PENALIZED CONIC RELAXATIONS FOR QUADRATICALLY-CONSTRAINED QUADRATIC PROGRAMMING *

RAMTIN MADANI, MOHSEN KHEIRANDISHFARD, JAVAD LAVAEI, AND ALPER ATAMTÜRK

Abstract. In this paper, we give a new penalized conic programming relaxation for non-convex quadratically-constrained quadratic programs (QCQPs). Incorporating the penalty terms into the objective of convex relaxations enables the retrieval of feasible and near-optimal solutions for non-convex QCQPs. We introduce a generalized linear independence constraint qualification (GLICQ) criterion and prove that any GLICQ regular point that is sufficiently close to the feasible set can be used to construct an appropriate penalty term and recover a feasible solution. As a consequence, we describe a simple sequential penalized conic optimization procedure that preserves feasibility and aims to improve the objective of the solutions at each iteration. Numerical experiments on large-scale system identification problems as well as benchmark instances from the QPLIB library of quadratic programming demonstrate the ability of the proposed penalized conic relaxations in finding near-optimal solutions for non-convex QCQPs.

Key words. Semidefinite programming, nonconvex optimization, nonlinear programming, penalty methods

AMS subject classifications. 90C22, 90C26, 90C30

1. Introduction.

Semi-definite programming (SDP) [39] has been critically important for constructing strong convex relaxations of non-convex optimization problems. In particular, forming hierarchies of SDP relaxations [11, 19, 25–28, 35, 40, 42] has been shown to yield the convex hull of non-convex problems. Geomans and Williamson [15] show that the SDP relaxation objective is within 14% of the optimal value for the MAXCUT problem. SDP relaxations have played a central role in developing numerous approximation algorithms for non-convex optimization problems [16, 17, 29, 38, 47–50]. They are also used within branch-and-bound algorithms [8, 10] for non-convex optimization. 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 [5, 22, 23, 36, 37]. Another direction, pursued in [1, 2, 7, 31, 34, 41], is to use lower-complexity relaxations as alternatives to computationally demanding semidefinite programming relaxations. A relaxation is said to be *exact* if it has the same optimal objective value as the original problem. The exactness of the SDP relaxation has been verified for a variety of problems [9, 22, 24, 44, 45].

1.1. Contributions.

This paper is concerned with non-convex quadratically-constrained quadratic programs (QCQPs) for which SDP or its low order conic relaxations are inexact. In order to recover feasible points to QCQP, we incorporate a linear penalty term into the objective of the conic relaxations and show that feasible and near-globally optimal points can be obtained for the original QCQP by solving the resulting penalized conic relaxation problem. The 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 conic relaxation has a unique solution that is feasible for the original QCQP and its objective value is not worse than that of the initial point. Our second result states that if the initial point is infeasible, but instead is sufficiently

*Ramtin Madani and Mohsen Kheirandishfard are with the Department of Electrical Engineering, the University of Texas at Arlington, Arlington, TX 76019, USA (emails: ramtin.madani@uta.edu, mohsen.kheirandishfard@uta.edu). Javad Lavaei, and Alper Atamtürk are with the Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, USA (emails: lavaei@berkeley.edu, atamturk@berkeley.edu). This work is in part supported by the NSF Award 1809454, a University of Texas System STARs award. Javad Lavaei has been supported by an AFOSR YIP Award and ONR N000141712933. Alper Atamtürk has been supported, in part, by grant FA9550-10-1-0168 from the Office of the Assistant Secretary of Defense for Research and Engineering and the NSF Award 1807260.

close to the feasible set and satisfies a generalized LICQ condition, then the unique optimal solution to the penalized relaxation is feasible for the QCQP. Lastly, motivated by these results on constructing a feasible solution, we propose a sequential procedure for QCQP and demonstrate its performance on benchmark instances from the QPLIB library as well as on large-scale system identification problems.

The success of sequential frameworks and penalized cone programming relaxations in solving bilinear matrix inequalities (BMIs) is demonstrated in [18, 20, 21]. In [4], it is shown that penalized SDP relaxation is able to find the roots of overdetermined systems of polynomial equations. Moreover, the incorporation of penalty terms into the objective of conic relaxations is proven to be effective for solving non-convex optimization problems in power systems [30, 33, 51, 52]. These papers show that penalizing certain physical quantities in power network optimization problems such as reactive power loss and thermal loss facilitates the recovery of feasible points from convex relaxations. In [18], a sequential framework is introduced for solving BMIs without theoretical guarantees. Papers [20, 21] investigate this approach further and offer theoretical results through the notion of generalized Mangasarian-Fromovitz regularity condition. However, these conditions are not valid in the presence of equality constraints and for general QCQPs. Motivated by the success of penalized relaxations, this paper offers a theoretical framework for penalized conic relaxation of general QCQP and, by extension, polynomial optimization problems.

1.2. 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 $\boldsymbol{a}$ and a matrix $\boldsymbol{A}$, the symbols a_i and A_{ij} respectively indicate the i^{th} element of $\boldsymbol{a}$ and the $(i,j)^{\text{th}}$ element of $\boldsymbol{A}$. The symbols $\langle \cdot\,, \cdot \rangle$ and $\|\cdot\|_{\text{F}}$ 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 ℓ_2 norm of vectors, matrices, and matrix pencils. The $n \times n$ identity matrix is denoted by $\boldsymbol{I}_n$. The origin of $\mathbb{R}^n$ is denoted by $\mathbf{0}_n$. The superscript $(\cdot)^\top$ and the symbol $\operatorname{tr}\{\cdot\}$ represent the transpose and trace operators, respectively. Given a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, the notation $\sigma_{\min}(\boldsymbol{A})$ represents the minimum singular value of $\boldsymbol{A}$. The notation $\boldsymbol{A} \succeq 0$ means that $\boldsymbol{A}$ is symmetric positive-semidefinite. For a pair of $n\times n$ symmetric matrices $(\boldsymbol{A}, \boldsymbol{B})$ and proper cone $\mathcal{C} \subseteq \mathbb{S}_n$, the notation $\boldsymbol{A} \succeq_{\mathcal{C}} \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B} \in \mathcal{C}$, whereas $\boldsymbol{A} \succ_{\mathcal{C}} \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ belongs to the interior of $\mathcal{C}$. Given an integer $r > 1$, define $\mathcal{C}_r$ as the cone of $n\times n$ symmetric matrices whose $r\times r$ principal submatrices are all positive semidefinite. Similarly, define $\mathcal{C}_r^*$ as the dual cone of $\mathcal{C}_r$, i.e., the cone of $n \times n$ symmetric matrices with factor-width bounded by r. Given a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and two sets of positive integers $\mathcal{S}_1$ and $\mathcal{S}_2$, define $\boldsymbol{A}\{\mathcal{S}_1, \mathcal{S}_2\}$ as the submatrix of $\boldsymbol{A}$ obtained by removing all rows of $\boldsymbol{A}$ whose indices do not belong to $\mathcal{S}_1$, and all columns of $\boldsymbol{A}$ whose indices do not belong to $\mathcal{S}_2$. Moreover, define $\boldsymbol{A}\{\mathcal{S}_1\}$ as the submatrix of $\boldsymbol{A}$ obtained by removing all rows of $\boldsymbol{A}$ that do not belong to $\mathcal{S}_1$. Given a vector $\boldsymbol{a} \in \mathbb{R}^n$ and a set $\mathcal{F} \subseteq \mathbb{R}^n$, define $d_{\mathcal{F}}(\boldsymbol{a})$ as the minimum distance between $\boldsymbol{a}$ and members of $\mathcal{F}$. Given a pair of integers (n, r), the binomial coefficient "n choose r" is denoted by C_r^n. The notations $\nabla_{\boldsymbol{x}} f(\boldsymbol{a})$ and $\nabla^2_{\boldsymbol{x}} f(\boldsymbol{a})$, respectively, represent the gradient and Hessian of the function f, with respect to the vector $\boldsymbol{x}$, at a point $\boldsymbol{a}$.

1.3. Outline. The remainder of the paper is organized as follows. In section 2, we review the basic lifted and RLT formulations as well as the standard conic relaxations. Section 3 presents the main results of the paper: the penalized conic relaxation, its theoretical analysis on producing a feasible solution along with a generalized linear independence constraint qualification, and finally the sequential penalization procedure. In Section 4 we present numerical experiments to test the effectiveness of the sequential penalization approach for non-convex QCQPs from the library of quadratic programming instances (QPLIB) as well as large-scale system identification problems. Finally, we conclude in section 5 with a few final remarks.

2. Preliminaries. In this section, we review the lifting and reformulation-linearization as well as the standard convex relaxations of QCQP that are necessary for the development of the main results on penalized conic relaxations in Section 3. Consider a general quadratically-constrained quadratic program (QCQP):

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad q_0(\boldsymbol{x}) \tag{2.1a}$$

$$\text{s.t.} \quad q_k(\boldsymbol{x}) \leq 0, \quad k \in \mathcal{I} \tag{2.1b}$$

$$q_k(\boldsymbol{x}) = 0, \quad k \in \mathcal{E}, \tag{2.1c}$$

where $\mathcal{I}$ and $\mathcal{E}$ index the sets of inequality and equality constraints, respectively. For every $k \in \{0\} \cup \mathcal{I} \cup \mathcal{E}$, $q_k : \mathbb{R}^n \to \mathbb{R}$ is a quadratic function of the form $q_k(\boldsymbol{x}) \triangleq \boldsymbol{x}^\top \boldsymbol{A}_k \boldsymbol{x} + 2\boldsymbol{b}_k^\top \boldsymbol{x} + c_k$, where $\boldsymbol{A}_k \in \mathbb{S}_n$, $\boldsymbol{b}_k \in \mathbb{R}^n$, and $c_k \in \mathbb{R}$. Denote $\mathcal{F}$ as the feasible set of the QCQP (2.1a)–(2.1c). To derive the optimality conditions for a given point, it is useful to define the Jacobian matrix of the constraint functions.

DEFINITION 2.1 (Jacobian Matrix). *For every $\hat{\boldsymbol{x}} \in \mathbb{R}^n$, the Jacobian matrix $\mathcal{J}(\hat{\boldsymbol{x}})$ for the constraint functions $\{q_k\}_{k\in\mathcal{I}\cup\mathcal{E}}$ is*

$$\mathcal{J}(\hat{\boldsymbol{x}}) \triangleq [\nabla_{\boldsymbol{x}} q_1(\hat{\boldsymbol{x}}), \ldots, \nabla_{\boldsymbol{x}} q_{|\mathcal{I}\cup\mathcal{E}|}(\hat{\boldsymbol{x}})]^\top. \tag{2.2a}$$

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

Given a feasible point for the QCQP (2.1a)–(2.1c), the well-known linear independence constraint qualification (LICQ) condition can be used as a regularity criterion.

DEFINITION 2.2 (LICQ Condition). *A feasible point $\hat{\boldsymbol{x}} \in \mathcal{F}$ is LICQ regular if the rows of $\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}})$ are linearly independent, where $\hat{\mathcal{B}} \triangleq \{k \in \mathcal{I} \cup \mathcal{E} \mid q_k(\hat{\boldsymbol{x}}) = 0\}$ denotes the set of binding constraints at $\hat{\boldsymbol{x}}$.*

Finding a feasible point for the QCQP (2.1a)–(2.1c), however, is NP-hard as the Boolean Satisfiability Problem (SAT) is a special case. Therefore, in Section 3, we introduce a notion of generalized LICQ as a regularity condition for both feasible and infeasible points.

