution \( \hat{x} \) of the problem (26a)-(26c) satisfies

\[
0 \leq \|\hat{x} - \hat{x}\|_2 - d_F(\hat{x}) \leq \varepsilon \tag{28}
\]

if \( \eta \) is sufficiently large.

**Proof.** Consider an optimal solution \( \hat{x} \). Due to Definition 3 the distance between \( \hat{x} \) and every member of \( F \) is not less than \( d_F(\hat{x}) \), which concludes the left side of (28). Let \( x_d \) be an arbitrary member of the set \( \{ x \in F \mid \|x - \hat{x}\|_2 = d_F(\hat{x}) \} \). Due to the optimality of \( \hat{x} \), we have

\[
q_0(x) + \eta\|x - x\|_2^2 \leq q_0(x_d) + \eta\|x_d - x\|_2^2. \tag{29}
\]

According to the inequalities (29) and (27), one can write

\[
(\eta - \alpha)\|x - x\|_2^2 - \alpha \leq (\eta + \alpha)\|x_d - x\|_2^2 + \alpha \tag{30a}
\]

\[
\Rightarrow \|\hat{x} - x\|_2^2 \leq \|x_d - \hat{x}\|_2^2 + \frac{2\alpha}{\eta - \alpha} (1 + \|x_d - \hat{x}\|_2^2) \tag{30b}
\]

\[
\Rightarrow \|\hat{x} - x\|_2^2 \leq d_F(\hat{x})^2 + \frac{2\alpha}{\eta - \alpha} (1 + d_F(\hat{x})^2), \tag{30c}
\]

which concludes the right side of (28), provided that \( \eta \geq \alpha + 2\alpha(1 + d_F(\hat{x})^2)[\varepsilon^2 + 2\varepsilon d_F(\hat{x})]^{-1} \). □

**Lemma 2.** Assume that \( \hat{x} \in \mathbb{R}^n \) satisfies the GLICQ condition for the problem (26a)-(26c). Given an arbitrary \( \varepsilon > 0 \), every optimal solution \( \hat{x} \) of the problem satisfies

\[
s(\hat{x}) - s(\hat{x}) \leq 2d_F(\hat{x})\|P\|_2 + \varepsilon, \tag{31}\]

if \( \eta \) is sufficiently large.

**Proof.** Let \( \hat{B} \) and \( \check{B} \) denote the sets of quasi-binding constraints for \( x \) and binding constraints for \( \hat{x} \), respectively (based on Definition 4). Due to Lemma 1 for every \( k \in I \setminus \hat{B} \) and every arbitrary \( \varepsilon_1 > 0 \), we have

\[
q_k(\hat{x}) - q_k(\hat{x}) \leq 2\left( A_k\hat{x} + b_k \right)^T (\hat{x} - \hat{x}) + (\hat{x} - \hat{x})^T A_k(\hat{x} - \hat{x}) \leq \|\nabla q_k(\hat{x})\|_2 \|\hat{x} - \hat{x}\|_2 + \|A_k\|_2 \|\hat{x} - \hat{x}\|_2^2 \leq \|\nabla q_k(\hat{x})\|_2 d_F(\hat{x}) + \|A_k\|_2 d_F(\hat{x})^2 + \varepsilon_1 \leq -q_k(\hat{x}), \tag{32}\]

if \( \eta \) is sufficiently large, which yields \( \hat{B} \subseteq \check{B} \). Let \( \nu \in \mathbb{R}^k \) be the left singular vector of \( J_{\text{GLICQ}}(\hat{x}) \), corresponding to the smallest singular value. Hence

\[
s(\hat{x}) = \sigma_{\min}(J_{\text{GLICQ}}(\hat{x})) \geq \sigma_{\min}(J_{\text{GLICQ}}(\hat{x})) \geq ||J_{\text{GLICQ}}(\hat{x})^T \nu\|_2 \tag{33a}
\]

