PENALIZED CONIC RELAXATIONS FOR QUADRATICALLY-CONSTRAINED QUADRATIC PROGRAMMING *

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 a generalized linear independence constraint qualification (GLICQ) criterion 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 synthetic system identification problems as well as 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.

Key words. example, LaTeX

AMS subject classifications. 68Q25, 68R10, 68U05

1. 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 [49, 52–54], imaging science [7, 15, 27, 68], signal processing [2, 6, 19, 46, 48, 56], automatic control [1, 26, 53, 72], quantum mechanics [14, 24, 35, 44], and cybersecurity [21–23, 60]. 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 [30], phase retrieval in signal processing [17], and matrix completion in machine learning [16, 58].

This paper advances a popular framework for the global analysis of polynomial optimization, which involves convex hull characterization by forming hierarchies of semidefinite programming (SDP) relaxations [20, 36, 42, 43, 57, 64, 66]. 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 branch-and-bound algorithms [12, 18, 63]. 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 [31]. The maximum possible gap between the solution of a graph optimization and its SDP relaxation is defined as the Grothendieck constant of the graph [4, 11]. This constant has been derived for some special cases in [45]. The paper [32] shows how a complex SDP relaxation may solve the max-3-cut problem. This approach has been generalized in several papers [33, 34, 47, 62, 73–76]. If the

*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. A short version of this work entitled "Polynomial Optimization via Penalized Conic Relaxation, " is submitted to the 2018 European Control Conference.

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 [13,39,41,69,70]. 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 [8, 39, 40, 59, 61]. Another direction is pursued in [1,3,10,50,55,65], which is based on lower-complexity relaxations as alternatives to computationally demanding semidefinite programming relaxations.

1.1. Contributions. This paper is concerned with non-convex QCQP for which standard convex relaxations are inexact and fail to produce feasible points. 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 3, 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 point 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 [37, 38] for convex relaxation of bilinear matrix inequalities, as well as in papers [77, 78] which demonstrate the ability of sequential penalized relaxation in solving computationally-hard power system 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 $\boldsymbol{0}_n$. The superscript $(\cdot)^\top$ and the symbol $\text{tr}\{\cdot\}$ 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 every pair of $n \times n$ symmetric matrices $(\boldsymbol{A}, \boldsymbol{B})$ and every 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 $\boldsymbol{a}$.

2. Problem Formulation. Consider a general quadratically-constrained quadratic program (QCQP):

$$
\begin{aligned}
&\text{(2.1a)} \qquad && \underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} && q_0(\boldsymbol{x}) && \\
&\text{(2.1b)} \qquad && \text{subject to} && q_k(\boldsymbol{x}) \le 0 && k \in \mathcal{I} \\
&\text{(2.1c)} \qquad && && q_k(\boldsymbol{x}) = 0 && k \in \mathcal{E},
\end{aligned}
$$

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*

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

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) is NP-hard in general as the Boolean Satisfiability Problem (SAT) is a special case. In this paper, we introduce the notion of generalized LICQ as a regularity condition for both feasible and infeasible points. First, the following definition will provide a distance measure from an arbitrary point in $\mathbb{R}^n$ to the feasible set of the problem.

DEFINITION 2.3 (Feasibility Distance). *The feasibility distance function $d_{\mathcal{F}} : \mathbb{R}^n \to \mathbb{R}$ is defined as*

$$\text{(2.3)} \qquad d_{\mathcal{F}}(\hat{\boldsymbol{x}}) \triangleq \min\{\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|_2 \,|\, \boldsymbol{x} \in \mathcal{F}\}.$$

DEFINITION 2.4 (Generalized LICQ Condition). *For every $\hat{\boldsymbol{x}} \in \mathbb{R}^n$, the set of quasi-binding constraints is defined as*

$$\text{(2.4)} \qquad \hat{\mathcal{B}} \triangleq \mathcal{E} \cup \left\{ k\in\mathcal{I} \,\middle|\, q_k(\hat{\boldsymbol{x}}) + \|\nabla q_k(\hat{\boldsymbol{x}})\|_2 d_{\mathcal{F}}(\hat{\boldsymbol{x}}) + \frac{\|\nabla^2 q_k(\hat{\boldsymbol{x}})\|_2}{2}\, d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2 \ge 0 \right\}.$$

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

$$\text{(2.5)} \qquad s(\hat{\boldsymbol{x}}) \triangleq \begin{cases} \sigma_{\min}(\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}})) & \textit{if } \hat{\boldsymbol{x}} \textit{ satisfies GLICQ} \\ 0 & \textit{otherwise,} \end{cases}$$

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 LICQ and GLICQ conditions are equivalent. Moreover, GLICQ is satisfied if and only if $s(\hat{\boldsymbol{x}}) > 0$. The next definition introduces the notion of matrix pencil corresponding to the QCQP (2.1a)-(2.1c), which will be used later as a sensitivity measure.

DEFINITION 2.5 (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:*

$$\boldsymbol{P}(\boldsymbol{\gamma}, \boldsymbol{\mu}) \triangleq \sum_{k \in \mathcal{I}} \gamma_k \boldsymbol{A}_k + \sum_{k \in \mathcal{E}} \mu_k \boldsymbol{A}_k. \tag{2.6}$$

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

$$\|\boldsymbol{P}\|_2 \triangleq \max\left\{\|\boldsymbol{P}(\boldsymbol{\gamma}, \boldsymbol{\mu})\|_2 \,\middle|\, \|\boldsymbol{\gamma}\|_2^2 + \|\boldsymbol{\mu}\|_2^2 = 1\right\}, \tag{2.7}$$

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

2.1. Lifting and Reformulation-Linearization. A common practice for tackling the non-convex QCQP (2.1a)-(2.1c) involves introducing an auxiliary variable $\boldsymbol{X} \in \mathbb{S}_n$ accounting for $\boldsymbol{x}\boldsymbol{x}^\top$. Using $\boldsymbol{X}$, the objective function (2.1a) and constraints (2.1b)-(2.1c) can be recast in a linear way. 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

$$\bar{q}_k(\boldsymbol{x}, \boldsymbol{X}) \triangleq \langle \boldsymbol{A}_k, \boldsymbol{X} \rangle + 2\boldsymbol{b}_k^\top \boldsymbol{x} + c_k. \tag{2.8}$$

Moreover, in the presence of affine constraints, the well-known reformulation-linearization technique (RLT) or Sherali and Adams [67] 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

$$\boldsymbol{H} \triangleq [\,\boldsymbol{B}\{\mathcal{L} \cap \mathcal{I}\}^\top, \boldsymbol{B}\{\mathcal{L} \cap \mathcal{E}\}^\top, -\boldsymbol{B}\{\mathcal{L} \cap \mathcal{E}\}^\top]^\top, \tag{2.9a}$$

$$\boldsymbol{h} \triangleq [\; \boldsymbol{c}\{\mathcal{L} \cap \mathcal{I}\}^\top,\; \boldsymbol{c}\{\mathcal{L} \cap \mathcal{E}\}^\top, -\; \boldsymbol{c}\{\mathcal{L} \cap \mathcal{E}\}^\top\;]^\top, \tag{2.9b}$$

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

$$\boldsymbol{H}\boldsymbol{x} + \boldsymbol{h} \leq 0, \tag{2.10}$$

and, as a result, all elements of the matrix

$$\boldsymbol{H}\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{H}^\top + \boldsymbol{h}\boldsymbol{x}^\top\boldsymbol{H}^\top + \boldsymbol{H}\boldsymbol{x}\boldsymbol{h}^\top + \boldsymbol{h}\boldsymbol{h}^\top \tag{2.11}$$

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

$$\boldsymbol{e}_i^\top \boldsymbol{V}(\boldsymbol{x}, \boldsymbol{x}\boldsymbol{x}^\top)\boldsymbol{e}_j \geq 0 \tag{2.12}$$

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

$$\boldsymbol{V}(\boldsymbol{x}, \boldsymbol{X}) \triangleq \boldsymbol{H}\boldsymbol{X}\boldsymbol{H}^\top + \boldsymbol{h}\boldsymbol{x}^\top\boldsymbol{H}^\top + \boldsymbol{H}\boldsymbol{x}\boldsymbol{h}^\top + \boldsymbol{h}\boldsymbol{h}^\top, \tag{2.13}$$