2.1. Lifting and reformulation-linearization. A common approach for tackling the non-convex QCQP (2.1a)–(2.1c) introduces an auxiliary variable $\boldsymbol{X} \in \mathbb{S}_n$ accounting for $\boldsymbol{x}\boldsymbol{x}^\top$. Then, the objective function (2.1a) and constraints (2.1b)–(2.1c) can be written as linear functions of $\boldsymbol{x}$ and $\boldsymbol{X}$. For every $k \in \{0\} \cup \mathcal{I} \cup \mathcal{E}$, define $\bar{q}_k : \mathbb{R}^n \times \mathbb{S}_n \to \mathbb{R}$ as

$$\bar{q}_k(\boldsymbol{x}, \boldsymbol{X}) \triangleq \langle \boldsymbol{A}_k, \boldsymbol{X} \rangle + 2\boldsymbol{b}_k^\top \boldsymbol{x} + c_k. \tag{2.3}$$

Moreover, in the presence of affine constraints, the reformulation-linearization technique (RLT) of Sherali and Adams [43] can be used to produce additional inequalities with respect to $\boldsymbol{x}$ and $\boldsymbol{X}$ to strengthen convex relaxations. Define $\mathcal{L}$ as the set of affine constrains in the QCQP (2.1a)–(2.1c), i.e., $\mathcal{L} \triangleq \{k \in \mathcal{I} \cup \mathcal{E} \mid \boldsymbol{A}_k = \boldsymbol{0}_n\}$. Define also

$$\boldsymbol{H} \triangleq [\,\boldsymbol{B}\{\mathcal{L}\cap\mathcal{I}\}^\top, \boldsymbol{B}\{\mathcal{L}\cap\mathcal{E}\}^\top, -\boldsymbol{B}\{\mathcal{L}\cap\mathcal{E}\}^\top]^\top, \tag{2.4a}$$

$$\boldsymbol{h} \triangleq [\; \boldsymbol{c}\{\mathcal{L}\cap\mathcal{I}\}^\top, \; \boldsymbol{c}\{\mathcal{L}\cap\mathcal{E}\}^\top, - \; \boldsymbol{c}\{\mathcal{L}\cap\mathcal{E}\}^\top\,]^\top, \tag{2.4b}$$

where $\boldsymbol{B} \triangleq [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{|\mathcal{I}\cap\mathcal{E}|}]^\top$ and $\boldsymbol{c} \triangleq [c_1, \ldots, c_{|\mathcal{I}\cap\mathcal{E}|}]^\top$. Every $\boldsymbol{x} \in \mathcal{F}$ satisfies

$$\boldsymbol{H}\boldsymbol{x} + \boldsymbol{h} \leq 0, \tag{2.5}$$

and, as a result, all elements of the matrix

$$\boldsymbol{H}\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{H}^\top + \boldsymbol{h}\boldsymbol{x}^\top\boldsymbol{H}^\top + \boldsymbol{H}\boldsymbol{x}\boldsymbol{h}^\top + \boldsymbol{h}\boldsymbol{h}^\top \tag{2.6}$$

are nonnegative if $\boldsymbol{x}$ is feasible. Hence, the inequality

$$\boldsymbol{e}_i^\top \boldsymbol{V}(\boldsymbol{x}, \boldsymbol{x}\boldsymbol{x}^\top)\boldsymbol{e}_j \geq 0 \tag{2.7}$$

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

$$\boldsymbol{V}(\boldsymbol{x}, \boldsymbol{X}) \triangleq \boldsymbol{H}\boldsymbol{X}\boldsymbol{H}^\top + \boldsymbol{h}\boldsymbol{x}^\top\boldsymbol{H}^\top + \boldsymbol{H}\boldsymbol{x}\boldsymbol{h}^\top + \boldsymbol{h}\boldsymbol{h}^\top, \tag{2.8}$$

$\mathcal{H} \triangleq \{1, 2, \ldots, |\mathcal{L}\cap\mathcal{I}| + 2|\mathcal{L}\cap\mathcal{E}|\}$, and $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{|\mathcal{H}|}$ denote the standard bases in $\mathbb{R}^{|\mathcal{H}|}$.

2.2. Convex relaxation. Consider the following relaxation of QCQP (2.1a)–(2.1c):

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{X}\in\mathbb{S}_n}{\text{minimize}} \quad \bar{q}_0(\boldsymbol{x}, \boldsymbol{X}) \tag{2.9a}$$

$$\text{s.t.} \quad \bar{q}_k(\boldsymbol{x}, \boldsymbol{X}) \leq 0, \qquad k \in \mathcal{I} \tag{2.9b}$$

$$\bar{q}_k(\boldsymbol{x}, \boldsymbol{X}) = 0, \qquad k \in \mathcal{E} \tag{2.9c}$$

$$\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top \succeq_{\mathcal{C}_r} 0 \tag{2.9d}$$

$$\boldsymbol{e}_i^\top \boldsymbol{V}(\boldsymbol{x}, \boldsymbol{X})\boldsymbol{e}_j \geq 0, \qquad (i,j) \in \mathcal{V} \tag{2.9e}$$

where $\mathcal{V} \subseteq \mathcal{H} \times \mathcal{H}$ is a selection of RLT inequalities, the additional conic constraint (2.9d) is a convex relaxation of the equation $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^\top$ and

$$\mathcal{C}_r \triangleq \left\{\boldsymbol{Y} \;\middle|\; \boldsymbol{Y}\{\mathcal{K},\mathcal{K}\} \succeq 0, \;\; \forall\, \mathcal{K} \subseteq \{1,\ldots,n\} \wedge |\mathcal{K}| = r\right\}. \tag{2.10}$$

If $\mathcal{V} \neq \emptyset$, we refer to the convex problem (2.9a)–(2.9e) as the rth-order conic programming relaxation of the QCQP (2.1a)–(2.1c) with RLT inequalities from $\mathcal{V}$. The choices $r = n$ and $r = 2$ yield the well-known semidefinite programming (SDP) and second-order conic programming (SOCP) relaxations, respectively.

If the relaxed problem (2.9a)–(2.9e) has an optimal solution $(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{X}})$ that satisfies $\mathring{\boldsymbol{X}} = \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top$, then the relaxation is said to be *exact* and $\mathring{\boldsymbol{x}}$ is a globally optimal solution for the QCQP (2.1a)–(2.1c). The next section offers a penalization method for addressing the case where the relaxation is not exact.

3. Penalized conic relaxation. If the conic relaxation problem (2.9a)–(2.9e) is not exact, the resulting solution is not necessarily feasible for the original QCQP (2.1a)–(2.1c). In this case, we use an initial point $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ (either feasible or infeasible) to revise the objective function, resulting in a *penalized conic programming relaxation* of the form:

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{X}\in\mathbb{S}_n}{\text{minimize}} \quad \bar{q}_0(\boldsymbol{x}, \boldsymbol{X}) + \eta(\text{tr}\{\boldsymbol{X}\} - 2\hat{\boldsymbol{x}}^\top\boldsymbol{x} + \hat{\boldsymbol{x}}^\top\hat{\boldsymbol{x}}) \tag{3.1a}$$

$$\text{s.t.} \quad \bar{q}_k(\boldsymbol{x}, \boldsymbol{X}) \leq 0, \qquad k \in \mathcal{I} \tag{3.1b}$$

$$\bar{q}_k(\boldsymbol{x}, \boldsymbol{X}) = 0, \qquad k \in \mathcal{E} \tag{3.1c}$$

$$\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top \succeq_{\mathcal{C}_r} 0 \tag{3.1d}$$

$$\boldsymbol{e}_i^\top \boldsymbol{V}(\boldsymbol{x}, \boldsymbol{X})\boldsymbol{e}_j \geq 0, \qquad (i,j) \in \mathcal{V}, \tag{3.1e}$$

where $\eta > 0$ is a fixed penalty parameter. Note that the penalty term $\text{tr}\{\boldsymbol{X}\} - 2\hat{\boldsymbol{x}}^\top\boldsymbol{x} + \hat{\boldsymbol{x}}^\top\hat{\boldsymbol{x}}$ equals zero for $\boldsymbol{X} = \hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^\top$. The penalization is said to be *tight* if problem (3.1a)–(3.1e) has a unique optimal solution $(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{X}})$ that satisfies $\mathring{\boldsymbol{X}} = \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top$. In the next section, we give conditions under which the penalized conic programming relaxation is tight.

3.1. Theoretical analysis. The following theorem guarantees that if $\hat{\boldsymbol{x}}$ is feasible and satisfies the LICQ regularity condition (in Section 2), then the solution of (3.1a)–(3.1e) is guaranteed to be feasible for the QCQP (2.1a)–(2.1c) for an appropriate choice of η.

THEOREM 3.1. *Let $\hat{\boldsymbol{x}}$ be a feasible point for the QCQP (2.1a)–(2.1b) that satisfies the LICQ condition. For sufficiently large $\eta > 0$, the convex problem (3.1a)–(3.1e) has a unique optimal solution $(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{X}})$ such that $\mathring{\boldsymbol{X}} = \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top$. Moreover, $\mathring{\boldsymbol{x}}$ is feasible for (2.1a)–(2.1c) and satisfies $q_0(\mathring{\boldsymbol{x}}) \leq q_0(\hat{\boldsymbol{x}})$.*

If $\hat{\boldsymbol{x}}$ is not feasible, but satisfies a generalized LICQ regularity condition, introduced below, and is close enough to the feasible set $\mathcal{F}$, then the penalization is still tight for large enough $\eta > 0$. This result is described formally in Theorem 3.4. First, we define a distance measure from an arbitrary point in $\mathbb{R}^n$ to the feasible set of the problem.

DEFINITION 3.2 (Feasibility Distance). *The feasibility distance function $d_{\mathcal{F}} : \mathbb{R}^n \to \mathbb{R}$ is defined as*

$$d_{\mathcal{F}}(\hat{\boldsymbol{x}}) \triangleq \min\{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|_2 \,|\, \boldsymbol{x} \in \mathcal{F}\}. \tag{3.2}$$

DEFINITION 3.3 (Generalized LICQ Condition). *For every $\hat{\boldsymbol{x}} \in \mathbb{R}^n$, the set of quasi-binding constraints is defined as*

$$\hat{\mathcal{B}} \triangleq \mathcal{E} \cup \left\{ k \in \mathcal{I} \,\middle|\, q_k(\hat{\boldsymbol{x}}) + \|\nabla q_k(\hat{\boldsymbol{x}})\|_2 d_{\mathcal{F}}(\hat{\boldsymbol{x}}) + \frac{\|\nabla^2 q_k(\hat{\boldsymbol{x}})\|_2}{2}\, d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2 \geq 0 \right\}. \tag{3.3}$$

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

$$s(\hat{\boldsymbol{x}}) \triangleq \begin{cases} \sigma_{\min}(\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}})) & \text{if } \hat{\boldsymbol{x}} \text{ satisfies GLICQ} \\ 0 & \text{otherwise,} \end{cases} \tag{3.4}$$

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

Observe that if $\hat{\boldsymbol{x}}$ is feasible, then $d_{\mathcal{F}}(\hat{\boldsymbol{x}}) = 0$, and GLICQ condition reduces to the LICQ condition. Moreover, GLICQ is satisfied if and only if $s(\hat{\boldsymbol{x}}) > 0$.

THEOREM 3.4. *Let $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ satisfy the GLICQ condition for the QCQP (2.1a)-(2.1b), and assume that*

$$d_{\mathcal{F}}(\hat{\boldsymbol{x}}) < \frac{s(\hat{\boldsymbol{x}})}{2\,(1 + C_{n-1,r-1})\,\|\boldsymbol{P}\|_2}. \tag{3.5}$$

If η is sufficiently large, then the convex problem (3.1a)–(3.1e) has a unique optimal solution $(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{X}})$ such that $\mathring{\boldsymbol{X}} = \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top$ and $\mathring{\boldsymbol{x}}$ is feasible for (2.1a)–(2.1c).

The rest of this section is devoted to proving Theorems 3.1 and 3.4. The next definition introduces the notion of matrix pencil corresponding to the QCQP (2.1a)–(2.1c), which will be used as a sensitivity measure.

DEFINITION 3.5 (Pencil Norm). *For the QCQP (2.1a)-(2.1c), define the corresponding matrix pencil $\boldsymbol{P} : \mathbb{R}^{|\mathcal{I}|} \times \mathbb{R}^{|\mathcal{E}|} \to \mathbb{S}_n$ as follows:*

$$\boldsymbol{P}(\boldsymbol{\gamma}, \boldsymbol{\mu}) \triangleq \sum_{k \in \mathcal{I}} \gamma_k \boldsymbol{A}_k + \sum_{k \in \mathcal{E}} \mu_k \boldsymbol{A}_k. \tag{3.6}$$

Moreover, define the pencil norm $\|\boldsymbol{P}\|_2$ as

$$\|\boldsymbol{P}\|_2 \triangleq \max\left\{\|\boldsymbol{P}(\boldsymbol{\gamma}, \boldsymbol{\mu})\|_2 \,\middle|\, \|\boldsymbol{\gamma}\|_2^2 + \|\boldsymbol{\mu}\|_2^2 = 1\right\}, \tag{3.7}$$

which is upperbounded by $\sqrt{\sum_{k \in \mathcal{I} \cup \mathcal{E}} \|\boldsymbol{A}_k\|_2^2}$.

In order to prove Theorems 3.1 and 3.4, it is convenient to consider the following optimization problem:

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad q_0(\boldsymbol{x}) + \eta\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|_2^2 \tag{3.8a}$$

$$\text{s.t.} \quad q_k(\boldsymbol{x}) \le 0, \qquad k \in \mathcal{I} \tag{3.8b}$$

$$q_k(\boldsymbol{x}) = 0, \qquad k \in \mathcal{E}. \tag{3.8c}$$

Consider an $\alpha > 0$ for which the inequality

$$|q_0(\boldsymbol{x})| \le \alpha\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|_2^2 + \alpha, \tag{3.9}$$

is satisfied for every $\boldsymbol{x}\in\mathbb{R}^n$. If $\eta > \alpha$, then the objective function (3.8a) is lower bounded by $-\alpha$ and its optimal value is attainable within any closed and nonempty subset of $\mathbb{R}^n$.