\[
\geq ||J_{\text{GLICQ}}(\hat{x})^T \nu\|_2 - 2\|P\|_2 \|\hat{x} - \hat{x}\|_2 \|\nu\|_2 \geq 2(\hat{x} - \hat{x})^T ||P\|_2 \|\hat{x} - \hat{x}\|_2 - \varepsilon, \tag{33c}\]

if \( \eta \) is large, which concludes the inequality (31). □

**Lemma 3.** Let \( \hat{x} \) be an optimal solution of the problem (26a)-(26c), and assume that \( \hat{x} \) is LICQ regular. There exists a pair of dual vectors \((\gamma, \mu) \in \mathbb{R}^{|I|} \times \mathbb{R}^{|E|}\) that satisfies the following Karush-Kuhn-Tucker (KKT) conditions:

\[
2(\eta I + A_0)(\hat{x} - \hat{x}) + 2(A_0\hat{x} + b) + J(\hat{x})^T [\gamma^T, \mu^T] = 0, \tag{34a}
\]

\[
\gamma_k q_k(\hat{x}) = 0, \quad \forall k \in I. \tag{34b}
\]

**Proof.** Due to the LICQ condition, there exists a pair of dual vectors \((\gamma, \mu) \in \mathbb{R}^{|I|} \times \mathbb{R}^{|E|}\), which satisfies the KKT stationarity and complementarity slackness conditions. Due to stationarity, we have

\[
0 = \nabla \mathcal{L}(\hat{x}, \gamma, \mu) = (\eta I + A_0)(\hat{x} - \hat{x}) + (A_0\hat{x} + b) + J(\hat{x})^T [\gamma^T, \mu^T] = 0, \tag{35}\]

Moreover, (34b) is concluded from the complementarity slackness. □

**Lemma 4.** Consider an arbitrary \( \varepsilon > 0 \) and assume that \( \hat{x} \in \mathbb{R}^n \) satisfies the inequality

\[
s(\hat{x}) > 2d_F(\hat{x})\|P\|_2. \tag{36}\]

If \( \eta \) is sufficiently large, for every optimal solution \( \hat{x} \) of the problem (26a)-(26c), there exists a pair of dual vectors \((\gamma, \mu) \in \mathbb{R}^{|I|} \times \mathbb{R}^{|E|}\) that satisfies the inequality

\[
\frac{1}{\eta} \sqrt{\|\gamma\|_2^2 + \|\mu\|_2^2} \leq \frac{2d_F(\hat{x})}{s(\hat{x}) - 2d_F(\hat{x})\|P\|_2} + \varepsilon \tag{37}\]

as well as the equations (34a) and (34b).
Proof. Due to Lemma 3, there exists \((\hat{\gamma}, \hat{\mu}) \in \mathbb{R}^{|\mathcal{I}|}_+ \times \mathbb{R}^{|\mathcal{K}|}\) that satisfies the equations (34a) and (34b). Let \(\tau \triangleq [\gamma^T, \mu^T]^T\) and let \(\mathcal{B}\) be the set of binding constraints for \(\hat{x}\). Due to equations (34a) and (34b), one can write
\[
2(\eta I + A_0)(\hat{x} - \hat{x}) + 2(A_0\hat{x} + b) + J_{\mathcal{B}}(\hat{\tau}) = 0. \tag{38}
\]
Let \(\phi \triangleq s(\hat{x}) - 2d_F(\hat{x})\|P\|_2\) and define
\[
\varepsilon_1 \triangleq \phi \times \frac{1 - 2\eta^{-1}||A_0||_2}{2 + 2\eta^{-1}||A_0||_2 + 2\eta^{-1}d_F(\hat{x})}. \tag{39}
\]
If \(\eta\) is sufficiently large, \(\varepsilon_1\) is positive and based on Lemmas 1 and 2 we have
\[
\|\tau\|_2 = \frac{\|\tau(\hat{\mathcal{B}})\|_2}{\eta} \leq \frac{2||\eta I + A_0||_2(\hat{x} - \hat{x}) + ||A_0\hat{x} + b\|_2}{\eta \min{\{J_{\mathcal{B}}(\hat{x})\}}} \leq \frac{2(\eta I + A_0)(\hat{x} - \hat{x}) + ||A_0\hat{x} + b\|_2}{\eta s(\hat{x}) - 2d_F(\hat{x})\|P\|_2 - \varepsilon_1} \tag{40}
\]
where the last equality is a result of the equation (39).