$\mathcal{H} \triangleq \{1, 2, \ldots, |\mathcal{L} \cap \mathcal{I}| + 2|\mathcal{L} \cap \mathcal{E}|\}$, 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 reformulation of QCQP (2.1a)-(2.1c):

$$\begin{aligned}
&\text{(2.14a)} \qquad && \underset{\substack{\boldsymbol{x}\in\mathbb{R}^n\\ \boldsymbol{X}\in\mathbb{S}_n}}{\text{minimize}} && \bar{q}_0(\boldsymbol{x},\boldsymbol{X}) && \\
&\text{(2.14b)} && \text{subject to} && \bar{q}_k(\boldsymbol{x},\boldsymbol{X}) \leq 0 && k\in\mathcal{I}\\
&\text{(2.14c)} && && \bar{q}_k(\boldsymbol{x},\boldsymbol{X}) = 0 && k\in\mathcal{E}\\
&\text{(2.14d)} && && \boldsymbol{X}-\boldsymbol{x}\boldsymbol{x}^\top \succeq_{\mathcal{C}_r} 0 && \\
&\text{(2.14e)} && && \boldsymbol{e}_i^\top \boldsymbol{V}(\boldsymbol{x},\boldsymbol{X})\boldsymbol{e}_j \geq 0, && (i,j)\in\mathcal{V}
\end{aligned}$$

where $\mathcal{V}\subseteq\mathcal{H}\times\mathcal{H}$ is a selection of RLT inequalities, the additional conic constraint (2.14d) is a convex relaxation of the equation $\boldsymbol{X}=\boldsymbol{x}\boldsymbol{x}^\top$, and

$$\text{(2.15)} \qquad \mathcal{C}_r \triangleq \left\{\boldsymbol{Y} \;\middle|\; \boldsymbol{Y}\{\mathcal{K},\mathcal{K}\}\succeq 0,\ \ \forall\,\mathcal{K}\subseteq\{1,\ldots,n\}\wedge|\mathcal{K}|=r\right\}.$$

If $\mathcal{V}\neq\emptyset$, we refer to the convex problem (2.14a)-(2.14e) as the r-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 (2.14a)-(2.14e) has an optimal solution $(\overset{*}{\boldsymbol{x}},\overset{*}{\boldsymbol{X}})$ that satisfies $\overset{*}{\boldsymbol{X}}=\overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top$, then the relaxation is said to be exact and $\overset{*}{\boldsymbol{x}}$ 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.

2.3. Penalization. If the conic relaxation problem (2.14a)-(2.14e) is not exact, the resulting solution is not necessarily feasible for the original QCQP (2.1a)-(2.1c). In this case, an initial point $\hat{\boldsymbol{x}}\in\mathbb{R}^n$ (either feasible or infeasible) can be used to revise the objective function, resulting in a penalized conic programming relaxation problem of the form:

$$\begin{aligned}
&\text{(2.16a)} \qquad && \underset{\substack{\boldsymbol{x}\in\mathbb{R}^n\\ \boldsymbol{X}\in\mathbb{S}_n}}{\text{minimize}} && \bar{q}_0(\boldsymbol{x},\boldsymbol{X})+\eta(\operatorname{tr}\{\boldsymbol{X}\}-2\hat{\boldsymbol{x}}^\top\boldsymbol{x}+\hat{\boldsymbol{x}}^\top\hat{\boldsymbol{x}}) && \\
&\text{(2.16b)} && \text{subject to} && \bar{q}_k(\boldsymbol{x},\boldsymbol{X}) \leq 0 && k\in\mathcal{I}\\
&\text{(2.16c)} && && \bar{q}_k(\boldsymbol{x},\boldsymbol{X}) = 0 && k\in\mathcal{E}\\
&\text{(2.16d)} && && \boldsymbol{X}-\boldsymbol{x}\boldsymbol{x}^\top \succeq_{\mathcal{C}_r} 0, && \\
&\text{(2.16e)} && && \boldsymbol{e}_i^\top \boldsymbol{V}(\boldsymbol{x},\boldsymbol{X})\boldsymbol{e}_j \geq 0 && (i,j)\in\mathcal{V}
\end{aligned}$$

where $\eta>0$ is a fixed penalty parameter. The penalization is said to be *tight* if the problem (2.16a)-(2.16e) has a unique optimal solution $(\overset{*}{\boldsymbol{x}},\overset{*}{\boldsymbol{X}})$ that satisfies $\overset{*}{\boldsymbol{X}}=\overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top$.

The following theorem guarantees that if $\hat{\boldsymbol{x}}$ is feasible and satisfies the LICQ regularity condition, then the solution of (2.16a)-(2.16e) is guaranteed to be feasible for the QCQP (2.1a)-(2.1c) for an appropriate choice of η.

THEOREM 2.6. *Let $\hat{\boldsymbol{x}}$ be a feasible point for the QCQP* (2.1a)-(2.1b) *that satisfies the LICQ condition. If η is sufficiently large, then the convex problem* (2.16a)-(2.16e) *has a unique optimal solution $(\overset{*}{\boldsymbol{x}},\overset{*}{\boldsymbol{X}})$ such that $\overset{*}{\boldsymbol{X}}=\overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top$. Moreover, $\overset{*}{\boldsymbol{x}}$ is feasible for* (2.1a)-(2.1c) *and satisfies $q_0(\overset{*}{\boldsymbol{x}})\leq q_0(\hat{\boldsymbol{x}})$.*

Proof. Please refer to Section 4 for the proof. ∎

If $\hat{\boldsymbol{x}}$ 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 η is large. This will be explained next.

THEOREM 2.7. *Let $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ satisfy the GLICQ condition for the QCQP (2.1a)-(2.1b), and assume that*

$$d_{\mathcal{F}}(\hat{\boldsymbol{x}}) < \frac{s(\hat{\boldsymbol{x}})}{2\,(1+C_{n-1,r-1})\,\|\boldsymbol{P}\|_2}. \tag{2.17}$$

If η is sufficiently large, then the convex problem (2.16a)-(2.16e) has a unique optimal solution $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{X}})$ such that $\overset{*}{\boldsymbol{X}} = \overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top$ and $\overset{*}{\boldsymbol{x}}$ is feasible for (2.1a)-(2.1c).

Proof. Please refer to Section 4 for the proof. ∎

2.4. Sequential Penalization. If the initial point $\hat{\boldsymbol{x}}$ does not result in a tight penalization, the convex problem (2.16a)-(2.16e) 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 2.1. According to Theorem (2.7), once we are close to the feasible set $\mathcal{F}$, the relaxation becomes tight. Then, according to Theorem (2.6), feasibility is preserved and the objective value does not increase.

Algorithm 2.1 Sequential Penalized Conic Relaxation.

initiate $r \geq 2$, $\eta > 0$, $\hat{\boldsymbol{x}} \in \mathbb{R}^n$, $\{q_k\}_{k\in\{0\}\cup\mathcal{I}\cup\mathcal{E}}$
while stopping criterion is met **do**
  solve the convex problem (2.16a)-(2.16e) with the initial point $\hat{\boldsymbol{x}}$ to obtain $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{X}})$
  set $\hat{\boldsymbol{x}} := \overset{*}{\boldsymbol{x}}$
end while
return $\overset{*}{\boldsymbol{x}}$

3. 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

$$\begin{aligned}
&\text{(3.1a)} && \underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} && u_0(\boldsymbol{x}) && \\
&\text{(3.1b)} && \text{subject to} && u_k(\boldsymbol{x}) \leq 0 && k \in \mathcal{I} \\
&\text{(3.1c)} && && u_k(\boldsymbol{x}) = 0 && k \in \mathcal{E},
\end{aligned}$$

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 (3.1a)-(3.1c) can be reformulated as a QCQP of the form:

$$\begin{aligned}
&\text{(3.2a)} && \underset{\substack{\boldsymbol{x}\in\mathbb{R}^n\\ \boldsymbol{y}\in\mathbb{R}^o}}{\text{minimize}} && w_0(\boldsymbol{x},\boldsymbol{y}) && \\
&\text{(3.2b)} && \text{subject to} && w_k(\boldsymbol{x},\boldsymbol{y}) \leq 0 && k \in \mathcal{I} \\
&\text{(3.2c)} && && w_k(\boldsymbol{x},\boldsymbol{y}) = 0 && k \in \mathcal{E} \\
&\text{(3.2d)} && && v_i(\boldsymbol{x},\boldsymbol{y}) = 0 && i \in \mathcal{O}
\end{aligned}$$

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 $\boldsymbol{v}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{0}_n$, 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 (3.1a)-(3.1c) and (3.2a)-(3.2d). Moreover, a feasible point $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{y}})$ is an optimal solution to the QCQP (3.2a)-(3.2d) if and only if $\overset{*}{\boldsymbol{x}}$ is an optimal solution to the polynomial optimization problem (3.1a)-(3.1c). The next theorem provides an upper bound on the number of parameters and constraints that are sufficient to transform a general polynomial optimization to a QCQP.