LEMMA 3.6. *Given an arbitrary $\hat{\boldsymbol{x}}\in\mathbb{R}^n$ and $\varepsilon > 0$, for sufficiently large $\eta > 0$, every optimal solution $\overset{*}{\boldsymbol{x}}$ of the problem* (3.8a)-(3.8c) *satisfies*

$$0 \le \|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2 - d_{\mathcal{F}}(\hat{\boldsymbol{x}}) \le \varepsilon. \tag{3.10}$$

Proof. Consider an optimal solution $\overset{*}{\boldsymbol{x}}$. Due to Definition 3.2, the distance between $\hat{\boldsymbol{x}}$ and every member of $\mathcal{F}$ is not less than $d_{\mathcal{F}}(\hat{\boldsymbol{x}})$, which concludes the left side of (3.10). Let $\boldsymbol{x}_d$ be an arbitrary member of the set $\{\boldsymbol{x}\in\mathcal{F} \mid \|\boldsymbol{x}-\hat{\boldsymbol{x}}\|_2 = d_{\mathcal{F}}(\hat{\boldsymbol{x}})\}$. Due to the optimality of $\overset{*}{\boldsymbol{x}}$, we have

$$q_0(\overset{*}{\boldsymbol{x}}) + \eta\|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2^2 \le q_0(\boldsymbol{x}_d) + \eta\|\boldsymbol{x}_d-\hat{\boldsymbol{x}}\|_2^2. \tag{3.11}$$

According to the inequalities (3.11) and (3.9), one can write

$$(\eta-\alpha)\|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2^2 - \alpha \le (\eta+\alpha)\|\boldsymbol{x}_d-\hat{\boldsymbol{x}}\|_2^2 + \alpha \tag{3.12a}$$

$$\Rightarrow\ \|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2^2 \le \|\boldsymbol{x}_d-\hat{\boldsymbol{x}}\|_2^2 + \frac{2\alpha}{\eta-\alpha}(1+\|\boldsymbol{x}_d-\hat{\boldsymbol{x}}\|_2^2) \tag{3.12b}$$

$$\Rightarrow\ \|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2^2 \le d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2 + \frac{2\alpha}{\eta-\alpha}(1+d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2), \tag{3.12c}$$

which concludes the right side of (3.10), provided that $\eta \ge \alpha + 2\alpha(1+d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2)[\varepsilon^2 + 2\varepsilon d_{\mathcal{F}}(\hat{\boldsymbol{x}})]^{-1}$. ∎

LEMMA 3.7. *Assume that $\hat{\boldsymbol{x}}\in\mathbb{R}^n$ satisfies the GLICQ condition for the problem* (3.8a)–(3.8c). *Given an arbitrary $\varepsilon > 0$, for sufficiently large $\eta > 0$, every optimal solution $\overset{*}{\boldsymbol{x}}$ of the problem satisfies*

$$s(\hat{\boldsymbol{x}}) - s(\overset{*}{\boldsymbol{x}}) \le 2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2 + \varepsilon. \tag{3.13}$$

Proof. Let $\hat{\mathcal{B}}$ and $\overset{*}{\mathcal{B}}$ denote the sets of quasi-binding constraints for $\hat{\boldsymbol{x}}$ and binding constraints for $\overset{*}{\boldsymbol{x}}$, respectively (based on Definition 3.3). Due to Lemma 3.6, for every $k\in\mathcal{I}\setminus\hat{\mathcal{B}}$ and every arbitrary $\varepsilon_1 > 0$, we have

$$\begin{aligned} q_k(\overset{*}{\boldsymbol{x}}) - q_k(\hat{\boldsymbol{x}}) &= 2(\boldsymbol{A}_k\hat{\boldsymbol{x}}+\boldsymbol{b}_k)^\top(\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}) + (\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}})^\top\boldsymbol{A}_k(\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}) \\ &\le \|\nabla q_k(\hat{\boldsymbol{x}})\|_2\|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2 + \|\boldsymbol{A}_k\|_2\|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2^2 \\ &\le \|\nabla q_k(\hat{\boldsymbol{x}})\|_2 d_{\mathcal{F}}(\hat{\boldsymbol{x}}) + \|\boldsymbol{A}_k\|_2 d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2 + \varepsilon_1 < -q_k(\hat{\boldsymbol{x}}), \end{aligned} \tag{3.14}$$

if η is sufficiently large, which yields $\mathring{\mathcal{B}} \subseteq \hat{\mathcal{B}}$. Let $\boldsymbol{\nu} \in \mathbb{R}^{|\hat{\mathcal{B}}|}$ be the left singular vector of $\mathcal{J}_{\hat{\mathcal{B}}}(\mathring{\boldsymbol{x}})$, corresponding to the smallest singular value. Hence

$$s(\mathring{\boldsymbol{x}}) = \sigma_{\min}\{\mathcal{J}_{\mathring{\mathcal{B}}}(\mathring{\boldsymbol{x}})\} \geq \sigma_{\min}\{\mathcal{J}_{\hat{\mathcal{B}}}(\mathring{\boldsymbol{x}})\} = \|\mathcal{J}_{\hat{\mathcal{B}}}(\mathring{\boldsymbol{x}})^\top \boldsymbol{\nu}\|_2 \tag{3.15a}$$

$$\geq \|\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}})^\top \boldsymbol{\nu}\|_2 - \|[\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}}) - \mathcal{J}_{\hat{\mathcal{B}}}(\mathring{\boldsymbol{x}})]^\top \boldsymbol{\nu}\|_2 \tag{3.15b}$$

$$\geq \sigma_{\min}\{\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}})\}\|\boldsymbol{\nu}\|_2 - 2\|\boldsymbol{P}\|_2\|\hat{\boldsymbol{x}} - \mathring{\boldsymbol{x}}\|_2\|\boldsymbol{\nu}\|_2 \tag{3.15c}$$

$$\geq s(\hat{\boldsymbol{x}}) - 2\|\boldsymbol{P}\|_2\|\hat{\boldsymbol{x}} - \mathring{\boldsymbol{x}}\|_2 \tag{3.15d}$$

$$\geq s(\hat{\boldsymbol{x}}) - 2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2 - \varepsilon, \tag{3.15e}$$

if η is large, which concludes the inequality (3.13). ∎

LEMMA 3.8. *Let $\mathring{\boldsymbol{x}}$ be an optimal solution of the problem* (3.8a)–(3.8c)*, and assume that $\mathring{\boldsymbol{x}}$ is LICQ regular. There exists a pair of dual vectors $(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}}) \in \mathbb{R}_+^{|\mathcal{I}|} \times \mathbb{R}^{|\mathcal{E}|}$ that satisfies the following Karush-Kuhn-Tucker (KKT) conditions:*

$$2(\eta\boldsymbol{I} + \boldsymbol{A}_0)(\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + 2(\boldsymbol{A}_0\hat{\boldsymbol{x}} + \boldsymbol{b}_0) + \mathcal{J}(\mathring{\boldsymbol{x}})^\top[\mathring{\boldsymbol{\gamma}}^\top, \mathring{\boldsymbol{\mu}}^\top]^\top = 0, \tag{3.16a}$$

$$\mathring{\gamma}_k q_k(\mathring{\boldsymbol{x}}) = 0, \qquad \forall k \in \mathcal{I}. \tag{3.16b}$$

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

$$\begin{aligned} 0 &= \nabla_{\boldsymbol{x}} \mathcal{L}(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}})/2 \\ &= \eta(\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + (\boldsymbol{A}_0\mathring{\boldsymbol{x}} + \boldsymbol{b}_0) + \boldsymbol{P}(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}})\mathring{\boldsymbol{x}} + \sum_{k\in\mathcal{I}} \mathring{\gamma}_k \boldsymbol{b}_k + \sum_{k\in\mathcal{E}} \mathring{\mu}_k \boldsymbol{b}_k \\ &= (\eta\boldsymbol{I} + \boldsymbol{A}_0)(\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + (\boldsymbol{A}_0\hat{\boldsymbol{x}} + \boldsymbol{b}_0) + \mathcal{J}(\mathring{\boldsymbol{x}})^\top[\mathring{\boldsymbol{\gamma}}^\top, \mathring{\boldsymbol{\mu}}^\top]^\top/2. \end{aligned} \tag{3.17}$$

Moreover, (3.16b) is concluded from the complementary slackness. ∎

LEMMA 3.9. *Consider an arbitrary $\varepsilon > 0$ and suppose $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ satisfies the inequality*

$$s(\hat{\boldsymbol{x}}) > 2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2. \tag{3.18}$$

If η is sufficiently large, for every optimal solution $\mathring{\boldsymbol{x}}$ of the problem (3.8a)–(3.8c)*, there exists a pair of dual vectors $(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}}) \in \mathbb{R}_+^{|\mathcal{I}|} \times \mathbb{R}^{|\mathcal{E}|}$ that satisfies the inequality*

$$\frac{1}{\eta}\sqrt{\|\mathring{\boldsymbol{\gamma}}\|_2^2 + \|\mathring{\boldsymbol{\mu}}\|_2^2} \leq \frac{2d_{\mathcal{F}}(\hat{\boldsymbol{x}})}{s(\hat{\boldsymbol{x}}) - 2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2} + \varepsilon \tag{3.19}$$

as well as the equations (3.16a) *and* (3.16b)*.*

Proof. Due to Lemma 3.8, there exists $(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}}) \in \mathbb{R}_+^{|\mathcal{I}|} \times \mathbb{R}^{|\mathcal{E}|}$ that satisfies the equations (3.16a) and (3.16b). Let $\boldsymbol{\tau} \triangleq [\mathring{\boldsymbol{\gamma}}^\top, \mathring{\boldsymbol{\mu}}^\top]^\top$ and let $\mathring{\mathcal{B}}$ be the set of binding constraints for $\mathring{\boldsymbol{x}}$. Due to equations (3.16a) and (3.16b), one can write

$$2(\eta\boldsymbol{I} + \boldsymbol{A}_0)(\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + 2(\boldsymbol{A}_0\hat{\boldsymbol{x}} + \boldsymbol{b}_0) + \mathcal{J}_{\mathring{\mathcal{B}}}(\mathring{\boldsymbol{x}})^\top \boldsymbol{\tau}\{\mathring{\mathcal{B}}\} = 0. \tag{3.20}$$

Let $\phi \triangleq s(\hat{\boldsymbol{x}}) - 2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2$ and define

$$\varepsilon_1 \triangleq \phi \times \frac{\varepsilon - 2\eta^{-1}\phi^{-1}(\|\boldsymbol{A}_0\hat{\boldsymbol{x}} + \boldsymbol{b}_0\|_2 + d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{A}_0\|_2)}{\varepsilon + 2 + 2\eta^{-1}\|\boldsymbol{A}_0\|_2 + 2\phi^{-1}d_{\mathcal{F}}(\hat{\boldsymbol{x}})}. \tag{3.21}$$

If η is sufficiently large, ε_1 is positive and based on Lemmas 3.6 and 3.7, we have

$$\begin{aligned}\frac{\|\boldsymbol{\tau}\|_2}{\eta} = \frac{\|\boldsymbol{\tau}\{\overset{*}{\mathcal{B}}\}\|_2}{\eta} &\leq \frac{2\|(\eta\boldsymbol{I}+\boldsymbol{A}_0)(\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}})+(\boldsymbol{A}_0\hat{\boldsymbol{x}}+\boldsymbol{b}_0)\|_2}{\eta\sigma_{\min}\{\mathcal{J}_{\overset{*}{\mathcal{B}}}(\overset{*}{\boldsymbol{x}})\}}\\
&\leq \frac{2\eta\|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2+2\|\boldsymbol{A}_0\|_2\|\overset{*}{\boldsymbol{x}}-\hat{\boldsymbol{x}}\|_2+2\|\boldsymbol{A}_0\hat{\boldsymbol{x}}+\boldsymbol{b}_0\|_2}{\eta s(\overset{*}{\boldsymbol{x}})}\\
&\leq \frac{2(d_{\mathcal{F}}(\hat{\boldsymbol{x}})+\varepsilon_1)+2\eta^{-1}[\|\boldsymbol{A}_0\|_2(d_{\mathcal{F}}(\hat{\boldsymbol{x}})+\varepsilon_1)+\|\boldsymbol{A}_0\hat{\boldsymbol{x}}+\boldsymbol{b}_0\|_2]}{s(\hat{\boldsymbol{x}})-2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2-\varepsilon_1}\\
&= \frac{2d_{\mathcal{F}}(\hat{\boldsymbol{x}})}{s(\hat{\boldsymbol{x}})-2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2}+\varepsilon, \qquad (3.22)\end{aligned}$$

where the last equality is a result of the equation (3.21). ∎

LEMMA 3.10. *Consider an optimal solution $\overset{*}{\boldsymbol{x}}$ of the problem* (3.8a)–(3.8c)*, and a pair of dual vectors $(\overset{*}{\boldsymbol{\gamma}},\overset{*}{\boldsymbol{\mu}}) \in \mathbb{R}_+^{|\mathcal{I}|}\times\mathbb{R}^{|\mathcal{E}|}$ that satisfies the conditions* (3.16a) *and* (3.16b)*. If the matrix inequality*

$$\eta\boldsymbol{I}+\boldsymbol{A}_0+\boldsymbol{P}(\overset{*}{\boldsymbol{\gamma}},\overset{*}{\boldsymbol{\mu}}) \succ_{\mathcal{D}_r} 0, \qquad (3.23)$$

holds true, then the pair $(\overset{*}{\boldsymbol{x}},\overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top)$ is the unique primal solution to the penalized convex relaxation problem (3.1a)–(3.1e).