Lemma 5. Consider an optimal solution \(\hat{x}\) of the problem (26a)-(26c), and a pair of dual vectors \((\hat{\gamma}, \hat{\mu}) \in \mathbb{R}^{|\mathcal{I}|}_+ \times \mathbb{R}^{|\mathcal{K}|}\) that satisfies the conditions (34a) and (34b). If the matrix inequality
\[
\eta I + A_0 + P(\hat{\gamma}, \hat{\mu}) \succcurlyeq 0, \tag{41}
\]
holds true, then the pair \((\hat{x}, \hat{x}^\top\hat{\gamma})\) is the unique primal solution to the penalized conic relaxation problem (16a)-(16c).

Proof. With no loss of generality, it suffices to prove the lemma for the case \(\mathcal{V} = 0\) only. Let \(\Lambda \in \mathbb{S}^n_+\) denote the dual variable associated with the conic constraint (16d). Then, the KKT conditions for the problem (16a)-(16c) can be written as follows:

\[
\nabla_x \hat{L}(x, X, \gamma, \mu, \Lambda) = 2(Ax - \hat{x}x + b) = 0, \tag{42a}
\]
\[
\nabla_X \hat{L}(x, X, \gamma, \mu, \Lambda) = \eta I + A_0 + P(\gamma, \mu) - \Lambda = 0, \tag{42b}
\]
\[
\gamma_k x_k(x) = 0, \quad \forall k \in \mathcal{I}, \tag{42c}
\]
\[
\langle A, xx^\top - X \rangle = 0, \tag{42d}
\]
where \(\hat{L} : \mathbb{R}^n \times \mathbb{S}^n \times \mathbb{R}^{|\mathcal{I}|}_+ \times \mathbb{R}^{|\mathcal{K}|} \rightarrow \mathbb{R}\) is the Lagrangian function, equations (42a) and (42b) account for stationarity with respect to \(x\) and \(X\), respectively, and equations (42c) and (42d) are the complementary slackness conditions for the constraints (16b) and (16d), respectively. Define
\[
\hat{\Lambda} = \eta I + A_0 + P(\hat{\gamma}, \hat{\mu}). \tag{43}
\]
Due to Lemma 3, if \(\eta\) is sufficiently large, \(\hat{x}\) and \((\hat{\gamma}, \hat{\mu})\) satisfy the equations (34a) and (34b), which yield the optimality conditions (42a)-(42d) if \(x = \hat{x}, X = \hat{x}\hat{x}^\top, \gamma = \hat{\gamma}, \mu = \hat{\mu}\), and \(\Lambda = \hat{\Lambda}\). Therefore, the pair \((\hat{x}, \hat{x}^\top\hat{\gamma})\) is a primal optimal point for the penalized conic relaxation problem (16a)-(16c).

Since the KKT conditions hold between every arbitrary pair of primal and dual solutions, we have
\[
x = \hat{\Lambda}^{-1}(\hat{\eta} \hat{x} - b) \tag{44}
\]
and \(X = \hat{x}\hat{x}^\top\), according to the equations (42a) and (42d), respectively, which imply the uniqueness of the solution.