THEOREM 3.1 ([51]). *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* (3.1a)-(3.1c) *in the form of* (3.2a)-(3.2d)*, where $|\mathcal{O}| \leq mn\left(\lfloor \log_2(d) \rfloor + 1\right)$.*

Proof. Please refer to [51] for the proof. □

The next theorem 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 (3.2a)-(3.2d).

THEOREM 3.2. *Consider a pair of vectors $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ and $\hat{\boldsymbol{y}} \in \mathbb{R}^{|\mathcal{O}|}$ satisfying $\boldsymbol{v}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) = \boldsymbol{0}_n$. The following two statements are equivalent:*

1. *$\hat{\boldsymbol{x}}$ is feasible and satisfies the LICQ condition for the polynomial optimization problem* (3.1a)-(3.1b).
2. *$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ is feasible and satisfies the LICQ condition for the QCQP* (3.2a)-(3.2d).

Proof. Please refer to Section 4 for the proof. □

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

Example 3.3. Consider the following three-dimensional polynomial optimization problem:

$$\underset{\boldsymbol{x}\in\mathbb{R}^3}{\text{minimize}} \quad x_1^3 + x_2^2 + 3x_1x_2x_3 \tag{3.3a}$$
$$\text{subject to} \quad x_1^2 - 1 \leq 0 \tag{3.3b}$$
$$x_3^2 - 1 = 0. \tag{3.3c}$$

To derive a QCQP reformulation of the problem (3.3a)-(3.3c), we introduce the auxiliary variables y_1 and y_2, accounting for the monomials x_1^2 and x_1x_2, respectively. Define

$$w_0(\boldsymbol{x}, \boldsymbol{y}) \triangleq x_1y_1 + x_2^2 + 3y_2x_3 \tag{3.4a}$$
$$w_1(\boldsymbol{x}, \boldsymbol{y}) \triangleq y_1 - 1, \qquad v_1(\boldsymbol{x}, \boldsymbol{y}) \triangleq y_1 - x_1^2 \tag{3.4b}$$
$$w_2(\boldsymbol{x}, \boldsymbol{y}) \triangleq x_3^2 - 1, \qquad v_2(\boldsymbol{x}, \boldsymbol{y}) \triangleq y_2 - x_1x_2, \tag{3.4c}$$

A QCQP reformulation (3.2a)-(3.2d) is as follows:

$$\underset{\substack{\boldsymbol{x}\in\mathbb{R}^3\\ \boldsymbol{y}\in\mathbb{R}^2}}{\text{minimize}} \quad x_1y_1 + x_2^2 + 3y_2x_3 \tag{3.5a}$$
$$\text{subject to} \quad y_1 - 1 \leq 0 \tag{3.5b}$$
$$x_3^2 - 1 = 0 \tag{3.5c}$$
$$y_1 - x_1^2 = 0 \tag{3.5d}$$
$$y_2 - x_1x_2 = 0. \tag{3.5e}$$

Observe that every feasible point $(x_1, x_2, x_3) \in \mathbb{R}^3$ of the problem (3.3a)-(3.3c) can be mapped into a feasible point $(x_1, x_2, x_3, x_1^2, x_1x_2) \in \mathbb{R}^5$ of (3.5a)-(3.5e) and vice versa.

Additionally, the constraint Jacobian matrices for the two problems are given as:

$$\mathcal{J}_{\mathrm{PO}}(\boldsymbol{x}) = \begin{bmatrix} 2x_1 & 0 & 0 \\ 0 & 0 & 2x_3 \end{bmatrix} \tag{3.6a}$$

$$\mathcal{J}_{\mathrm{QCQP}}(\boldsymbol{x},\boldsymbol{y}) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2x_3 & 0 & 0 \\ -2x_1 & 0 & 0 & 1 & 0 \\ -x_1 & -x_2 & 0 & 0 & 1 \end{bmatrix}. \tag{3.6b}$$

It is straightforward to verify that $\mathcal{J}_{\mathrm{PO}}$ has full row rank if and only if $\mathcal{J}_{\mathrm{QCQP}}(\boldsymbol{x},\boldsymbol{y})$ has full row rank.

In light of Theorems 3.1 and 3.2, with no loss of generality, the proposed penalization method can be applied to the class of polynomial optimization problems as well.

4. Proofs.

Proof of Theorem 3.2. Due to the equality $\boldsymbol{u}(\hat{\boldsymbol{x}}) = \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})$ as well as the invertiblity assumption for $\boldsymbol{v}(\hat{\boldsymbol{x}},\cdot)$, we have

$$\begin{aligned} \frac{\partial \boldsymbol{u}(\hat{\boldsymbol{x}})}{\partial \boldsymbol{x}} &= \begin{bmatrix} \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & -\left(\frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}}\right)^{-1} \frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} \end{bmatrix}^{\top} \\ &= \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} - \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}} \left(\frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}}\right)^{-1} \frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}}. \end{aligned} \tag{4.1}$$

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

$$\mathcal{J}_{\mathrm{QCQP}}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}}) = \begin{bmatrix} \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{w}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}} \\ \frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{v}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})}{\partial \boldsymbol{y}} \end{bmatrix}, \tag{4.2}$$

which is the Jacobian matrix of the QCQP (3.2a)-(3.2d) 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. □

In order to prove Theorems 2.6 and 2.7, it is useful to consider the following optimization problem:

$$\begin{aligned} &\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} && q_0(\boldsymbol{x}) + \eta\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|_2^2 && && (4.3a)\\ &\text{subject to} && q_k(\boldsymbol{x}) \le 0 && k \in \mathcal{I} && (4.3b)\\ & && q_k(\boldsymbol{x}) = 0 && k \in \mathcal{E} && (4.3c) \end{aligned}$$

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

$$|q_0(\boldsymbol{x})| \le \alpha\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|_2^2 + \alpha, \tag{4.4}$$

is satisfied for every $\boldsymbol{x} \in \mathbb{R}^n$. If $\eta > \alpha$, then the objective function (4.3a) is lower bounded by $-\alpha$ and its optimal value is attainable within any closed and nonempty subset of $\mathbb{R}^n$.

LEMMA 4.1. *Given an arbitrary $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ and $\varepsilon > 0$, every optimal solution $\overset{*}{\boldsymbol{x}}$ of the problem* (4.3a)-(4.3c) *satisfies*

$$0 \le \|\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2 - d_{\mathcal{F}}(\hat{\boldsymbol{x}}) \le \varepsilon \tag{4.5}$$

if η is sufficiently large.

Proof. Consider an optimal solution $\overset{*}{\boldsymbol{x}}$. Due to Definition 2.3, 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 (4.5). 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

$$q_0(\overset{*}{\boldsymbol{x}}) + \eta\|\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2^2 \leq q_0(\boldsymbol{x}_d) + \eta\|\boldsymbol{x}_d - \hat{\boldsymbol{x}}\|_2^2. \tag{4.6}$$

According to the inequalities (4.6) and (4.4), one can write

$$(\eta - \alpha)\|\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2^2 - \alpha \leq (\eta + \alpha)\|\boldsymbol{x}_d - \hat{\boldsymbol{x}}\|_2^2 + \alpha \tag{4.7a}$$

$$\Rightarrow \ \|\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2^2 \leq \|\boldsymbol{x}_d - \hat{\boldsymbol{x}}\|_2^2 + \frac{2\alpha}{\eta - \alpha}(1 + \|\boldsymbol{x}_d - \hat{\boldsymbol{x}}\|_2^2) \tag{4.7b}$$

$$\Rightarrow \ \|\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2^2 \leq d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2 + \frac{2\alpha}{\eta - \alpha}(1 + d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2), \tag{4.7c}$$

which concludes the right side of (4.5), provided that $\eta \geq \alpha + 2\alpha(1 + d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2)[\varepsilon^2 + 2\varepsilon d_{\mathcal{F}}(\hat{\boldsymbol{x}})]^{-1}$. ∎

LEMMA 4.2. *Assume that $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ satisfies the GLICQ condition for the problem* (4.3a)-(4.3c)*. Given an arbitrary $\varepsilon > 0$, every optimal solution $\overset{*}{\boldsymbol{x}}$ of the problem satisfies*

$$s(\hat{\boldsymbol{x}}) - s(\overset{*}{\boldsymbol{x}}) \leq 2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2 + \varepsilon, \tag{4.8}$$

if η is sufficiently large.