Proof. With no loss of generality, it suffices to prove the lemma for the case $\mathcal{V}=\emptyset$ only. Let $\boldsymbol{\Lambda}\in\mathbb{S}_n^+$ denotes the dual variable associated with the conic constraint (3.1d). Then, the KKT conditions for the problem (3.1a)-(3.1e) can be written as follows:

$$\nabla_{\boldsymbol{x}}\ \bar{\mathcal{L}}(\boldsymbol{x},\boldsymbol{X},\boldsymbol{\gamma},\boldsymbol{\mu},\boldsymbol{\Lambda}) = 2\left(\boldsymbol{\Lambda x}-\eta\hat{\boldsymbol{x}}+\boldsymbol{b}_0+\sum_{k\in\mathcal{I}}\overset{*}{\gamma}_k\boldsymbol{b}_k+\sum_{k\in\mathcal{E}}\overset{*}{\mu}_k\boldsymbol{b}_k\right)=0, \qquad (3.24a)$$

$$\nabla_{\boldsymbol{X}}\ \bar{\mathcal{L}}(\boldsymbol{x},\boldsymbol{X},\boldsymbol{\gamma},\boldsymbol{\mu},\boldsymbol{\Lambda}) = \eta\boldsymbol{I}+\boldsymbol{A}_0+\boldsymbol{P}(\boldsymbol{\gamma},\boldsymbol{\mu})-\boldsymbol{\Lambda}=0, \qquad (3.24b)$$

$$\gamma_k q_k(\boldsymbol{x})=0, \qquad \forall k\in\mathcal{I} \qquad (3.24c)$$

$$\langle\boldsymbol{\Lambda},\, \boldsymbol{x}\boldsymbol{x}^\top-\boldsymbol{X}\rangle = 0, \qquad (3.24d)$$

where $\bar{\mathcal{L}}:\mathbb{R}^n\times\mathbb{S}_n\times\mathbb{R}^{|\mathcal{I}|}\times\mathbb{R}^{|\mathcal{E}|}\times\mathbb{S}_n\to\mathbb{R}$ is the Lagrangian function, equations (3.24a) and (3.24b) account for stationarity with respect to $\boldsymbol{x}$ and $\boldsymbol{X}$, respectively, and equations (3.24c) and (3.24d) are the complementary slackness conditions for the constraints (3.1b) and (3.1d), respectively. Define

$$\overset{*}{\boldsymbol{\Lambda}} \triangleq \eta\boldsymbol{I}+\boldsymbol{A}_0+\boldsymbol{P}(\overset{*}{\boldsymbol{\gamma}},\overset{*}{\boldsymbol{\mu}}). \qquad (3.25)$$

Due to Lemma (3.8), if η is sufficiently large, $\overset{*}{\boldsymbol{x}}$ and $(\overset{*}{\boldsymbol{\gamma}},\overset{*}{\boldsymbol{\mu}})$ satisfy the equations (3.16a) and (3.16b), which yield the optimality conditions (3.24a)-(3.24d), if $\boldsymbol{x}=\overset{*}{\boldsymbol{x}}$, $\boldsymbol{X}=\overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top$, $\boldsymbol{\gamma}=\overset{*}{\boldsymbol{\gamma}}$, $\boldsymbol{\mu}=\overset{*}{\boldsymbol{\mu}}$, and $\boldsymbol{\Lambda}=\overset{*}{\boldsymbol{\Lambda}}$. Therefore, the pair $(\overset{*}{\boldsymbol{x}},\overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top)$ is a primal optimal points for the penalized convex relaxation problem (3.1a)-(3.1e).

Since the KKT conditions hold for every pair of primal and dual solutions, we have

$$\overset{*}{\boldsymbol{x}} = \overset{*}{\boldsymbol{\Lambda}}^{-1}\left(\eta\hat{\boldsymbol{x}}-\boldsymbol{b}_0-\sum_{k\in\mathcal{I}}\overset{*}{\gamma}_k\boldsymbol{b}_k-\sum_{k\in\mathcal{E}}\overset{*}{\mu}_k\boldsymbol{b}_k\right) \qquad (3.26)$$

and $\overset{*}{\boldsymbol{X}}=\overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top$, according to the equations (3.24a) and (3.24d), respectively, which implies the uniqueness of the solution. ∎

LEMMA 3.11. *Consider an optimal solution* $\mathring{\boldsymbol{x}}$ *of the problem* (3.8a)-(3.8c)*, and a pair of dual vectors* $(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}}) \in \mathbb{R}_+^{|\mathcal{I}|} \times \mathbb{R}^{|\mathcal{E}|}$ *that satisfies the conditions* (3.16a) *and* (3.16b)*. If the inequality,*

$$\frac{1}{\eta}\sqrt{\|\mathring{\boldsymbol{\gamma}}\|_2^2 + \|\mathring{\boldsymbol{\mu}}\|_2^2} < \frac{1}{C_{n-1,r-1}\|\boldsymbol{P}\|_2} - \frac{\|\boldsymbol{A}_0\|_2}{\eta\|\boldsymbol{P}\|_2} \tag{3.27}$$

holds true, then the pair $(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top)$ *is the unique primal solution to the penalized convex relaxation problem* (3.1a)–(3.1e).

Proof. Based on Lemma 3.10, it suffices to prove the conic inequality (3.23). Define

$$\boldsymbol{K} \triangleq \boldsymbol{A}_0 + \boldsymbol{P}(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}}). \tag{3.28}$$

It follows that

$$\|\boldsymbol{K}\|_2 \leq \|\boldsymbol{A}_0\|_2 + \sum_{k\in\mathcal{I}} \mathring{\gamma}_k \|\boldsymbol{A}_k\|_2 + \sum_{k\in\mathcal{E}} \mathring{\mu}_k \|\boldsymbol{A}_k\|_2, \tag{3.29a}$$

$$\leq \|\boldsymbol{A}_0\|_2 + \|\boldsymbol{P}\|_2 \sqrt{\|\mathring{\boldsymbol{\gamma}}\|_2^2 + \|\mathring{\boldsymbol{\mu}}\|_2^2}\,. \tag{3.29b}$$

Let $\mathcal{R}$ be the set of all r-member subsets of $\{1, 2, \ldots, n\}$. Hence,

$$\eta \boldsymbol{I} + \boldsymbol{K} = \sum_{\mathcal{K}\in\mathcal{R}} \boldsymbol{I}\{\mathcal{K}\}^\top \boldsymbol{R}_{\mathcal{K}}\, \boldsymbol{I}\{\mathcal{K}\}, \tag{3.30}$$

where

$$\boldsymbol{R}_{\mathcal{K}} = \binom{n-1}{r-1}^{-1} [\eta \boldsymbol{I}\{\mathcal{K},\mathcal{K}\} + \boldsymbol{K}\{\mathcal{K},\mathcal{K}\}]. \tag{3.31}$$

Due to the inequalities (3.27) and (3.29), we have $\boldsymbol{R}_{\mathcal{K}} \succ 0$ for every $\mathcal{K} \in \mathcal{R}$, which proves that $\eta\boldsymbol{I} + \boldsymbol{K} \succ_{\mathcal{D}_r} 0$. ∎

Proof of Theorem 3.4. Let $\mathring{\boldsymbol{x}}$ be an optimal solution of the problem (3.8a)–(3.8c). According to the assumption (3.5), the inequality (3.18) holds true, and due to Lemma 3.9, if η is sufficiently large, there exists a corresponding pair of dual vectors $(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}})$ that satisfies the inequality (3.19). Now, according to the inequality (3.5), we have

$$\frac{2d_{\mathcal{F}}(\hat{\boldsymbol{x}})}{s(\hat{\boldsymbol{x}}) - 2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2} \leq \frac{1}{C_{n-1,r-1}\|\boldsymbol{P}\|_2} \tag{3.32}$$

and therefore (3.19) concludes (3.27). Hence, according to Lemma 3.11, the pair $(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top)$ is the unique primal solution to the penalized convex relaxation problem (3.1a)–(3.1e). ∎

Proof of Theorem 3.1. If $\hat{\boldsymbol{x}}$ is feasible, then $d_{\mathcal{F}}(\hat{\boldsymbol{x}}) = 0$. Therefore, the tightness of the penalization for Theorem 3.1 is a direct consequence of Theorem 3.4. Denote the unique optimal solution of the penalized relaxation as $(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top)$. Then it is straightforward to verify the inequality $q_0(\mathring{\boldsymbol{x}}) \leq q_0(\hat{\boldsymbol{x}})$ by evaluating the objective function (3.1a) at the point $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^\top)$. ∎

3.2. Sequential penalization procedure. In practice, the penalized conic programming relaxation (3.1a)–(3.1e) can be initialized by a point that may not satisfy the conditions of Theorem 3.1 or Theorem 3.4 as these conditions are only sufficient, but not necessary. If the chosen initial point $\hat{\boldsymbol{x}}$ does not result in a tight penalization, the penalized convex relaxation

Algorithm 3.1 Sequential Penalized Conic Relaxation.

initiate $\{q_k\}_{k\in\{0\}\cup\mathcal{I}\cup\mathcal{E}}$, $r \geq 2$, $\hat{\boldsymbol{x}} \in \mathbb{R}^n$, and the fixed parameter $\eta > 0$
while stopping criterion is not met **do**
  solve the penalized problem (3.1a)–(3.1e) with the initial point $\hat{\boldsymbol{x}}$ to obtain $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{X}})$
  set $\hat{\boldsymbol{x}} \leftarrow \overset{*}{\boldsymbol{x}}$
end while
return $\overset{*}{\boldsymbol{x}}$

(3.1a)–(3.1e) can be solved sequentially by updating the initial point until a feasible and near-optimal point is obtained. This procedure is described in Algorithm 3.1.

According to Theorem (3.4), once $\hat{\boldsymbol{x}}$ is close enough to the feasible set $\mathcal{F}$, the relaxation becomes tight, i.e., a feasible solution $\overset{*}{\boldsymbol{x}}$ is recovered as the unique optima solution to (3.1a)–(3.1e). Afterwards, in the subsequent iterations, according to Theorem (3.1), feasibility is preserved and the objective value does not increase. The following example illustrates the application of Algorithm 3.1 for a polynomial optimization problem.