Lemma 6. Consider an optimal solution \(\hat{x}\) of the problem (26a)-(26c), and a pair of dual vectors \((\hat{\gamma}, \hat{\mu}) \in \mathbb{R}^{|\mathcal{I}|}_+ \times \mathbb{R}^{|\mathcal{K}|}\) that satisfies the conditions (34a) and (34b). If the inequality
\[
\frac{1}{\eta} \sqrt{\|\hat{\gamma}\|_2^2 + \|\hat{\mu}\|_2^2} < \frac{1}{\eta \|P\|_2 - \|A_0\|_2} \tag{45}
\]
holds true, then the pair \((\hat{x}, \hat{x}^\top\hat{\gamma})\) is the unique primal solution to the penalized convex relaxation problem (16a)-(16c).

Proof. Based on Lemma 5, it suffices to prove the conic inequality (41). Define
\[
K \triangleq A_0 + P(\hat{\gamma}, \hat{\mu}). \tag{46}
\]
It follows that
\[
\|K\|_2 \leq \|A_0\|_2 + \sum_{k \in \mathcal{I}} \hat{\gamma}_k \|A_k\|_2 + \sum_{k \in \mathcal{K}} \hat{\mu}_k \|A_k\|_2, \tag{47a}
\]
\[
\leq \|A_0\|_2 + \|P\|_2 \sqrt{\|\hat{\gamma}\|_2^2 + \|\hat{\mu}\|_2^2}. \tag{47b}
\]
Let \(\mathcal{R}\) be the set of all \(r\)-member subsets of \(\{1, 2, \ldots , n\}\). Hence,
\[
\eta I + K = \sum_{K \in \mathcal{R}} I\{K\}^\top R_K I\{K\}, \tag{48}
\]
where
\[
R_K = \left(\begin{array}{c}
(n - 1) \quad -1
\end{array}\right)^{-1} \left(\begin{array}{c}
\|I\{K\} + K\{K, K\}\|
\end{array}\right). \tag{49}
\]
Due to the inequalities (45) and (47), we have \(R_K > 0\) for every \(K \in \mathcal{R}\), which proves that \(\eta I + K \succcurlyeq 0\).

Proof of Theorem 2 Let \(\hat{x}\) be an optimal solution of the problem (26a)-(26c). According to the assumption (17), the inequality (36) holds true, and due to Lemma 4 if \(\eta\) is sufficiently large, there exists a corresponding pair of dual vectors \((\gamma, \mu)\) that satisfies the inequality (37). Now, according to the inequality (17), we have
\[
\frac{2d_F(\hat{x})}{s(\hat{x}) - 2d_F(\hat{x})\|P\|_2} \leq \frac{1}{C_{n-1,r-1}\|P\|_2} \tag{50}
\]
and therefore (37) concludes (45). Hence, according to Lemma 6, the pair \((\hat{x}, \hat{x}^\top\hat{\gamma})\) is the unique primal solution to the penalized convex relaxation problem (16a)-(16c).

Proof of Theorem 7 If \(\hat{x}\) is feasible, then \(d_F(\hat{x}) = 0\). Therefore, the tightness of the penalization for Theorem 1 is a direct consequence of Theorem 2. Denote the unique optimal solution of the penalized relaxation as \((\hat{x}, \hat{x}^\top\hat{\gamma})\). Then it is straightforward to verify the inequality \(q_0(\hat{x}) \leq q_0(\hat{x})\) by evaluating the objective function (16a) at the point \((\hat{x}, \hat{x}^\top\hat{\gamma})\).
V. NUMERICAL EXPERIMENTS

In this section, the effectiveness of the proposed method is verified by extensive experiments on non-convex QCQPs from the library of quadratic programming instances (QPLIB) [73]. The experiments are all performed on a desktop computer with a 12-core 3.0GHz CPU and 256GB RAM. MOSEK v8.1 [74] is used through MATLAB 2017a to solve the resulting convex relaxations.