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 2.4). Due to Lemma 4.1, for every $k \in \mathcal{I} \setminus \hat{\mathcal{B}}$ and every arbitrary $\varepsilon_1 > 0$, we have

$$\begin{aligned} q_k(\overset{*}{\boldsymbol{x}}) - q_k(\hat{\boldsymbol{x}}) &= 2(\boldsymbol{A}_k\hat{\boldsymbol{x}} + \boldsymbol{b}_k)^\top(\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + (\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}})^\top\boldsymbol{A}_k(\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}) \\ &\leq \|\nabla q_k(\hat{\boldsymbol{x}})\|_2\|\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2 + \|\boldsymbol{A}_k\|_2\|\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2^2 \\ &\leq \|\nabla q_k(\hat{\boldsymbol{x}})\|_2 d_{\mathcal{F}}(\hat{\boldsymbol{x}}) + \|\boldsymbol{A}_k\|_2 d_{\mathcal{F}}(\hat{\boldsymbol{x}})^2 + \varepsilon_1 < -q_k(\hat{\boldsymbol{x}}), \end{aligned} \tag{4.9}$$

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

$$s(\overset{*}{\boldsymbol{x}}) = \sigma_{\min}\{\mathcal{J}_{\overset{*}{\mathcal{B}}}(\overset{*}{\boldsymbol{x}})\} \geq \sigma_{\min}\{\mathcal{J}_{\hat{\mathcal{B}}}(\overset{*}{\boldsymbol{x}})\} = \|\mathcal{J}_{\hat{\mathcal{B}}}(\overset{*}{\boldsymbol{x}})^\top\boldsymbol{\nu}\|_2 \tag{4.10a}$$

$$\geq \|\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}})^\top\boldsymbol{\nu}\|_2 - \|[\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}}) - \mathcal{J}_{\hat{\mathcal{B}}}(\overset{*}{\boldsymbol{x}})]^\top\boldsymbol{\nu}\|_2 \tag{4.10b}$$

$$\geq \sigma_{\min}\{\mathcal{J}_{\hat{\mathcal{B}}}(\hat{\boldsymbol{x}})\}\|\boldsymbol{\nu}\|_2 - 2\|\boldsymbol{P}\|_2\|\hat{\boldsymbol{x}} - \overset{*}{\boldsymbol{x}}\|_2\|\boldsymbol{\nu}\|_2 \tag{4.10c}$$

$$\geq s(\hat{\boldsymbol{x}}) - 2\|\boldsymbol{P}\|_2\|\hat{\boldsymbol{x}} - \overset{*}{\boldsymbol{x}}\|_2 \tag{4.10d}$$

$$\geq s(\hat{\boldsymbol{x}}) - 2d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2 - \varepsilon, \tag{4.10e}$$

if η is large, which concludes the inequality (4.8). ∎

LEMMA 4.3. *Let $\overset{*}{\boldsymbol{x}}$ be an optimal solution of the problem* (4.3a)-(4.3c)*, and assume that $\overset{*}{\boldsymbol{x}}$ is LICQ regular. There exists a pair of dual vectors $(\overset{*}{\boldsymbol{\gamma}}, \overset{*}{\boldsymbol{\mu}}) \in \mathbb{R}_+^{|\mathcal{I}|} \times \mathbb{R}^{|\mathcal{E}|}$ that satisfies the following Karush-Kuhn-Tucker (KKT) conditions:*

$$2(\eta\boldsymbol{I} + \boldsymbol{A}_0)(\overset{*}{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + 2(\boldsymbol{A}_0\hat{\boldsymbol{x}} + \boldsymbol{b}_0) + \mathcal{J}(\overset{*}{\boldsymbol{x}})^\top[\overset{*}{\boldsymbol{\gamma}}^\top, \overset{*}{\boldsymbol{\mu}}^\top]^\top = 0, \tag{4.11a}$$

$$\overset{*}{\gamma}_k q_k(\overset{*}{\boldsymbol{x}}) = 0, \qquad \forall k \in \mathcal{I}. \tag{4.11b}$$

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

$$
\begin{aligned}
0 &= \nabla_{\boldsymbol{x}} \mathcal{L}(\mathring{\boldsymbol{x}}, \mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}})/2 \\
&= \eta(\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + (\boldsymbol{A}_0 \mathring{\boldsymbol{x}} + \boldsymbol{b}_0) + \boldsymbol{P}(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}})\mathring{\boldsymbol{x}} + \sum_{k \in \mathcal{I}} \mathring{\gamma}_k \boldsymbol{b}_k + \sum_{k \in \mathcal{E}} \mathring{\mu}_k \boldsymbol{b}_k \\
&= (\eta \boldsymbol{I} + \boldsymbol{A}_0)(\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + (\boldsymbol{A}_0 \hat{\boldsymbol{x}} + \boldsymbol{b}_0) + \mathcal{J}(\mathring{\boldsymbol{x}})^{\top} [\mathring{\boldsymbol{\gamma}}^{\top}, \mathring{\boldsymbol{\mu}}^{\top}]^{\top}/2. \qquad (4.12)
\end{aligned}
$$

Moreover, (4.11b) is concluded from the complementary slackness. ∎

LEMMA 4.4. *Consider an arbitrary $\varepsilon > 0$ and assume that $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ satisfies the inequality*

$$
s(\hat{\boldsymbol{x}}) > 2 d_{\mathcal{F}}(\hat{\boldsymbol{x}}) \|\boldsymbol{P}\|_2. \tag{4.13}
$$

If η is sufficiently large, for every optimal solution $\mathring{\boldsymbol{x}}$ of the problem (4.3a)-(4.3c)*, 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*

$$
\frac{1}{\eta} \sqrt{\|\mathring{\boldsymbol{\gamma}}\|_2^2 + \|\mathring{\boldsymbol{\mu}}\|_2^2} \leq \frac{2 d_{\mathcal{F}}(\hat{\boldsymbol{x}})}{s(\hat{\boldsymbol{x}}) - 2 d_{\mathcal{F}}(\hat{\boldsymbol{x}}) \|\boldsymbol{P}\|_2} + \varepsilon \tag{4.14}
$$

as well as the equations (4.11a) *and* (4.11b)*.*

Proof. Due to Lemma 4.3, there exists $(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}}) \in \mathbb{R}_{+}^{|\mathcal{I}|} \times \mathbb{R}^{|\mathcal{E}|}$ that satisfies the equations (4.11a) and (4.11b). 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 (4.11a) and (4.11b), one can write

$$
2(\eta \boldsymbol{I} + \boldsymbol{A}_0)(\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + 2(\boldsymbol{A}_0 \hat{\boldsymbol{x}} + \boldsymbol{b}_0) + \mathcal{J}_{\mathring{\mathcal{B}}}(\mathring{\boldsymbol{x}})^{\top} \boldsymbol{\tau}\{\mathring{\mathcal{B}}\} = 0. \tag{4.15}
$$

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

$$
\varepsilon_1 \triangleq \phi \times \frac{\varepsilon - 2\eta^{-1}\phi^{-1}(\|\boldsymbol{A}_0 \hat{\boldsymbol{x}} + \boldsymbol{b}_0\|_2 + d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{A}_0\|_2)}{\varepsilon + 2 + 2\eta^{-1}\|\boldsymbol{A}_0\|_2 + 2\phi^{-1} d_{\mathcal{F}}(\hat{\boldsymbol{x}})}. \tag{4.16}
$$