Example 3.12. Consider the following three-dimensional polynomial optimization:

$$\underset{a,b,c\in\mathbb{R}}{\text{minimize}} \quad a \tag{3.33a}$$

$$\text{s.t.} \quad a^5 - b^4 - c^4 + 2a^3 + 2a^2b - 2ab^2 + 6abc - 2 = 0 \tag{3.33b}$$

To derive a QCQP reformulation of the problem (3.33a)–(3.33b), we consider a variable $x \in \mathbb{R}^8$, whose elements account for the monomials a, b, c, a^2, b^2, c^2, ab, and a^3, respectively. This leads to the following QCQP:

$$\underset{\boldsymbol{x}\in\mathbb{R}^8,}{\text{minimize}} \quad x_1 \tag{3.34a}$$

$$\text{s.t.} \quad x_4x_8 - x_5^2 - x_6^2 + 2x_1x_4 + 2x_2x_4 - 2x_1x_5 + 6x_3x_7 - 2 = 0 \tag{3.34b}$$

$$x_4 - x_1^2 = 0 \tag{3.34c}$$

$$x_5 - x_2^2 = 0 \tag{3.34d}$$

$$x_6 - x_3^3 = 0 \tag{3.34e}$$

$$x_7 - x_1x_2 = 0 \tag{3.34f}$$

$$x_8 - x_1x_4 = 0 \tag{3.34g}$$

The transformation of the polynomial optimization to QCQP is standard and it is described in Appendix A for completeness. The global optimal objective value of the above QCQP equals -2.0198 and the lower-bound, offered by the standard SDP relaxation equals -89.8901. In order to solve the above QCQP, we run Algorithm 3.1, equipped with the SDP relaxation (no additional valid inequalities) and penalty term $\eta = 0.025$. The trajectory with three different initializations $\hat{\boldsymbol{x}}^1 = [0, 0, 0, 0, 0, 0, 0]^\top$, $\hat{\boldsymbol{x}}^2 = [-3, 0, 2, 9, 0, 4, 0, 27]^\top$, and $\hat{\boldsymbol{x}}^3 = [0, 4, 0, 0, 16, 0, 0, 0]^\top$ are given in Table 1 and shown in Fig. 1. In all three cases, the algorithm achieves feasibility in 1–8 rounds. Moreover, a feasible solution with less than $\%0.2$ gap from global optimality is attained within 10 rounds in all three cases. The example illustrates that Appendix A is not sensitive to the initial point.

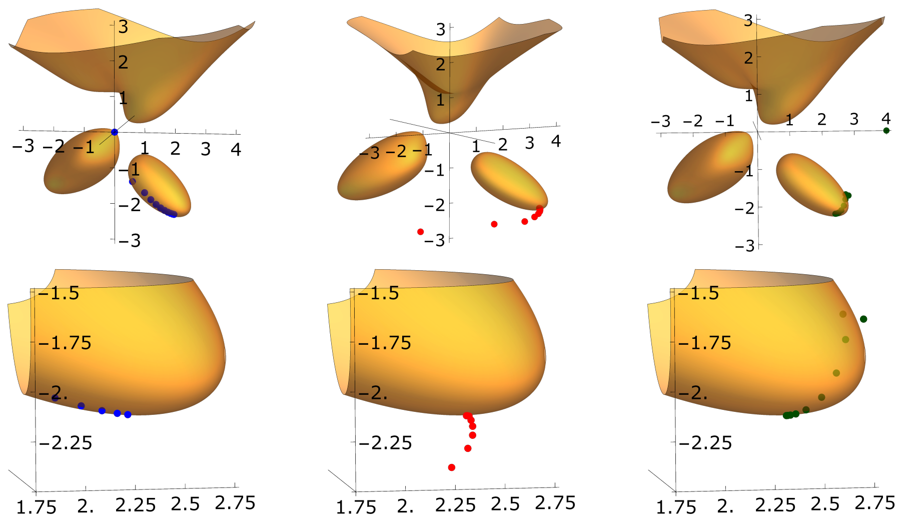

Fig. 1: Trajectory of Algorithm 3.1 for three different initializations. The yellow surface represents the feasible set and the blue, red and green points correspond to $\hat{\boldsymbol{x}}^1$, $\hat{\boldsymbol{x}}^2$ and $\hat{\boldsymbol{x}}^3$, respectively.

Table 1: Trajectory of Algorithm 3.1 for three different initializations.

Round	$\hat{\boldsymbol{x}}^1$				$\hat{\boldsymbol{x}}^2$				$\hat{\boldsymbol{x}}^3$			
	a (obj.)	b	c	$\mathrm{tr}\{\overset{*}{\boldsymbol{X}} - \overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top\}$	a (obj.)	b	c	$\mathrm{tr}\{\overset{*}{\boldsymbol{X}} - \overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top\}$	a (obj.)	b	c	$\mathrm{tr}\{\overset{*}{\boldsymbol{X}} - \overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top\}$
0	0.0000	0.0000	0.0000	-	-3.0000	0.0000	2.0000	-	0.0000	4.0000	0.0000	-
1	-1.2739	0.6601	-0.4697	2.1884	-2.5377	1.2831	-0.7380	138.9796	-1.5721	2.6848	-0.9492	39.2455
2	-1.5173	1.1445	-1.0128	$< 10^{-11}$	-2.4389	2.0715	-1.3946	51.1170	-1.5749	2.7588	-1.3854	13.5140
3	-1.6882	1.3773	-1.2015	$< 10^{-11}$	-2.2889	2.2685	-1.7098	23.0050	-1.6678	2.6583	-1.5228	0.9995
4	-1.8021	1.5739	-1.3561	$< 10^{-11}$	-2.1878	2.3416	-1.8442	11.4963	-1.8322	2.6083	-1.5587	$< 10^{-11}$
5	-1.8824	1.7447	-1.4873	$< 10^{-11}$	-2.1194	2.3621	-1.9007	5.9206	-1.9460	2.5261	-1.6624	$< 10^{-11}$
6	-1.9386	1.8930	-1.5992	$< 10^{-11}$	-2.0733	2.3611	-1.9250	2.9082	-2.0002	2.4391	-1.7847	$< 10^{-11}$
7	-1.9760	2.0180	-1.6923	$< 10^{-11}$	-2.0423	2.3526	-1.9352	1.1594	-2.0156	2.3824	-1.8598	$< 10^{-11}$
8	-1.9985	2.1175	-1.7656	$< 10^{-11}$	-2.0214	2.3426	-1.9393	0.0938	-2.0189	2.3532	-1.8938	$< 10^{-11}$
9	-2.0104	2.1907	-1.8193	$< 10^{-11}$	-2.0197	2.3352	-1.9302	$< 10^{-11}$	-2.0196	2.3387	-1.9079	$< 10^{-11}$
10	-2.0160	2.2408	-1.8559	$< 10^{-11}$	-2.0198	2.3304	-1.9240	$< 10^{-11}$	-2.0197	2.3313	-1.9135	$< 10^{-11}$

4. Numerical experiments. In this section we describe numerical experiments to test the effectiveness of the sequential penalization method for non-convex QCQPs from the library of quadratic programming instances (QPLIB) [13] as well as large-scale system identification problems [12].

4.1. QPLIB problems. The experiments are performed on a desktop computer with a 12-core 3.0GHz CPU and 256GB RAM. MOSEK v8.1 [3] is used through MATLAB 2017a to solve the resulting convex relaxations.

4.1.1. Sequential penalization. Tables 2, 3, 4, and 5 report the results of Algorithm 3.1 for SOCP, SOCP+RLT, SDP, and SDP+RLT relaxations, respectively. The following valid inequalities are imposed on all of the convex relaxations:

$$X_{kk} - (x_k^{\mathrm{lb}} + x_k^{\mathrm{ub}})x_k + x_k^{\mathrm{lb}} x_k^{\mathrm{ub}} \leq 0, \qquad \forall k \in \{1, \dots, n\} \tag{4.1a}$$

$$X_{kk} - (x_k^{\mathrm{ub}} + x_k^{\mathrm{ub}})x_k + x_k^{\mathrm{ub}} x_k^{\mathrm{ub}} \geq 0, \qquad \forall k \in \{1, \dots, n\} \tag{4.1b}$$

$$X_{kk} - (x_k^{\mathrm{lb}} + x_k^{\mathrm{lb}})x_k + x_k^{\mathrm{lb}} x_k^{\mathrm{lb}} \geq 0, \qquad \forall k \in \{1, \dots, n\} \tag{4.1c}$$

where $\boldsymbol{l}, \boldsymbol{u} \in \mathbb{R}^n$ are given lower and upper bounds on $\boldsymbol{x}$. Problem (2.9a)–(2.9e) is solved with the following four settings:

- *SOCP relaxation:* $r = 2$ and valid inequalities (4.1a) – (4.1c).
- *SOCP+RLT relaxation:* $\mathcal{V} = \mathcal{H} \times \mathcal{H}$ and $r = 2$.
- *SDP relaxation:* $r = n$ and valid inequalities (4.1a) – (4.1c).
- *SDP+RLT relaxation:* $\mathcal{V} = \mathcal{H} \times \mathcal{H}$ and $r = n$.

Let $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{X}})$ denote the optimal solution of the convex relaxation (2.9a)-(2.9e). We use the point $\hat{\boldsymbol{x}} = \overset{*}{\boldsymbol{x}}$ as the initial point of the algorithm. For each benchmark QCQP and convex relaxation, the optimal cost of convex relaxation is reported as $\mathrm{LB} \triangleq q_0(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{X}})$.

The penalty parameter η is chosen via bisection as the smallest number of the form $\alpha \times 10^{\beta}$, which results in a tight relaxation during the first six rounds, where $\alpha \in \{1, 2, 5\}$ and β is an integer. In all of the experiments, the value of η has remained static throughout Algorithm 3.1. Denote the sequence of penalized relaxation solutions obtained by Algorithm 3.1 as

$$(\boldsymbol{x}^{(1)}, \boldsymbol{X}^{(1)}), \;\; (\boldsymbol{x}^{(2)}, \boldsymbol{X}^{(2)}), \;\; (\boldsymbol{x}^{(3)}, \boldsymbol{X}^{(3)}), \;\; \dots$$

The smallest i such that

$$\operatorname{tr}\{\boldsymbol{X}^{(i)} - \boldsymbol{x}^{(i)}(\boldsymbol{x}^{(i)})^{\top}\} < 10^{-7} \tag{4.2}$$

is denoted by i^{feas}, i.e., it is the number of rounds that Algorithm 3.1 needs to attain a tight penalization. Moreover, the smallest i such that

$$\frac{q_0(\boldsymbol{x}^{(i-1)}) - q_0(\boldsymbol{x}^{(i)})}{|q_0(\boldsymbol{x}^{(i)})|} \leq 5 \times 10^{-4} \tag{4.3}$$

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

$$\mathrm{GAP}(\%) = 100 \times \frac{q_0^{\text{stop}} - q_0(\boldsymbol{x}^{\text{QPLIB}})}{|q_0(\boldsymbol{x}^{\text{QPLIB}})|}. \tag{4.4}$$

Moreover, t(s) denotes the cumulative solver time in seconds for the i^{stop} rounds. Our results are compared with BARON [46] and COUENNE [6] by fixing the maximum solver times equal to the accumulative solver times spent by Algorithm 3.1. We ran BARON and COUENNE through GAMS v25.1.2 [14]. The resulting lower bounds, upper bounds and GAPs (from the equation (4.4)) are reported in Tables 2, 3, 4, and 5.

As demonstrated in the tables, penalized SOCP+RLT, SDP, and SDP+RLT relaxations have successfully obtained feasible points within 4% gaps from QPLIB solutions. Sequential SDP requires a smaller number of rounds compared sequential SOCP to meet the stopping criterion (4.3). Using any of the relaxations, the infeasible initial points can be rounded to a feasible point with only two round of Algorithm 3.1 and all relaxations arrive at satisfactory gaps percentages.

Figures 2, shows the convergence of Algorithm 3.1 for cases 1507. The choice of η for all curves are taken from the corresponding rows of the Tables 2, 3, 4, and 5.

4.1.2. Choice of the penalty parameter η. In this experiment the sensitivity of different convex relaxations to the choice of the penalty parameter η is tested. To this end, one round of the penalized relaxation problem (3.1a)-(3.1e) is solved for a wide range of η values. The benchmark case 1143 is used for this experiment. If η is small, none of the proposed penalized relaxations are tight for the case 1143. As the value of η increases, the feasibility violation

Table 2: Sequential penalized SOCP relaxation.