A. Convex Relaxation without Penalization

Table I reports the outcome of the unpenalized convex relaxation problem (14a)-(14c) with the following settings:

- V = Ø and r = n (SDP relaxation),
- V = H × H and no conic inequality (RLT relaxation),
- V = H × H and r = 2 (SOCP+RLT relaxation),
- V = H × H and r = 2 (SOCP+RLT relaxation).

For each benchmark QCQP and each convex relaxation, the following measurements are reported:

- “t(s)”: solver time,
- “FV”: feasibility violation of the resulting output,
- “LB”: optimal cost of convex relaxation, and
- “GAP(%)”: percentage gap from the solution reported by the QPLIB library.

More precisely, we have LB ≜ q0(ŷ, Ź), and

\[ \text{FV} \triangleq \text{tr}\{\hat{X} - \hat{x}\hat{x}^T\} \]

\[ \text{GAP(\%)} \triangleq 100 \times \frac{q_0(x^{QPLIB}) - q_0(ŷ, Ź)}{q_0(x^{QPLIB})}, \]

where (ŷ, Ź) and x^{QPLIB} denote solutions from convex relaxation and QPLIB, respectively.

The cases 0911, 0975, 1055, 1886, 1972, 1931, and 1967 do not have affine constraints (i.e., H = Ø ) and, therefore, imposing RLT cuts do not make any difference for these cases. As demonstrated in Table I the SDP and SOCP relaxations perform poorly in the absence of RLT cuts. Additionally, there is no major difference between the SOCP and SOCP+RLT relaxations in terms of the resulting gaps. However, the SDP+RLT combination performs satisfactorily for all of the cases with affine constraints. Observe that for the cases 0343, 1157, 1353 and 1507 the unpenalized SDP+RLT relaxation is exact and produces a globally optimal solution.

B. Sequential Penalization with Flat Start

Based on the previous experiment, a question arises as to whether we can obtain feasible and near globally-optimal solutions for the test cases in Table I using penalization and without high-quality choices for the initial points (flat start). Tables II and IV report the results of Algorithm I for SOCP, SDP and SDP+RLT relaxations, respectively. For this experiment, we use the given lower and upper bound vectors \( x^{lb} \in (\{-\infty\} \cup \mathbb{R})^n \) and \( x^{ub} \in (\mathbb{R} \cup \{+\infty\})^n \) to create the initial point \( x^{flat} \) for Algorithm I where

\[ x^{flat}_k = \begin{cases} \frac{x^{lb}_k + x^{ub}_k}{2} & \text{if } x^{lb}_k > -\infty \land x^{ub}_k < \infty \\ x^{lb}_k & \text{if } x^{lb}_k > -\infty \land x^{ub}_k = \infty \\ x^{ub}_k & \text{if } x^{lb}_k = -\infty \land x^{ub}_k < \infty \\ 0 & \text{if } x^{lb}_k = -\infty \land x^{ub}_k = \infty \end{cases}, \]

for every \( k \in \{1, \ldots, n\} \). The penalty parameter \( \eta \) is chosen via bisection as the smallest number of the form \( \alpha \times 10^\beta \), which results in a tight relaxation during the first three rounds, where \( \alpha \in \{1, 2, 5\} \) and \( \beta \) is an integer. In all of the experiments, the value of \( \eta \) has remained static throughout Algorithm I. Denote the sequence of penalized relaxation solutions obtained by Algorithm I as

\[ (x^{(1)}, X^{(1)}), (x^{(2)}, X^{(2)}), (x^{(3)}, X^{(3)}), \ldots \]