If η is sufficiently large, ε_1 is positive and based on Lemmas 4.1 and 4.2, we have

$$
\begin{aligned}
\frac{\|\boldsymbol{\tau}\|_2}{\eta} = \frac{\|\boldsymbol{\tau}\{\mathring{\mathcal{B}}\}\|_2}{\eta} &\leq \frac{2\|(\eta \boldsymbol{I} + \boldsymbol{A}_0)(\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}) + (\boldsymbol{A}_0 \hat{\boldsymbol{x}} + \boldsymbol{b}_0)\|_2}{\eta \sigma_{\min}\{\mathcal{J}_{\mathring{\mathcal{B}}}(\mathring{\boldsymbol{x}})\}} \\
&\leq \frac{2\eta\|\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2 + 2\|\boldsymbol{A}_0\|_2 \|\mathring{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|_2 + 2\|\boldsymbol{A}_0 \hat{\boldsymbol{x}} + \boldsymbol{b}_0\|_2}{\eta s(\mathring{\boldsymbol{x}})} \\
&\leq \frac{2(d_{\mathcal{F}}(\hat{\boldsymbol{x}}) + \varepsilon_1) + 2\eta^{-1}[\|\boldsymbol{A}_0\|_2(d_{\mathcal{F}}(\hat{\boldsymbol{x}}) + \varepsilon_1) + \|\boldsymbol{A}_0 \hat{\boldsymbol{x}} + \boldsymbol{b}_0\|_2]}{s(\hat{\boldsymbol{x}}) - 2 d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2 - \varepsilon_1} \\
&= \frac{2 d_{\mathcal{F}}(\hat{\boldsymbol{x}})}{s(\hat{\boldsymbol{x}}) - 2 d_{\mathcal{F}}(\hat{\boldsymbol{x}})\|\boldsymbol{P}\|_2} + \varepsilon, \qquad (4.17)
\end{aligned}
$$

where the last equality is a result of the equation (4.16). ∎

LEMMA 4.5. *Consider an optimal solution $\mathring{\boldsymbol{x}}$ of the problem* (4.3a)-(4.3c)*, 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* (4.11a) *and* (4.11b)*. If the matrix inequality*

$$
\eta \boldsymbol{I} + \boldsymbol{A}_0 + \boldsymbol{P}(\mathring{\boldsymbol{\gamma}}, \mathring{\boldsymbol{\mu}}) \succ_{\mathcal{D}_r} 0, \tag{4.18}
$$

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 (2.16a)-(2.16e)*.*

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 (2.16d). Then, the KKT conditions for the problem (2.16a)-(2.16e) can be written as follows:

$$\nabla_{\boldsymbol{x}}\ \bar{\mathcal{L}}(\boldsymbol{x}, \boldsymbol{X}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = 2\left(\boldsymbol{\Lambda}\boldsymbol{x} - \eta\hat{\boldsymbol{x}} + \boldsymbol{b}_0 + \sum_{k\in\mathcal{I}} \overset{*}{\gamma}_k \boldsymbol{b}_k + \sum_{k\in\mathcal{E}} \overset{*}{\mu}_k \boldsymbol{b}_k\right) = 0, \tag{4.19a}$$

$$\nabla_{\boldsymbol{X}}\ \bar{\mathcal{L}}(\boldsymbol{x}, \boldsymbol{X}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \eta\boldsymbol{I} + \boldsymbol{A}_0 + \boldsymbol{P}(\boldsymbol{\gamma}, \boldsymbol{\mu}) - \boldsymbol{\Lambda} = 0, \tag{4.19b}$$

$$\gamma_k q_k(\boldsymbol{x}) = 0, \qquad \forall k \in \mathcal{I} \tag{4.19c}$$

$$\langle \boldsymbol{\Lambda},\ \boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{X}\rangle = 0, \tag{4.19d}$$

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 (4.19a) and (4.19b) account for stationarity with respect to $\boldsymbol{x}$ and $\boldsymbol{X}$, respectively, and equations (4.19c) and (4.19d) are the complementary slackness conditions for the constraints (2.16b) and (2.16d), respectively. Define

$$\overset{*}{\boldsymbol{\Lambda}} \triangleq \eta\boldsymbol{I} + \boldsymbol{A}_0 + \boldsymbol{P}(\overset{*}{\boldsymbol{\gamma}}, \overset{*}{\boldsymbol{\mu}}). \tag{4.20}$$

Due to Lemma (4.3), if η is sufficiently large, $\overset{*}{\boldsymbol{x}}$ and $(\overset{*}{\boldsymbol{\gamma}}, \overset{*}{\boldsymbol{\mu}})$ satisfy the equations (4.11a) and (4.11b), which yield the optimality conditions (4.19a)-(4.19d), 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 (2.16a)-(2.16e).

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

$$\overset{*}{\boldsymbol{x}} = \overset{*}{\boldsymbol{\Lambda}}^{-1}\left(\eta\hat{\boldsymbol{x}} - \boldsymbol{b}_0 - \sum_{k\in\mathcal{I}} \overset{*}{\gamma}_k \boldsymbol{b}_k - \sum_{k\in\mathcal{E}} \overset{*}{\mu}_k \boldsymbol{b}_k\right) \tag{4.21}$$

and $\overset{*}{\boldsymbol{X}} = \overset{*}{\boldsymbol{x}}\overset{*}{\boldsymbol{x}}^\top$, according to the equations (4.19a) and (4.19d), respectively, which implies the uniqueness of the solution. ◻

LEMMA 4.6. *Consider an optimal solution $\overset{*}{\boldsymbol{x}}$ of the problem* (4.3a)-(4.3c)*, 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* (4.11a) *and* (4.11b)*. If the inequality,*

$$\frac{1}{\eta}\sqrt{\|\overset{*}{\boldsymbol{\gamma}}\|_2^2 + \|\overset{*}{\boldsymbol{\mu}}\|_2^2} < \frac{1}{C_{n-1,r-1}\|\boldsymbol{P}\|_2} - \frac{\|\boldsymbol{A}_0\|_2}{\eta\|\boldsymbol{P}\|_2} \tag{4.22}$$

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 (2.16a)-(2.16e)*.*

Proof. Based on Lemma 4.5, it suffices to prove the conic inequality (4.18). Define

$$\boldsymbol{K} \triangleq \boldsymbol{A}_0 + \boldsymbol{P}(\overset{*}{\boldsymbol{\gamma}}, \overset{*}{\boldsymbol{\mu}}). \tag{4.23}$$

It follows that

$$\|\boldsymbol{K}\|_2 \le \|\boldsymbol{A}_0\|_2 + \sum_{k\in\mathcal{I}} \overset{*}{\gamma}_k \|\boldsymbol{A}_k\|_2 + \sum_{k\in\mathcal{E}} \overset{*}{\mu}_k \|\boldsymbol{A}_k\|_2, \tag{4.24a}$$

$$\le \|\boldsymbol{A}_0\|_2 + \|\boldsymbol{P}\|_2\sqrt{\|\overset{*}{\boldsymbol{\gamma}}\|_2^2 + \|\overset{*}{\boldsymbol{\mu}}\|_2^2}\ . \tag{4.24b}$$

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

$$\eta \boldsymbol{I} + \boldsymbol{K} = \sum_{\mathcal{K} \in \mathcal{R}} \boldsymbol{I}\{\mathcal{K}\}^\top \boldsymbol{R}_{\mathcal{K}} \boldsymbol{I}\{\mathcal{K}\}, \tag{4.25}$$

where

$$\boldsymbol{R}_{\mathcal{K}} = \binom{n-1}{r-1}^{-1} [\eta \boldsymbol{I}\{\mathcal{K}, \mathcal{K}\} + \boldsymbol{K}\{\mathcal{K}, \mathcal{K}\}]. \tag{4.26}$$