Inst	Sequential SOCP relaxation							BARON			COUENNE		
	η	i^{feas}	i^{stop}	t(s)	LB	UB	GAP(%)	LB	UB	GAP(%)	LB	UB	GAP(%)
0343	5e+2	1	100	75.27	-223.281	-5.882	7.89	-95.372	-6.386	0.00	-7668.005	-6.386	0.00
0911	1e+1	1	29	22.91	-76.432	-30.675	4.58	-172.777	0.000	100	-172.777	-31.026	3.49
0975	5e+0	6	18	46.36	-78.263	-36.434	3.75	-47.428	-37.801	0.14	-171.113	-37.213	1.69
1055	1e+1	1	22	14.39	-94.532	-32.620	1.26	-37.841	-33.037	0.00	-199.457	-33.037	0.00
1143	2e+1	1	44	25.68	-178.842	-55.417	3.20	-69.522	-57.247	0.00	-384.45	-56.237	1.76
1157	2e+0	2	9	9.01	-18.715	-10.938	0.10	-11.414	-10.948	0.00	-80.51	-10.948	0.00
1353	5e+0	1	48	84.90	-22.310	-7.700	0.19	-7.925	-7.714	0.00	-73.28	-7.714	0.00
1423	5e+0	1	29	17.44	-31.719	-14.684	1.90	-16.313	-14.968	0.00	-76.13	-14.871	0.65
1437	5e+0	1	36	54.57	-26.473	-7.785	0.06	-9.601	-7.789	0.00	-87.58	-7.789	0.00
1451	2e+1	4	21	20.86	-226.152	-85.598	2.26	-135.140	-87.577	0.00	-468.04	-86.860	0.82
1493	2e+1	1	18	14.49	-137.428	-41.910	2.90	-47.239	-43.160	0.00	-395.69	-43.160	0.00
1507	2e+0	1	15	8.98	-16.635	-8.289	0.15	-49.709	-8.301	0.00	-44.37	-8.301	0.00
1535	5e+0	1	26	28.16	-40.236	-10.948	5.51	-13.407	-11.397	1.63	-107.86	-11.398	1.63
1619	5e+0	1	39	32.34	-31.294	-9.210	0.08	-10.302	-9.217	0.00	-74.55	-9.217	0.00
1661	5e+0	1	32	87.50	-44.147	-15.666	1.81	-19.667	-15.955	0.00	-139.25	-15.955	0.00
1675	2e+1	1	21	36.38	-197.509	-75.485	0.24	-96.864	-75.669	0.00	-435.48	-75.669	0.00
1703	5e+1	2	30	31.82	-408.812	-130.902	1.43	-180.935	-132.802	0.00	-929.92	-132.802	0.00
1745	2e+1	1	26	22.15	-133.719	-71.704	0.93	-77.465	-72.377	0.00	-317.99	-72.377	0.00
1773	5e+0	1	56	148.79	-48.971	-14.154	3.34	-21.581	-14.642	0.00	-118.65	-14.642	0.00
1886	2e+1	1	34	26.82	-163.362	-78.604	0.09	-135.615	-78.672	0.00	-324.87	-78.672	0.00
1913	1e+1	1	28	21.91	-82.384	-51.889	0.42	-68.555	-52.109	0.00	-164.26	-51.478	1.21
1922	1e+1	1	23	11.16	-62.466	-35.437	1.43	-121.872	-35.951	0.00	-123.2	-35.951	0.00
1931	1e+1	1	13	8.78	-102.943	-53.684	3.64	-85.196	-55.709	0.00	-204.08	-54.290	2.55
1967	5e+1	1	32	27.23	-306.859	-105.570	1.87	-136.098	0.000	100	-622.57	-107.581	0.00
Avg	33.9	1.4	31.2	36.68			2.04			8.41			0.58
Max	500	6	100	148.79			7.89			100			3.34

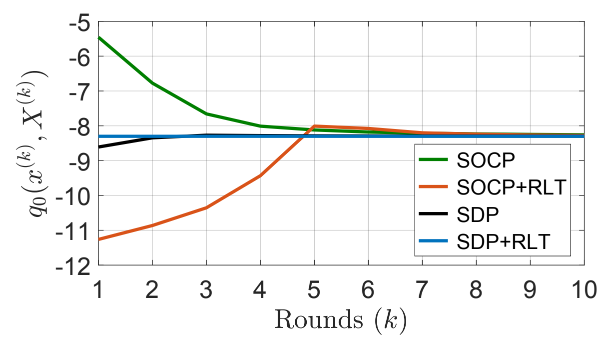

Fig. 2: Convergence of sequential SOCP, SOCP+RLT, SDP, and SDP+RLT relaxations for inst. 1507.

$\operatorname{tr}\{\mathring{\boldsymbol{X}} - \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top\}$ abruptly vanishes once crossing $\eta = 1.9$, $\eta = 7.7$, and $\eta = 19.6$, for the penalized SOCP, SDP and SDP+RLT relaxations, respectively. Remarkably, if $\mathring{\boldsymbol{x}}^{\text{SDP+RLT}}$ is used as the initial point and $\eta \simeq 2$, then the penalized SDP+RLT relaxation (3.1a)-(3.1e) produces a feasible point for the benchmark case 1143 whose objective value is within %0.2 of the reported optimal cost $q_0(\boldsymbol{x}^{\text{QPLIB}})$.

4.2. Large-scale system identification problems. Following [12], this case study is concerned with the problem of identifying the parameters of a linear dynamical system given limited observation and non-uniform snapshots of the state vector. Consider a discrete-time linear system described by the system of equations:

$$\text{(4.5a)} \qquad \boldsymbol{z}[\tau+1] = \boldsymbol{A}\boldsymbol{z}[\tau] + \boldsymbol{B}\boldsymbol{u}[\tau] + \boldsymbol{w}[\tau] \qquad \tau = 1, 2, \ldots, T-1$$

Table 3: Sequential penalized SOCP+RLT relaxation.

Inst	Sequential SOCP+RLT relaxation							BARON			COUENNE		
	η	i^{feas}	i^{stop}	t(s)	LB	UB	GAP(%)	LB	UB	GAP(%)	LB	UB	GAP(%)
0343	1e+2	4	24	25.23	-7.269	-5.945	6.91	-95.372	-6.386	0.00	-7668.005	-6.386	0.00
0911	1e+1	1	33	27.69	-73.061	-30.923	3.81	-172.777	-32.148	0.00	-172.777	-31.026	3.49
0975	5e+0	6	15	4.10	-74.194	-36.300	13.17	-47.428	-37.794	0.16	-171.113	-36.812	2.75
1055	1e+1	1	24	16.78	-90.430	-32.666	1.12	-37.841	-33.037	0.00	-199.457	-33.037	0.00
1143	2e+1	1	30	32.66	-109.302	-55.507	3.04	-69.522	-57.247	0.00	-384.45	-56.237	1.76
1157	2e+0	1	0	1.14	-10.948	-10.948	0.00	-11.414	-10.948	0.00	-80.51	-10.948	0.00
1353	1e+0	3	11	19.41	-10.256	-7.711	0.05	-7.925	-7.714	0.00	-73.28	-7.714	0.00
1423	2e+0	3	14	16.41	-22.462	-14.730	1.59	-16.313	-14.968	0.00	-76.13	-14.871	0.65
1437	5e-1	4	8	21.62	-9.268	-7.788	0.02	-9.601	-7.789	0.00	-87.58	-7.789	0.00
1451	2e+1	2	36	100.50	-185.434	-87.502	0.09	-135.140	-87.577	0.00	-468.04	-87.283	0.34
1493	1e+1	3	13	13.69	-61.053	-41.804	3.14	-47.239	-43.160	0.00	-395.69	-43.160	0.00
1507	1e+0	6	13	10.31	-11.862	-8.295	0.08	-49.709	-8.301	0.00	-44.37	-8.301	0.00
1535	2e+0	3	23	83.47	-21.065	-11.241	2.98	-13.407	-11.586	0.00	-107.86	-11.398	1.62
1619	2e+0	3	20	35.62	-17.163	-9.213	0.05	-10.302	-9.217	0.00	-74.55	-9.217	0.00
1661	1e+0	3	8	35.85	-19.439	-15.666	1.81	-19.667	-15.955	0.00	-139.25	-15.955	0.00
1675	1e+1	3	11	41.30	-121.753	-75.537	0.17	-96.864	-75.669	0.00	-435.48	-75.669	0.00
1703	2e+1	5	22	62.63	-250.703	-131.330	1.11	-180.935	-132.802	0.00	-929.92	-132.802	0.00
1745	5e+0	4	19	40.44	-92.924	-72.351	0.04	-77.465	-72.377	0.00	-317.99	-72.377	0.00
1773	5e+0	1	56	120.65	-29.962	-14.176	3.19	-21.581	-14.642	0.00	-118.65	-14.642	0.00
1886	2e+1	1	35	28.19	-155.747	-78.620	0.07	-135.615	-78.672	0.00	-324.87	-78.672	0.00
1913	5e+0	4	18	15.10	-75.555	-51.879	0.44	-68.555	-52.109	0.00	-164.26	-51.348	1.46
1922	1e+1	1	26	13.22	-57.575	-35.451	1.39	-121.872	-35.951	0.00	-123.2	-35.951	0.00
1931	1e+1	1	13	8.59	-97.100	-53.709	3.59	-85.196	-55.709	0.00	-204.08	-54.290	2.55
1967	5e+1	1	38	33.01	-297.981	-105.616	1.83	-136.098	0.000	100	-622.57	-107.581	0.00
Avg	13.4	2.7	21.3	33.65			2.07			4.17			0.61
Max	100	6	56	120.65			13.17			100			3.49

where

- $\{\boldsymbol{z}[\tau] \in \mathbb{R}^n\}_{\tau=1}^{T}$ are the state vectors that are known at times $\tau \in \{\tau_1, \ldots, \tau_o\}$,
- $\{\boldsymbol{u}[\tau] \in \mathbb{R}^m\}_{\tau=1}^{T}$ are the known control command vectors.
- $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times m}$ are fixed unknown matrices, and
- $\{\boldsymbol{w}[\tau] \in \mathbb{R}^n\}_{\tau=1}^{T}$ account for the unknown disturbance vectors.

Our goal is to estimate the pair of ground truth matrices $(\bar{\boldsymbol{A}}, \bar{\boldsymbol{B}})$, given a sample trajectory of the control commands $\{\bar{\boldsymbol{u}}[\tau] \in \mathbb{R}^n\}_{\tau=1}^{T}$ and the incomplete state vectors $\{\bar{\boldsymbol{z}}[\tau] \in \mathbb{R}^n\}_{\tau\in\{\tau_1,\ldots,\tau_o\}}$. To this end, we employ the minimum least absolute value estimator which amounts to the following QCQP:

$$\text{(4.6a)} \qquad \underset{\substack{\{\boldsymbol{y}[\tau]\in\mathbb{R}^n\}_{\tau=1}^{T-1}\\ \{\boldsymbol{z}[\tau]\in\mathbb{R}^n\}_{\tau=1}^{T}\\ \boldsymbol{A}\in\mathbb{R}^{n\times n}\\ \boldsymbol{B}\in\mathbb{R}^{n\times m}}}{\text{minimize}} \quad \sum_{\tau=1}^{T-1} \mathbf{1}_n^\top \boldsymbol{y}[\tau]$$

$$\text{(4.6b)} \qquad \text{subject to} \quad \boldsymbol{y}[\tau] \geq +\boldsymbol{z}[\tau+1] - \boldsymbol{A}\boldsymbol{z}[\tau] - \boldsymbol{B}\bar{\boldsymbol{u}}[\tau] \qquad \tau \in \{1, 2, \ldots, T-1\},$$

$$\text{(4.6c)} \qquad \boldsymbol{y}[\tau] \geq -\boldsymbol{z}[\tau+1] + \boldsymbol{A}\boldsymbol{z}[\tau] + \boldsymbol{B}\bar{\boldsymbol{u}}[\tau] \qquad \tau \in \{1, 2, \ldots, T-1\},$$

$$\text{(4.6d)} \qquad \boldsymbol{z}[\tau] = \bar{\boldsymbol{z}}[\tau] \qquad \tau \in \{\tau_1, \ldots, \tau_o\}.$$

For every $\tau \in \{1, 2, \ldots, T-1\}$, the auxiliary variable $\boldsymbol{y}[\tau] \in \mathbb{R}^n$ accounts for $|\boldsymbol{z}[\tau+1] - \boldsymbol{A}\boldsymbol{z}[\tau] - \boldsymbol{B}\bar{\boldsymbol{u}}[\tau]|$. This relation is imposed through the pair of constraints (4.6b) and (4.6c).

The problem (4.6a)–(4.6d), can be cast in the form of (2.1a)-(2.1c), with respect to the vector

$$\text{(4.7)} \qquad \boldsymbol{x} \triangleq [\boldsymbol{z}[1]^\top, \ldots, \boldsymbol{z}[T]^\top, \text{vec}\{\boldsymbol{A}\}^\top, \alpha\boldsymbol{y}[1]^\top, \ldots, \alpha\boldsymbol{y}[T-1]^\top, \alpha\text{vec}\{\boldsymbol{B}\}^\top],$$

Table 4: Sequential penalized SDP relaxation.