The smallest \( i \) such that

\[ \text{tr}\{X^{(i)} - x^{(i)}(x^{(i)})^T\} < 10^{-7} \]

is denoted by \( i^{\text{light}} \), i.e., it is the number of rounds that Algorithm I needs to attain a tight penalization, and the resulting objective value is denoted by \( q_0^{\text{stop}} \triangleq q_0(x^{(i^{\text{stop}})}). \) Moreover, the smallest \( i \) such that

\[ \frac{|q_0(x^{(i-1)}) - q_0(x^{(i)})|}{|q_0(x^{(i)})|} \leq 5 \times 10^{-4} \]

is denoted by \( i^{\text{stop}} \), and \( q_0^{\text{stop}} \triangleq q_0(x^{(i^{\text{stop}})}). \) The following formula is used to calculate the final percentage gaps from the optimal costs reported by the QPLIB library:

\[ \text{GAP(\%)} = 100 \times \frac{q_0^{\text{stop}} - q_0(x^{QPLIB})}{q_0(x^{QPLIB})}. \]

Moreover, t(s) denotes the accumulative solver times in seconds for all of the \( i^{\text{stop}} \) rounds. Our results are compared with BARON [75] and COUENNE [76] by fixing the maximum solver times equal to the accumulative solver times spent by Algorithm I. We ran BARON and COUENNE through GAMS v25.1.2 [77]. The resulting lower bounds, upper bounds and GAPS are reported in Tables II and IV.

As demonstrated in these tables, all three penalized relaxations have successfully obtained feasible points within 4% gaps from QPLIB solutions. Interestingly, with penalization, there is no major benefit in using better convex relaxations in terms of gaps. However, sequential SDP requires a smaller number of rounds compared sequential SOCP to meet the stopping criterion (54). Using any of the relaxations, the infeasible initial points can be rounded to a feasible point with only two round of Algorithm I and all relaxations arrive at satisfactory gaps percentages eventually.

C. Sequential Penalization with SDP+RLT initialization

In this experiment, we examine a better initialization of Algorithm I. Let \( \hat{x}^{\text{SDP+RLT}}, \hat{X}^{\text{SDP+RLT}} \) denote the optimal solution of the unpenalized SDP+RLT relaxation (14a)-(14c). We use the point \( \hat{x} = \hat{x}^{\text{SDP+RLT}} \) as the initial point of the algorithm. The 12 cases for which the unpenalized SDP+RLT relaxation is either unbounded or exact are excluded from this experiment as well as the case 1619 with small SDP+RLT
TABLE I: Unpenalized convex relaxations.

<table>
<thead>
<tr>
<th>Inst</th>
<th>LV</th>
<th>LB</th>
<th>UB</th>
<th>GAP(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>044</td>
<td>545</td>
<td>2.56 15 -12.785 0.00 100.00</td>
<td>-2.874 94.98</td>
<td>0.000 100.00</td>
</tr>
<tr>
<td>057</td>
<td>561</td>
<td>2.89 50 -2.914 1.41 97.37</td>
<td>-1.689 97.37</td>
<td>0.000 100.00</td>
</tr>
<tr>
<td>195</td>
<td>205</td>
<td>2.89 50 -2.914 1.41 97.37</td>
<td>0.000 100.00</td>
<td>-1.468 97.37</td>
</tr>
</tbody>
</table>

TABLE II: Sequential SOCP with flat start.

<table>
<thead>
<tr>
<th>Inst</th>
<th>Sequential SOCP with flat start</th>
<th>BARON</th>
<th>COENE</th>
</tr>
</thead>
<tbody>
<tr>
<td>044</td>
<td>18.56 10 145.78 0.00 100.00</td>
<td>0.280 100.00</td>
<td>0.280 100.00</td>
</tr>
<tr>
<td>057</td>
<td>18.56 10 145.78 0.00 100.00</td>
<td>0.280 100.00</td>
<td>0.280 100.00</td>
</tr>
<tr>
<td>195</td>
<td>18.56 10 145.78 0.00 100.00</td>
<td>0.280 100.00</td>
<td>0.280 100.00</td>
</tr>
</tbody>
</table>

D. Choice of the Penalty Parameter $\eta$

In this experiment the sensitivity of different convex relaxations to the choice of the penalty parameter $\eta$ is tested. To this end, one round of the penalized relaxation problem (16a) and (16c) is solved for a wide range of $\eta$ values. The benchmark case 1143 is used for this purpose. Figures 1 and 2 show the results for $\bar{x} = x^{\text{flat}}$ and $\bar{x} = x^{\text{SDP+RLT}}$, respectively.