Due to the inequalities (4.22) and (4.24), 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 2.7. Let $\overset{*}{\boldsymbol{x}}$ be an optimal solution of the problem (4.3a)-(4.3c). According to the assumption (2.17), the inequality (4.13) holds true, and due to Lemma 4.4, if η is sufficiently large, there exists a corresponding pair of dual vectors $(\overset{*}{\boldsymbol{\gamma}}, \overset{*}{\boldsymbol{\mu}})$ that satisfies the inequality (4.14). Now, according to the inequality (2.17), we have

$$\frac{2 d_{\mathcal{F}}(\hat{\boldsymbol{x}})}{s(\hat{\boldsymbol{x}}) - 2 d_{\mathcal{F}}(\hat{\boldsymbol{x}}) \|\boldsymbol{P}\|_2} \leq \frac{1}{C_{n-1,r-1} \|\boldsymbol{P}\|_2} \tag{4.27}$$

and therefore (4.14) concludes (4.22). Hence, according to Lemma 4.6, the pair $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{x}} \overset{*}{\boldsymbol{x}}^\top)$ is the unique primal solution to the penalized convex relaxation problem (2.16a)-(2.16e). ∎

Proof of Theorem 2.6. If $\hat{\boldsymbol{x}}$ is feasible, then $d_{\mathcal{F}}(\hat{\boldsymbol{x}}) = 0$. Therefore, the tightness of the penalization for Theorem 2.6 is a direct consequence of Theorem 2.7. Denote the unique optimal solution of the penalized relaxation as $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{x}} \overset{*}{\boldsymbol{x}}^\top)$. Then it is straightforward to verify the inequality $q_0(\overset{*}{\boldsymbol{x}}) \leq q_0(\hat{\boldsymbol{x}})$ by evaluating the objective function (2.16a) at the point $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}} \hat{\boldsymbol{x}}^\top)$. ∎

5. Numerical Experiments on QPLIB Cases.

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) [28]. The experiments are all performed on a desktop computer with a 12-core 3.0GHz CPU and 256GB RAM. MOSEK v8.1 [5] is used through MATLAB 2017a to solve the resulting convex relaxations.

5.1. Convex Relaxation without Penalization.

Table 1 reports the outcome of the unpenalized convex relaxation problem (2.14a)-(2.14e) with the following settings:

- $\mathcal{V} = \emptyset$ and $r = n$ (SDP relaxation),
- $\mathcal{V} = \emptyset$ and $r = 2$ (SOCP relaxation),
- $\mathcal{V} = \mathcal{H} \times \mathcal{H}$ and no conic inequality (RLT relaxation),
- $\mathcal{V} = \mathcal{H} \times \mathcal{H}$ and $r = n$ (SDP+RLT relaxation),
- $\mathcal{V} = \mathcal{H} \times \mathcal{H}$ 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 $\mathrm{LB} \triangleq q_0(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{X}})$, and

$$\mathrm{FV} \triangleq \mathrm{tr}\{\overset{*}{\boldsymbol{X}} - \overset{*}{\boldsymbol{x}} \overset{*}{\boldsymbol{x}}^\top\} \tag{5.1a}$$

$$\mathrm{GAP}(\%) \triangleq 100 \times \frac{q_0(\boldsymbol{x}^{\mathrm{QPLIB}}) - q_0(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{X}})}{q_0(\boldsymbol{x}^{\mathrm{QPLIB}})}, \tag{5.1b}$$

Table 1: Unpenalized convex relaxations.

Inst	RLT				SOCP+RLT				SDP+RLT				BARON			COUENNE		
	t(s)	FV	LB	GAP(%)	t(s)	FV	LB	GAP(%)	t(s)	FV	LB	GAP(%)	t(s)	LB	GAP(%)	t(s)	LB	GAP(%)
0343	0.23	-5e-1	-9.655	51.20	0.70	9e-2	-7.269	13.83	0.67	6e-7	-6.386	0.00	$\leq$0.64	-6.386	0.00	2.97	-6.386	0.00
0911	0.13	–	$-\infty$	–	0.22	–	$-\infty$	–	0.14	–	$-\infty$	–	40.09	-32.147	0.00	74.49	-32.147	0.00
0975	0.05	–	$-\infty$	–	0.80	–	$-\infty$	–	0.11	–	$-\infty$	–	$\leq$0.94	-37.791	0.15	52.78	-37.853	0.00
1055	0.13	–	$-\infty$	–	0.16	–	$-\infty$	–	0.11	–	$-\infty$	–	2.42	-33.037	0.00	$\leq$3.49	-33.037	0.00
1143	0.53	+1e-1	-114.706	100.37	0.70	9e+0	-109.302	90.93	0.61	5e+0	-58.908	2.90	$\leq$0.82	-57.246	0.00	56.13	-57.246	0.00
1157	0.52	-3e+0	-11.771	7.52	0.67	2e-6	-10.948	0.00	0.52	6e-8	-10.948	0.00	$\leq$0.32	-10.948	0.00	$\leq$1.49	-10.948	0.00
1353	0.64	-7e+0	-13.591	76.18	1.13	8e+0	-10.256	32.96	1.25	5e-5	-7.714	0.00	$\leq$0.49	-7.714	0.00	$\leq$3.15	-7.714	0.00
1423	0.42	+6e+0	-22.609	51.06	0.77	8e+0	-22.462	50.07	0.64	3e+0	-15.154	1.25	$\leq$0.91	-14.967	0.00	47.03	-14.929	0.25
1437	0.91	-4e-1	-10.141	30.19	1.23	6e+0	-9.268	18.99	1.55	2e+0	-7.794	0.07	$\leq$0.39	-7.789	0.00	$\leq$3.04	-7.789	0.00
1451	1.41	+1e+0	-189.539	116.43	1.81	1e+1	-185.434	111.74	2.53	9e+0	-94.346	7.73	$\leq$10.93	-87.576	0.00	70.20	-87.283	0.33
1493	0.41	+1e+0	-65.591	51.97	0.56	6e+0	-61.053	41.46	0.48	3e+0	-43.883	1.67	$\leq$0.56	-43.160	0.00	1.32	-43.160	0.00
1507	0.20	-5e+0	-15.807	90.41	0.33	6e+0	-11.862	42.89	0.23	1e-4	-8.301	0.00	$\leq$0.42	-8.301	0.00	$\leq$1.61	-8.301	0.00
1535	1.30	+7e+0	-21.561	86.10	2.77	1e+1	-21.064	81.81	3.08	7e+0	-12.203	5.32	$\leq$14.97	-11.454	1.14	141.56	-11.403	1.57
1619	0.70	+3e+0	-18.058	95.92	1.24	1e+1	-17.163	86.20	1.83	3e-4	-9.216	0.01	$\leq$4.02	-9.217	0.00	$\leq$4.02	-9.217	0.00
1661	2.42	-3e+0	-21.208	32.92	2.84	9e+0	-19.439	21.84	3.25	3e+0	-16.028	0.46	$\leq$0.84	-15.954	0.00	5.87	-15.954	0.00
1675	1.92	+1e+0	-127.977	69.13	2.38	1e+1	-121.753	60.90	3.08	5e+0	-76.342	0.89	$\leq$0.30	-75.668	0.00	4.12	-75.668	0.00
1703	1.36	+3e+0	-257.014	93.53	2.81	1e+1	-250.703	88.78	3.36	6e+0	-136.110	2.49	905.55	-132.802	0.00	29.60	-132.738	0.04
1745	0.70	+1e+0	-94.297	30.29	1.16	9e+0	-92.924	28.39	1.69	4e+0	-73.773	1.93	6.14	-72.376	0.00	$\leq$6.26	-72.376	0.00
1773	0.98	+2e-1	-31.049	112.06	2.13	1e+1	-29.962	104.63	2.25	7e+0	-15.490	5.79	$\leq$0.69	-14.641	0.00	$\leq$46.97	-14.641	0.00
1886	0.14	–	$-\infty$	–	0.20	–	$-\infty$	–	0.14	–	$-\infty$	–	$\leq$6.20	-78.671	0.00	$\leq$6.20	-78.671	0.00
1913	0.11	–	$-\infty$	–	0.20	–	$-\infty$	–	0.13	–	$-\infty$	–	3.67	-52.108	0.00	875.42	-51.954	0.29
1922	0.13	–	$-\infty$	–	0.14	–	$-\infty$	–	0.11	–	$-\infty$	–	1.55	-35.916	0.09	2.07	-35.950	0.00
1931	0.16	–	$-\infty$	–	0.22	–	$-\infty$	–	0.13	–	$-\infty$	–	$\leq$1.11	-55.708	0.00	398.95	-55.708	0.00
1967	0.20	–	$-\infty$	–	0.27	–	$-\infty$	–	0.16	–	$-\infty$	–	15.70	-107.373	0.19	1.51	-107.581	0.00

where $(\overset{*}{\boldsymbol{x}}, \overset{*}{\boldsymbol{X}})$ and $\boldsymbol{x}^{\mathrm{QPLIB}}$ denote solutions from convex relaxation and QPLIB, respectively.