Inst	Sequential SDP relaxation							BARON			COUENNE		
	η	i^{feas}	i^{stop}	t(s)	LB	UB	GAP(%)	LB	UB	GAP(%)	LB	UB	GAP(%)
0343	1e+2	1	53	29.24	-99.082	-6.379	0.12	-95.372	-6.386	0.00	-7668.005	-6.386	0.00
0911	2e+0	1	9	5.19	-36.068	-31.811	1.05	-172.777	0.000	100	-172.777	-31.026	3.49
0975	2e+0	2	13	8.18	-41.989	-37.845	0.02	-47.428	-37.794	0.16	-171.113	-36.812	2.75
1055	5e+0	1	8	4.36	-36.760	-32.528	1.54	-37.841	-33.037	0.00	-199.457	-33.037	0.00
1143	5e+0	4	15	7.89	-68.328	-55.606	2.87	-69.522	-57.247	0.00	-384.45	-53.367	6.78
1157	1e+0	1	5	3.15	-12.392	-10.945	0.03	-11.414	-10.948	0.00	-80.51	-10.948	0.00
1353	1e+0	1	10	6.12	-9.047	-7.712	0.03	-7.925	-7.714	0.00	-73.28	-7.714	0.00
1423	1e+0	1	5	3.28	-15.933	-14.676	1.95	-16.313	-14.968	0.00	-76.13	-14.078	5.94
1437	1e+0	1	7	4.30	-10.185	-7.787	0.03	-9.601	-7.789	0.00	-87.58	-7.789	0.00
1451†	5e+0	2	6	5.09	-109.318	-85.972	1.83	-135.140	-	-	-468.04	-	-
1493	5e+0	1	6	4.10	-52.396	-43.160	0.00	-47.239	-43.160	0.00	-395.69	-43.160	0.00
1507	5e-1	3	6	3.28	-9.433	-8.291	0.12	-49.709	-8.301	0.00	-44.37	-8.301	0.00
1535	1e+0	1	16	13.05	-13.916	-11.363	1.93	-13.407	-11.397	1.63	-107.86	-11.398	1.63
1619	1e+0	1	7	4.64	-10.376	-9.213	0.05	-10.302	-9.217	0.00	-74.55	-9.217	0.00
1661	1e+0	1	12	7.57	-18.440	-15.955	0.00	-19.667	-15.955	0.00	-139.25	-15.955	0.00
1675	5e+0	1	5	3.75	-93.125	-75.550	0.16	-96.864	-75.669	0.00	-435.48	-75.669	0.00
1703	1e+1	1	10	6.96	-152.774	-132.539	0.20	-180.935	-131.466	1.01	-929.92	-	-
1745†	5e+0	1	8	4.75	-81.668	-71.828	0.76	-77.465	-72.377	0.00	-317.99	-72.377	0.00
1773	1e+0	1	8	5.44	-17.307	-14.633	0.06	-21.581	-14.642	0.00	-118.65	-14.636	0.04
1886	5e+0	2	9	5.84	-87.184	-78.659	0.02	-135.615	-49.684	36.84	-324.87	-78.672	0.00
1913	5e+0	1	20	12.48	-57.441	-51.866	0.47	-68.555	-52.109	0.00	-164.26	-51.348	1.46
1922	5e+0	1	7	4.34	-39.969	-35.452	1.39	-121.872	-35.916	0.10	-123.2	-35.951	0.00
1931	5e+0	1	10	5.87	-60.460	-54.894	1.46	-85.196	-55.709	0.00	-204.08	-54.290	2.55
1967	1e+1	1	6	5.49	-121.990	-104.752	2.63	-136.098	0.000	100	-622.57	-107.581	0.00
Avg	7.6	1.3	11.1	6.92			0.76			10.85			1.12
Max	100	4	53	29.24			2.87			100			6.78

† Rows 1751 and 1745 are excluded from average and maximum computations due to missing entries.

where α is a preconditioning constant. To solve the resulting problem, we use the sequential Algorithm 3.1 equipped with the SOCP relaxation and the initial point $\hat{\boldsymbol{x}} = \boldsymbol{0}$.

We consider system identification problems with $n = 25$, $m = 20$, $T = 500$ and $o = 400$. In every experiment, $\{\tau_1, \ldots, \tau_o\}$ is a uniformly selected subset of $\{1, 2, \ldots, T\}$. The resulting QCQP variable $\boldsymbol{x}$ is 23605-dimensional and the problem is 16100-dimensional if we exclude the known state vectors $\{\bar{\boldsymbol{z}}[\tau] \in \mathbb{R}^n\}_{\tau \in \{\tau_1, \ldots, \tau_o\}}$. Due to sparsity of the QCQP (4.6a)-(4.6d) each round of the penalized SOCP relaxation is solved within 30 minutes, by omitting the elements of the lifted variable $\boldsymbol{X}$ that do not appear in the objective and constraints. All of the convex relaxations are solved using MOSEK v8.1 [3] through MATLAB 2017a and on a desktop computer with a 12-core 3.0GHz CPU and 256GB RAM.

The ground truth values are chosen as follows:

- The elements of $\bar{\boldsymbol{A}} \in \mathbb{R}^{25 \times 25}$ have zero-mean Gaussian distribution and the matrix is scaled in such a way that the largest singular value is equal to 0.5.
- Every element of $\bar{\boldsymbol{B}} \in \mathbb{R}^{25 \times 20}$, $\{\bar{\boldsymbol{u}}[\tau] \in \mathbb{R}^{20}\}_{\tau=1}^{T}$ and $\bar{\boldsymbol{z}}[1] \in \mathbb{R}^{25}$ have standard normal distribution.
- The elements of $\{\bar{\boldsymbol{w}}[\tau] \in \mathbb{R}^{25}\}_{\tau=1}^{T-1}$ have independent zero-mean Gaussian distribution with the standard deviation $\sigma \in \{0.01, 0.02, 0.05, 0.10\}$.

For each experiment, we ran Algorithm 3.1 for 10 rounds. The preconditioning and penalty terms are set to $\alpha = 10^{-3}$ and $\eta = 40$, respectively. For each $\sigma \in \{0.01, 0.02, 0.05, 0.10\}$, we have run 10 random experiments resulting in the average recovery errors 0.0005, 0.0010, 0.0026, and 0.0062, respectively, for $\|\bar{\boldsymbol{A}} - \boldsymbol{A}^{(10)}\|_F / n$, and the average errors 0.0014, 0.0028, 0.0070, and 0.0141, respectively, for $\|\bar{\boldsymbol{B}} - \boldsymbol{B}^{(10)}\|_F / \sqrt{mn}$. In all of the trials, a feasible point is obtained in the first round of Algorithm 3.1. Figure 3 illustrates the convergence behavior of the objective functions for one of the trials for each disturbance level.

Table 5: Sequential penalized SDP+RLT relaxation.

Inst	Sequential SDP+RLT relaxation							BARON			COUENNE		
	η	i^{feas}	i^{stop}	t(s)	LB	UB	GAP(%)	LB	UB	GAP(%)	LB	UB	GAP(%)
0343	0e+0	0	0	1.42	-6.386	-6.386	0.00	-95.372	-6.386	0.00	-7668.005	-6.386	0.00
0911	2e-1	4	5	13.08	-32.982	-32.147	0.00	-172.777	0.000	100	-172.777	-31.026	3.49
0975	2e-1	3	5	12.75	-38.633	-37.852	0.00	-47.428	-37.794	0.16	-171.113	-36.812	2.75
1055	1e+0	5	8	9.56	-33.909	-32.874	0.49	-37.841	-33.037	0.00	-199.457	-33.037	0.00
1143	5e-1	4	5	7.27	-58.908	-57.241	0.01	-69.522	-57.247	0.00	-384.45	-53.367	6.78
1157	0e+0	0	0	0.88	-10.948	-10.948	0.00	-11.414	-10.948	0.00	-80.51	-10.948	0.00
1353	0e+0	0	0	0.45	-7.714	-7.714	0.00	-7.925	-7.714	0.00	-73.28	-7.714	0.00
1423	2e-1	1	2	2.82	-15.154	-14.929	0.25	-16.313	-14.968	0.00	-76.13	-14.078	5.94
1437	1e-2	1	2	7.02	-7.795	-7.789	0.00	-9.601	-7.789	0.00	-87.58	-7.789	0.00
1451	2e+0	2	5	24.45	-94.346	-87.573	0.01	-135.140	-87.577	0.00	-468.04	-86.860	0.82
1493	5e-1	1	2	2.76	-43.883	-43.160	0.00	-47.239	-43.160	0.00	-395.69	-43.160	0.00
1507	0e+0	0	0	0.61	-8.301	-8.301	0.00	-49.709	-8.301	0.00	-44.37	-8.301	0.00
1535	5e-1	1	10	38.01	-12.203	-11.536	0.43	-13.407	-11.397	1.63	-107.86	-11.398	1.62
1619	0e+0	0	0	2.38	-9.217	-9.217	0.00	-10.302	-9.217	0.00	-74.55	-9.217	0.00
1661	1e-1	1	2	12.88	-16.028	-15.955	0.00	-19.667	-15.955	0.00	-139.25	-15.955	0.00
1675	5e-1	4	0	4.22	-76.342	-75.669	0.00	-96.864	-75.669	0.00	-435.48	-75.669	0.00
1703	2e+0	1	3	13.50	-137.543	-132.626	0.13	-180.935	-132.381	0.32	-929.92	-132.802	0.00
1745†	1e+0	6	0	2.53	-73.773	-72.376	0.00	-77.465	-	-	-317.99	-72.377	0.00
1773	2e-1	3	4	18.01	-15.490	-14.626	0.11	-21.581	-14.642	0.00	-118.65	-14.636	0.04
1886	2e+0	2	4	9.05	-81.846	-78.643	0.04	-135.615	-78.672	0.00	-324.87	-78.672	0.00
1913	1e+0	2	6	11.49	-53.290	-52.108	0.00	-68.555	-52.109	0.00	-164.26	-51.348	1.46
1922	2e+0	1	5	3.35	-38.075	-35.556	1.10	-121.872	-35.741	0.58	-123.2	-35.951	0.00
1931	1e+0	1	2	2.99	-56.165	-55.674	0.06	-85.196	-53.760	3.50	-204.08	-54.290	2.55
1967	5e+0	1	8	16.11	-113.802	-107.052	0.49	-136.098	0.000	100	-622.57	-107.581	0.00
Avg	0.8	1.7	3.39	9.35			0.14			8.96			1.11
Max	5	5	10	38			1.1			100			6.78

† Row 1745 is excluded from average and maximum computations due to missing entries.

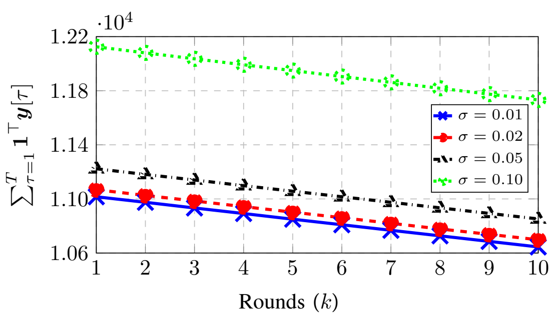

Fig. 3: Convergence of the sequential penalized SOCP relaxation for large-scale system identification with different disturbance levels.

5. Conclusions. This paper introduces a penalized conic relaxation approach for constructing feasible and near-optimal solutions to nonconvex quadratically-constrained quadratic programming (QCQP) problems. Given an arbitrary initial point (feasible or infeasible) for the original QCQP, a penalized relaxation is formulated by adding a linear term to the objective. A generalized linear independence constraint qualification (LICQ) condition is introduced as a regularity criterion for the initial points, and it is shown that the solution of the penalized relaxation is feasible for QCQP if the initial 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 the objective of the feasible solution. Numerical experiments on QPLIB benchmark cases demonstrate that the proposed sequential approach compares favorably with nonconvex optimizers BARON and COUENNE. Moreover, the scalability of the proposed method is demonstrated on large-scale system identification problems.

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

[1] A. A. Ahmadi and A. Majumdar, *DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization*, SIAM Journal on Applied Algebraic Geometry,, (2018).

[2] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming, 95 (2003), pp. 3–51.

[3] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 8.1.*, 2017, http://docs.mosek.com/8.1/toolbox/index.html.

[4] M. Ashraphijuo, R. Madani, and J. Lavaei, *Characterization of rank-constrained feasibility problems via a finite number of convex programs*, in 2016 IEEE 55th Conference on Decision and Control (CDC), IEEE, 2016, pp. 6544–6550.

[5] X. Bao, N. V. Sahinidis, and M. Tawarmalani, *Semidefinite relaxations for quadratically constrained quadratic programming: A review and comparisons*, Mathematical Programming, 129 (2011), pp. 129–157.

[6] P. Belotti, *COUENNE: A user's manual*, tech. report, Technical report, Lehigh University, 2013.

[7] D. Bienstock and G. Munoz, *LP formulations for polynomial optimization problems*, SIAM Journal on Optimization, 28 (2018), pp. 1121–1150.

[8] S. Burer and D. Vandenbussche, *A finite branch-and-bound algorithm for nonconvex quadratic programming via semidefinite relaxations*, Mathematical Programming, 113 (2008), pp. 259–282.