If $\eta$ is small, none of the proposed penalized relaxations are tight for the case 1143. As the value of $\eta$ increases, the feasibility violation $tr\{X - xx^T\}$ abruptly vanishes once crossing a certain threshold. According to the figures, all three relaxations produce feasible points for a wide range of $\eta$ values. Remarkably, if $\bar{x}^{\text{SDP+RLT}}$ is used as the initial point and $\mu \sim 1$, then the penalized SDP+RLT relaxation (16a) produces a feasible point for the benchmark case 1143 whose objective value is within $\%0.2$ of the reported optimal cost $q^* (x^{QLP})$.

VI. Conclusions

This paper investigates the design of a global optimization technique for the class of quadratically-constrained quadratic programming (QCQP) problems. A penalization technique is introduced to enhance the solution quality of the conic programming relaxations. Given an arbitrary nominal point (feasible or infeasible) for the original QCQP, a penalized relaxation is formulated by adding a linear term to the objective.
The generalized linear independence constraint qualification (LICQ) condition is introduced as a regularity criterion for nominal points, and it is shown that the solution of the penalized relaxation is feasible for QCQP if the nominal point is regular and close to the feasible set. We show that the proposed penalized conic programming relaxations can be solved sequentially in order to improve solution points. Numerical experiments on QPLIB benchmark cases demonstrate a satisfactory performance of the proposed sequential approach compared to the off-the-shelf solvers BARON and COUENNE.

### VII. ACKNOWLEDGMENT

The authors are grateful to GAMS Development Corporation for providing them with unrestricted access to a full set of solvers throughout the project.

## REFERENCES


### TABLE IV: Sequential SDP+RLT with flat start.

<table>
<thead>
<tr>
<th>Inst</th>
<th>$\eta$</th>
<th>$x^{10^{-5}}$</th>
<th>$x^{10^{-6}}$</th>
<th>$x^{10^{-7}}$</th>
<th>$x^{10^{-8}}$</th>
<th>$x^{10^{-9}}$</th>
<th>GAP(%)</th>
<th>t (s)</th>
<th>LB</th>
<th>UB</th>
<th>UB GAP(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1411</td>
<td>1</td>
<td>-1.000</td>
<td>1</td>
<td>-31.264</td>
<td>1</td>
<td>-75.664</td>
<td>0.00</td>
<td>8.77</td>
<td>-147.837</td>
<td>-75.669</td>
<td>0.00</td>
</tr>
<tr>
<td>1441</td>
<td>2</td>
<td>-75.664</td>
<td>2</td>
<td>-72.375</td>
<td>0.00</td>
<td>6.26</td>
<td>-314.314</td>
<td>32.777</td>
<td>-96.317</td>
<td>-72.377</td>
<td>0.00</td>
</tr>
<tr>
<td>1743</td>
<td>3</td>
<td>-75.664</td>
<td>3</td>
<td>-75.664</td>
<td>0.00</td>
<td>8.77</td>
<td>-98.000</td>
<td>-79.660</td>
<td>147.830</td>
<td>-73.600</td>
<td>0.00</td>
</tr>
<tr>
<td>1744</td>
<td>1</td>
<td>-14.508</td>
<td>8</td>
<td>-14.848</td>
<td>0.80</td>
<td>0.61</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
<tr>
<td>1745</td>
<td>2</td>
<td>-13.035</td>
<td>11</td>
<td>-14.560</td>
<td>0.56</td>
<td>0.69</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
<tr>
<td>1746</td>
<td>3</td>
<td>-12.368</td>
<td>12</td>
<td>-13.641</td>
<td>0.56</td>
<td>0.31</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
<tr>
<td>1747</td>
<td>4</td>
<td>-11.780</td>
<td>13</td>
<td>-13.060</td>
<td>0.36</td>
<td>0.12</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### TABLE V: Sequential SOCP initiated from unpenalized SDP+RLT.