For all of the cases except 1507, the SDP and SOCP relaxations are unbounded in the absence of RLT cuts. The SOCP relaxation of the case 1507 is unbounded as well, while the lower bound produced by the SDP relaxation is equal to -79.600. The cases 0911, 0975, 1055, 1886, 1913, 1922, 1931, and 1967 do not have affine constraints (i.e., $\mathcal{H} = \emptyset$) and therefore, imposing RLT cuts do not make any difference for these cases. There is no major difference between the RLT 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.

5.2. 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 1 using penalization and without high-quality choices for the initial points (flat start). Tables 2, 3, and 4 report the results of Algorithm 2.1 for SOCP, SDP and SDP+RLT relaxations, respectively. For this experiment, we use the given lower and upper bound vectors $\boldsymbol{x}^{\mathrm{lb}} \in (\{-\infty\} \cup \mathbb{R})^n$ and $\boldsymbol{x}^{\mathrm{ub}} \in (\mathbb{R} \cup \{+\infty\})^n$ to create the initial point $\boldsymbol{x}^{\mathrm{flat}}$ for Algorithm 2.1, where

$$(5.2) \qquad x_k^{\mathrm{flat}} = \begin{cases} (x_k^{\mathrm{lb}} + x_k^{\mathrm{ub}})/2 & \text{if} \quad x_k^{\mathrm{lb}} > -\infty \wedge x_k^{\mathrm{ub}} < \infty \\ x_k^{\mathrm{lb}} & \text{if} \quad x_k^{\mathrm{lb}} > -\infty \wedge x_k^{\mathrm{ub}} = \infty \\ x_k^{\mathrm{ub}} & \text{if} \quad x_k^{\mathrm{lb}} = -\infty \wedge x_k^{\mathrm{ub}} < \infty \\ 0 & \text{if} \quad x_k^{\mathrm{lb}} = -\infty \wedge x_k^{\mathrm{ub}} = \infty \end{cases},$$

for every $k \in \{1, \ldots, n\}$. 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 three rounds, where $\alpha \in \{1, 2, 5\}$ and β is an integer. In all of the experiments, the value of η has remained

Table 2: Sequential SOCP with flat start.

Inst	Sequential SOCP with flat start							BARON			COUENNE		
	η	i^{tight}	q_0^{tight}	i^{stop}	q_0^{stop}	GAP(%)	t(s)	LB	UB	GAP(%)	LB	UP	GAP(%)
0343	5e+2	1	-1.730	55	-6.382	0.07	8.62	-94.841	-6.386	0.00	-146.550	-6.386	0.00
0911	1e+1	1	-1.110	17	-30.948	3.73	4.24	-172.777	-28.206	12.26	-76.385	-30.967	3.67
0975	1e+1	1	-4.671	15	-36.424	3.78	16.83	-47.008	-37.794	0.16	-76.476	-36.812	2.75
1055	2e+1	1	4.127	24	-32.769	0.81	3.49	-37.759	-33.037	0.00	-92.544	-33.037	0.00
1143	2e+1	1	-3.471	24	-55.571	2.93	4.73	-69.745	-57.247	0.00	-138.560	-53.366	6.78
1157	5e+0	1	-5.066	12	-10.943	0.05	4.40	-12.165	-10.948	0.00	-13.748	-10.948	0.00
1353	5e+0	1	1.025	23	-7.714	0.00	15.57	-8.571	-7.714	0.00	-14.980	-7.714	0.00
1423	5e+0	1	-4.561	17	-14.682	1.91	2.56	-16.631	-14.967	0.00	-25.662	-14.059	6.07
1437	5e+0	1	-0.323	19	-7.779	0.13	13.57	-9.622	-7.789	0.00	-15.251	-7.789	0.00
1451	2e+1	3	-52.124	24	-87.332	0.28	10.93	-466.322	-87.576	0.00	-190.510	-86.859	0.82
1493	2e+1	1	-5.376	12	-41.807	3.14	3.77	-49.673	-43.160	0.00	-84.615	-43.160	0.00
1507	5e+0	1	-2.841	17	-8.295	0.08	1.61	-9.224	-8.301	0.00	-16.210	-8.301	0.00
1535	5e+0	2	-2.465	39	-11.516	0.61	17.94	-65.995	-11.454	1.14	-26.499	-11.398	1.62
1619	5e+0	1	6.199	19	-9.211	0.06	4.81	-72.315	-9.217	0.00	-22.310	-9.217	0.00
1661	5e+0	1	-6.084	32	-15.954	0.01	44.95	-18.719	-15.955	0.00	-29.410	-15.954	0.01
1675	2e+1	1	-8.886	14	-75.518	0.20	9.74	-99.009	-75.669	0.00	-147.830	-75.668	0.00
1703	5e+1	2	-33.359	26	-132.229	0.43	9.73	-926.764	-132.381	0.32	$-\infty$	–	–
1745	2e+1	1	-18.929	25	-72.360	0.02	6.85	-80.418	-72.377	0.00	-96.439	-72.376	0.00
1773	5e+0	1	-1.856	32	-13.916	4.95	46.97	-21.705	-14.642	0.00	-36.391	-14.641	0.01
1886	2e+1	1	-7.982	18	-78.610	0.08	6.20	-135.615	-78.672	0.00	-155.740	-78.671	0.00
1913	1e+1	1	-10.180	18	-51.871	0.46	6.02	-65.850	-52.108	0.00	-75.555	-51.348	1.46
1922	1e+1	1	-10.284	19	-35.725	0.63	3.59	-47.245	-35.916	0.10	-57.575	-35.950	0.00
1931	2e+1	1	-5.215	18	-54.290	2.55	3.72	-82.922	-55.709	0.00	-97.100	-54.290	2.55
1967	5e+1	1	-1.650	26	-106.826	0.70	8.99	-621.561	-106.91	0.62	-301.690	-107.580	0.00

static throughout Algorithm 2.1. Denote the sequence of penalized relaxation solutions obtained by Algorithm 2.1 as

$$(\boldsymbol{x}^{(1)}, \boldsymbol{X}^{(1)}), \quad (\boldsymbol{x}^{(2)}, \boldsymbol{X}^{(2)}), \quad (\boldsymbol{x}^{(3)}, \boldsymbol{X}^{(3)}), \quad \ldots$$

The smallest i such that

$$\text{tr}\{\boldsymbol{X}^{(i)} - \boldsymbol{x}^{(i)}(\boldsymbol{x}^{(i)})^\top\} < 10^{-7} \tag{5.3}$$

is denoted by i^{tight}, i.e., it is the number of rounds that Algorithm 2.1 needs to attain a tight penalization, and the resulting objective value is denoted by $q_0^{\text{tight}} \triangleq q_0(\boldsymbol{x}^{(i^{\text{tight}})})$. Moreover, the smallest i such that

$$\frac{q_0(\boldsymbol{x}^{(i-1)}) - q_0(\boldsymbol{x}^{(i)})}{|q_0(\boldsymbol{x}^{(i)})|} \leq 5 \times 10^{-4} \tag{5.4}$$

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

$$\text{GAP}(\%) = 100 \times \frac{q_0^{\text{stop}} - q_0(\boldsymbol{x}^{\text{QPLIB}})}{|q_0(\boldsymbol{x}^{\text{QPLIB}})|}. \tag{5.5}$$

Moreover, t(s) denotes the accumulative solver times in seconds for all of the i^{stop} rounds. Our results are compared with BARON [71] and COUENNE [9] by fixing the maximum solver times equal to the accumulative solver times spent by Algorithm 2.1. We ran BARON and COUENNE through GAMS v25.1.2 [29]. The resulting lower bounds, upper bounds and GAPs are reported in Tables 2, 3, and 4

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,

Table 3: Sequential SDP with flat start.