[9] S. Burer and Y. Ye, *Exact semidefinite formulations for a class of (random and non-random) nonconvex quadratic programs*, arXiv preprint arXiv:1802.02688, (2018).

[10] C. Chen, A. Atamtürk, and S. S. Oren, *A spatial branch-and-cut method for nonconvex QCQP with bounded complex variables*, Mathematical Programming, 165 (2017), pp. 549–577.

[11] J. Chen and S. Burer, *Globally solving nonconvex quadratic programming problems via completely positive programming*, Mathematical Programming Computation, 4 (2012), pp. 33–52.

[12] S. Fattahi and S. Sojoudi, *Data-driven sparse system identification*, in 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, 2018.

[13] F. Furini, E. Traversi, P. Belotti, A. Frangioni, A. Gleixner, N. Gould, L. Liberti, A. Lodi, R. Misener, H. Mittelmann, N. Sahinidis, S. Vigerske, and A. Wiegele, *QPLIB: A library of quadratic programming instances*, tech. report, February 2017, http://www.optimization-online.org/DB_HTML/2017/02/5846.html. Available at Optimization Online.

[14] GAMS Development Corporation, *General Algebraic Modeling System (GAMS) Release 24.2.1*. Washington, DC, USA, 2013, http://www.gams.com/.

[15] M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, Journal of the ACM (JACM), 42 (1995), pp. 1115–1145.

[16] S. He, Z. Li, and S. Zhang, *Approximation algorithms for homogeneous polynomial optimization with quadratic constraints*, Mathematical Programming, 125 (2010), pp. 353–383.

[17] S. He, Z. Luo, J. Nie, and S. Zhang, *Semidefinite relaxation bounds for indefinite homogeneous quadratic optimization*, SIAM Journal on Optimization, 19 (2008), pp. 503–523.

[18] S. Ibaraki and M. Tomizuka, *Rank minimization approach for solving BMI problems with random search*, in Proceedings of the 2001 American Control Conference.(Cat. No. 01CH37148), vol. 3, IEEE, 2001, pp. 1870–1875.

[19] C. Josz and D. K. Molzahn, *Lasserre hierarchy for large scale polynomial optimization in real and complex variables*, SIAM Journal on Optimization, 28 (2018), pp. 1017–1048.

[20] M. Kheirandishfard, F. Zohrizadeh, M. Adil, and R. Madani, *Convex relaxation of bilinear matrix inequalities part II: Applications to optimal control synthesis*, in IEEE 57th Annual Conference on Decision and Control (CDC), 2018.

[21] M. Kheirandishfard, F. Zohrizadeh, and R. Madani, *Convex relaxation of bilinear matrix inequalities part I: Theoretical results*, in IEEE 57th Annual Conference on Decision and Control (CDC),, 2018.

[22] S. Kim and M. Kojima, *Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations*, Computational Optimization and Applications, 26 (2003), pp. 143–154.

[23] S. Kim, M. Kojima, and M. Yamashita, *Second order cone programming relaxation of a positive semidefinite constraint*, Optimization Methods and Software, 18 (2003), pp. 535–541.

[24] J. B. Lasserre, *An explicit exact SDP relaxation for nonlinear 0-1 programs*, in Integer Programming and

Combinatorial Optimization, Springer, 2001, pp. 293–303.

[25] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization, 11 (2001), pp. 796–817.

[26] J. B. Lasserre, *Convergent SDP-relaxations in polynomial optimization with sparsity*, SIAM Journal on Optimization, 17 (2006), pp. 822–843.

[27] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and 0–1 optimization*, SIAM journal on optimization, 1 (1991), pp. 166–190.

[28] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and 0–1 optimization*, SIAM Journal on Optimization, 1 (1991), pp. 166–190.

[29] Z. Luo, N. Sidiropoulos, P. Tseng, and S. Zhang, *Approximation bounds for quadratic optimization with homogeneous quadratic constraints*, SIAM Journal on Optimization, 18 (2007), pp. 1–28.

[30] R. Madani, M. Ashraphijuo, and J. Lavaei, *Promises of conic relaxation for contingency-constrained optimal power flow problem*, IEEE Transactions on Power Systems, 31 (2016), pp. 1297–1307.

[31] R. Madani, A. Atamturk, and A. Davoudi, *A scalable semidefinite relaxation approach to grid scheduling*, arXiv preprint arXiv:1707.03541, (2017).

[32] R. Madani, G. Fazelnia, and J. Lavaei, *Rank-2 matrix solution for semidefinite relaxations of arbitrary polynomial optimization problems*, Preprint, (2014).

[33] R. Madani, S. Sojoudi, and J. Lavaei, *Convex relaxation for optimal power flow problem: Mesh networks*, IEEE Transactions on Power Systems, 30 (2015), pp. 199–211.

[34] A. Majumdar, A. A. Ahmadi, and R. Tedrake, *Control and verification of high-dimensional systems with DSOS and SDSOS programming*, in Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on, IEEE, 2014, pp. 394–401.

[35] A. Mohammad-Nezhad and T. Terlaky, *A rounding procedure for semidefinite optimization*, Submitted to Operations Research Letters, (2017).

[36] M. Muramatsu and T. Suzuki, *A new second-order cone programming relaxation for max-cut problems*, Journal of the Operations Research Society of Japan, 46 (2003), pp. 164–177.

[37] K. Natarajan, D. Shi, and K.-C. Toh, *A penalized quadratic convex reformulation method for random quadratic unconstrained binary optimization*, Optimization Online, (2013).

[38] Y. Nesterov, *Semidefinite relaxation and nonconvex quadratic optimization*, Optimization Methods and Software, 9 (1998), pp. 141–160.

[39] Y. Nesterov, A. S. Nemirovskii, and Y. Ye, *Interior-point polynomial algorithms in convex programming*, SIAM, 1994.

[40] D. Papp and F. Alizadeh, *Semidefinite characterization of sum-of-squares cones in algebras*, SIAM Journal on Optimization, 23 (2013), pp. 1398–1423.

[41] F. Permenter and P. Parrilo, *Partial facial reduction: simplified, equivalent sdps via approximations of the psd cone*, Mathematical Programming, (2014), pp. 1–54.

[42] H. D. Sherali and W. P. Adams, *A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems*, SIAM Journal on Discrete Mathematics, 3 (1990), pp. 411–430.

[43] H. D. Sherali and W. P. Adams, *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, vol. 31, Springer Science & Business Media, 2013.

[44] S. Sojoudi and J. Lavaei, *On the exactness of semidefinite relaxation for nonlinear optimization over graphs: Part I*, in Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on, IEEE, 2013, pp. 1043–1050.

[45] S. Sojoudi and J. Lavaei, *On the exactness of semidefinite relaxation for nonlinear optimization over graphs: Part II*, in Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on, IEEE, 2013, pp. 1043–1050.

[46] M. Tawarmalani and N. V. Sahinidis, *A polyhedral branch-and-cut approach to global optimization*, Mathematical Programming, 103 (2005), pp. 225–249.

[47] Y. Ye, *Approximating global quadratic optimization with convex quadratic constraints*, Journal of Global Optimization, 15 (1999), pp. 1–17.

[48] Y. Ye, *Approximating quadratic programming with bound and quadratic constraints*, Mathematical Programming, 84 (1999), pp. 219–226.

[49] S. Zhang, *Quadratic maximization and semidefinite relaxation*, Mathematical Programming, 87 (2000), pp. 453–465.

[50] S. Zhang and Y. Huang, *Complex quadratic optimization and semidefinite programming*, SIAM Journal on Optimization, 87 (2006), pp. 871–890.

[51] F. Zohrizadeh, M. Kheirandishfard, A. Nasir, and R. Madani, *Sequential relaxation of unit commitment with AC transmission constraints*, in IEEE 57th Annual Conference on Decision and Control (CDC), 2018.

[52] F. Zohrizadeh, M. Kheirandishfard, E. Quarm, and R. Madani, *Penalized parabolic relaxation for optimal power flow problem*, in IEEE 57th Annual Conference on Decision and Control (CDC), 2018.

Appendix A. Application 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

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad u_0(\boldsymbol{x}) \tag{A.1a}$$

$$\text{s.t.} \quad u_k(\boldsymbol{x}) \le 0, \qquad k \in \mathcal{I} \tag{A.1b}$$

$$u_k(\boldsymbol{x}) = 0, \qquad k \in \mathcal{E}, \tag{A.1c}$$

for every $k \in \{0\} \cup \mathcal{I} \cup \mathcal{E}$, where each function $u_k : \mathbb{R}^n \to \mathbb{R}$ is a polynomial of arbitrary degree. Problem (A.1a)–(A.1c) can be reformulated as a QCQP of the form:

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{y}\in\mathbb{R}^o}{\text{minimize}} \quad w_0(\boldsymbol{x},\boldsymbol{y}) \tag{A.2a}$$

$$\text{s.t.} \quad w_k(\boldsymbol{x},\boldsymbol{y}) \le 0, \qquad k \in \mathcal{I} \tag{A.2b}$$

$$w_k(\boldsymbol{x},\boldsymbol{y}) = 0, \qquad k \in \mathcal{E} \tag{A.2c}$$

$$v_i(\boldsymbol{x},\boldsymbol{y}) = 0, \qquad i \in \mathcal{O}, \tag{A.2d}$$

where $\boldsymbol{y} \in \mathbb{R}^{|\mathcal{O}|}$ is an auxiliary variable, and $v_1, \ldots, v_{|\mathcal{O}|}$ and $w_0, w_1, \ldots, w_{|\{0\}\cup\mathcal{I}\cup\mathcal{E}|}$ are quadratic functions with the following properties:

- For every $\boldsymbol{x} \in \mathbb{R}^n$, the function $\boldsymbol{v}(\boldsymbol{x},\cdot) : \mathbb{R}^{|\mathcal{O}|} \to \mathbb{R}^{|\mathcal{O}|}$ is invertible,
- If $\boldsymbol{v}(\boldsymbol{x},\boldsymbol{y}) = \mathbf{0}_n$, then $w_k(\boldsymbol{x},\boldsymbol{y}) = u_k(\boldsymbol{x})$ 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 (A.1a)–(A.1c) and (A.2a)–(A.2d). Moreover, a feasible point $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{y}})$ is an optimal solution to the QCQP (A.2a)–(A.2d) if and only if $\overset{*}{\boldsymbol{x}}$ is an optimal solution to the polynomial optimization problem (A.1a)–(A.1c).

THEOREM A.1 ([32]). *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* (A.1a)–(A.1c) *in the form of* (A.2a)–(A.2d)*, where $|\mathcal{O}| \le mn\,(\lfloor \log_2(d) \rfloor + 1)$.*

The next proposition shows that the LICQ regularity of a point $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ is inherited by the corresponding point $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) \in \mathbb{R}^n \times \mathbb{R}^o$ of the QCQP reformulation (A.2a)-(A.2d).

PROPOSITION A.2. *Consider a pair of vectors $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ and $\hat{\boldsymbol{y}} \in \mathbb{R}^{|\mathcal{O}|}$ satisfying $\boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}}) = \mathbf{0}_n$. The following two statements are equivalent:*

1. *$\hat{\boldsymbol{x}}$ is feasible and satisfies the LICQ condition for the polynomial optimization problem* (A.1a)–(A.1b).
2. *$(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})$ is feasible and satisfies the LICQ condition for the QCQP* (A.2a)–(A.2d).

Proof. From $\boldsymbol{u}(\hat{\boldsymbol{x}}) = \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})$ and the invertiblity assumption for $\boldsymbol{v}(\hat{\boldsymbol{x}},\cdot)$, we have

$$\frac{\partial \boldsymbol{u}(\hat{\boldsymbol{x}})}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & -\left(\frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}}\right)^{-1} \frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} \end{bmatrix}^\top$$

$$= \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} - \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}} \left(\frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}}\right)^{-1} \frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}}. \tag{A.3}$$

Therefore, $\mathcal{J}_{\mathrm{PO}}(\hat{\boldsymbol{x}}) = \frac{\partial \boldsymbol{u}(\hat{\boldsymbol{x}})}{\partial \boldsymbol{x}}$ is equal to the Schur complement of

$$\mathcal{J}_{\mathrm{QCQP}}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}}) = \begin{bmatrix} \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}} \\ \frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}} \end{bmatrix}, \tag{A.4}$$

which is the Jacobian matrix of the QCQP (A.2a)–(A.2d) at the point $(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})$. As a result, the matrix $\mathcal{J}_{\mathrm{PO}}(\hat{\boldsymbol{x}})$ is singular if and only if $\mathcal{J}_{\mathrm{QCQP}}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})$ is singular. ∎