<table>
<thead>
<tr>
<th>Inst</th>
<th>$\eta$</th>
<th>$x^{10^{-5}}$</th>
<th>$x^{10^{-6}}$</th>
<th>$x^{10^{-7}}$</th>
<th>$x^{10^{-8}}$</th>
<th>$x^{10^{-9}}$</th>
<th>GAP(%)</th>
<th>t (s)</th>
<th>LB</th>
<th>UB</th>
<th>UB GAP(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1414</td>
<td>1</td>
<td>-1.000</td>
<td>1</td>
<td>-31.264</td>
<td>1</td>
<td>-75.664</td>
<td>0.00</td>
<td>8.77</td>
<td>-147.837</td>
<td>-75.669</td>
<td>0.00</td>
</tr>
<tr>
<td>1415</td>
<td>2</td>
<td>-75.664</td>
<td>2</td>
<td>-72.375</td>
<td>0.00</td>
<td>6.26</td>
<td>-314.314</td>
<td>32.777</td>
<td>-96.317</td>
<td>-72.377</td>
<td>0.00</td>
</tr>
<tr>
<td>1748</td>
<td>3</td>
<td>-75.664</td>
<td>3</td>
<td>-75.664</td>
<td>0.00</td>
<td>8.77</td>
<td>-98.000</td>
<td>-79.660</td>
<td>147.830</td>
<td>-73.600</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### TABLE VI: Sequential SDP initiated from unpenalized SDP+RLT.

<table>
<thead>
<tr>
<th>Inst</th>
<th>$\eta$</th>
<th>$x^{10^{-5}}$</th>
<th>$x^{10^{-6}}$</th>
<th>$x^{10^{-7}}$</th>
<th>$x^{10^{-8}}$</th>
<th>$x^{10^{-9}}$</th>
<th>GAP(%)</th>
<th>t (s)</th>
<th>LB</th>
<th>UB</th>
<th>UB GAP(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1412</td>
<td>1</td>
<td>-14.508</td>
<td>8</td>
<td>-14.848</td>
<td>0.80</td>
<td>0.61</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
<tr>
<td>1413</td>
<td>2</td>
<td>-13.035</td>
<td>11</td>
<td>-14.560</td>
<td>0.56</td>
<td>0.69</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
<tr>
<td>1747</td>
<td>3</td>
<td>-12.368</td>
<td>12</td>
<td>-13.641</td>
<td>0.56</td>
<td>0.31</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### TABLE VII: Sequential SDP+RLT with flat start.

<table>
<thead>
<tr>
<th>Inst</th>
<th>$\eta$</th>
<th>$x^{10^{-5}}$</th>
<th>$x^{10^{-6}}$</th>
<th>$x^{10^{-7}}$</th>
<th>$x^{10^{-8}}$</th>
<th>$x^{10^{-9}}$</th>
<th>GAP(%)</th>
<th>t (s)</th>
<th>LB</th>
<th>UB</th>
<th>UB GAP(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1412</td>
<td>1</td>
<td>-14.508</td>
<td>8</td>
<td>-14.848</td>
<td>0.80</td>
<td>0.61</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
<tr>
<td>1413</td>
<td>2</td>
<td>-13.035</td>
<td>11</td>
<td>-14.560</td>
<td>0.56</td>
<td>0.69</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
<tr>
<td>1747</td>
<td>3</td>
<td>-12.368</td>
<td>12</td>
<td>-13.641</td>
<td>0.56</td>
<td>0.31</td>
<td>-138.223</td>
<td>-12.448</td>
<td>-19.610</td>
<td>-15.955</td>
<td>0.00</td>
</tr>
</tbody>
</table>