Inst	Sequential SDP with flat start							BARON			COUENNE		
	η	i^{tight}	q_0^{tight}	i^{stop}	q_0^{stop}	GAP(%)	t(s)	LB	UB	GAP(%)	LB	UP	GAP(%)
0343	1e+2	1	-2.487	23	-6.382	0.06	1.43	-7668.005	-6.386	0.00	$-\infty$	-1.688	73.56
0911	5e+0	1	-5.653	16	-30.934	3.78	2.52	$-\infty$	0.000	100.00	$-\infty$	0.000	100.00
0975	5e+0	1	-11.619	12	-36.436	3.74	0.94	-171.113	-37.794	0.16	-76.477	-0.913	97.59
1055	5e+0	1	-12.157	14	-32.868	0.51	1.27	-199.457	-32.879	0.48	$-\infty$	0.000	100.00
1143	1e+1	1	-19.447	11	-55.628	2.83	1.07	-383.133	-57.247	0.00	$-\infty$	-2.874	94.98
1157	1e+0	1	-10.154	6	-10.947	0.01	0.32	-12.165	-10.948	0.00	$-\infty$	-7.925	27.61
1353	1e+0	1	-4.953	9	-7.714	0.00	0.49	-70.152	-7.714	0.00	-15.014	-0.298	96.13
1423	2e+0	1	-7.842	11	-14.689	1.86	0.91	-74.783	-14.967	0.00	$-\infty$	-5.574	62.76
1437	1e+0	1	-6.170	7	-7.789	0.01	0.39	-83.849	-7.789	0.00	$-\infty$	-1.895	75.67
1451	1e+1	1	-38.399	14	-87.405	0.20	3.31	$-\infty$	–	–	$-\infty$	–	–
1493	1e+1	1	-16.634	14	-42.141	2.36	0.70	-393.302	-43.160	0.00	$-\infty$	-12.785	70.38
1507	1e+0	1	-6.186	7	-8.301	0.01	0.42	-38.444	-8.301	0.00	$-\infty$	-3.325	59.95
1535	2e+0	1	-3.621	13	-11.133	3.91	4.20	$-\infty$	–	–	$-\infty$	–	–
1619	1e+0	1	-7.166	7	-9.217	0.01	0.89	$-\infty$	–	–	$-\infty$	–	–
1661	1e+0	1	-13.543	14	-15.953	0.01	0.84	-133.138	-15.955	0.00	$-\infty$	-6.964	56.35
1675	1e+1	1	-28.592	9	-75.526	0.19	0.49	-431.719	-75.669	0.00	$-\infty$	-6.525	91.38
1703	1e+1	1	-88.887	10	-132.637	0.12	1.84	$-\infty$	–	–	$-\infty$	–	–
1745	1e+1	1	-38.030	17	-72.366	0.02	3.62	$-\infty$	–	–	$-\infty$	–	–
1773	2e+0	1	-4.967	21	-14.638	0.02	1.21	-115.845	-14.642	0.00	$-\infty$	–	–
1886	1e+1	1	-24.032	11	-78.651	0.03	1.31	$-\infty$	0.000	100.00	$-\infty$	0.000	100.00
1913	5e+0	1	-27.656	10	-51.881	0.44	1.11	$-\infty$	0.000	100.00	$-\infty$	0.000	100.00
1922	1e+1	1	-10.284	19	-35.725	0.63	1.31	-51.020	-35.916	0.10	$-\infty$	-7.695	78.59
1931	5e+0	1	-38.753	9	-54.290	2.55	0.78	$-\infty$	0.000	100.00	$-\infty$	-1.468	97.37
1967	1e+1	1	-73.098	13	-106.895	0.64	3.45	-621.561	-106.91	0.62	$-\infty$	0.000	100.00

Table 4: Sequential SDP+RLT with flat start.

Inst	Sequential SDP+RLT with flat start							BARON			COUENNE		
	η	i^{tight}	q_0^{tight}	i^{stop}	q_0^{stop}	GAP(%)	t(s)	LB	UB	GAP(%)	LB	UP	GAP(%)
0343	1e-1	1	-6.386	2	-6.386	0.00	0.64	-7668.0	-6.386	0.00	$-\infty$	-2.040	68.06
0911	5e+0	1	-5.653	16	-30.934	3.78	2.80	-172.78	0.000	100.00	$-\infty$	0.000	100.00
0975	5e+0	1	-11.619	12	-36.436	3.74	1.23	-171.11	-37.794	0.16	$-\infty$	-1.296	96.58
1055	5e+0	1	-12.157	14	-32.868	0.51	1.62	-199.46	-32.879	0.48	$-\infty$	0.000	100.00
1143	1e+0	2	-57.102	4	-57.230	0.03	2.20	-383.13	-57.247	0.00	$-\infty$	-2.874	94.98
1157	2e-1	1	-10.906	3	-10.947	0.01	1.49	-12.165	-10.948	0.00	-13.796	-10.948	0.00
1353	1e-1	1	-7.709	3	-7.714	0.00	3.15	-8.571	-7.714	0.00	-15.014	-7.714	0.00
1423	2e-1	2	-14.929	3	-14.929	0.25	1.61	-74.783	-14.967	0.00	$-\infty$	-5.574	62.76
1437	1e-1	1	-7.775	3	-7.789	0.00	3.04	-9.622	-7.789	0.00	-15.267	-7.789	0.00
1451	1e+1	1	-38.399	14	-87.405	0.20	47.19	-133.15	-87.576	0.00	-190.474	-86.860	0.82
1493	1e+0	2	-43.160	3	-43.160	0.00	1.38	-49.673	-43.160	0.00	-84.705	-43.160	0.00
1507	1e-1	1	-8.291	3	-8.301	0.00	0.63	-38.444	-8.301	0.00	$-\infty$	-3.325	59.95
1535	1e+0	2	-10.071	12	-11.396	1.64	33.87	-65.995	-11.454	1.14	-26.499	-11.398	1.62
1619	1e-1	1	-9.208	3	-9.217	0.01	4.02	-72.315	-9.217	0.00	-22.310	-9.217	0.00
1661	2e-1	2	-15.952	3	-15.955	0.00	8.68	-19.610	-15.955	0.00	-29.410	-15.954	0.01
1675	1e+0	2	-75.664	3	-75.668	0.00	8.77	-99.009	-75.669	0.00	-147.830	-75.668	0.00
1703	2e+0	3	-132.618	4	-132.641	0.12	10.50	-926.76	-132.38	0.32	$-\infty$	–	–
1745	2e+0	2	-71.833	5	-72.375	0.00	6.26	-314.154	-72.377	0.00	-96.439	-72.376	0.00
1773	5e-1	2	-14.413	7	-14.626	0.11	15.04	-21.705	-14.642	0.00	-36.844	-14.635	0.05
1886	1e+1	1	-24.032	11	-78.651	0.03	1.78	-324.870	0.000	100.00	$-\infty$	0.000	100.00
1913	5e+0	1	-27.656	10	-51.881	0.44	1.35	-164.260	0.000	100.00	$-\infty$	0.000	100.00
1922	1e+1	1	-10.284	19	-35.725	0.63	1.87	-51.020	-35.916	0.10	-57.575	-34.855	3.05
1931	5e+0	1	-38.753	9	-54.290	2.55	1.11	-204.080	-55.709	0.00	$-\infty$	-1.468	97.37
1967	1e+1	1	-73.098	13	-106.895	0.64	3.91	-621.561	-106.91	0.62	$-\infty$	0.000	100.00

sequential SDP requires a smaller number of rounds compared sequential SOCP to meet the stopping criterion (5.4). Using any of the relaxations, the infeasible initial points can be rounded to a feasible point with only two round of Algorithm 2.1 and all relaxations arrive at satisfactory gaps percentages eventually.

Table 5: Sequential SOCP initiated from unpenalized SDP+RLT.

Inst	Sequential SOCP initiated from SDP+RLT							BARON			COUENNE		
	η	i^{tight}	q_0^{tight}	i^{stop}	q_0^{stop}	GAP(%)	t(s)	LB	UB	GAP(%)	LB	UP	GAP(%)
1143	2e+1	1	-51.703	15	-57.212	0.06	3.37	-69.745	-57.247	0.00	-138.560	-52.342	8.57
1423	5e+0	1	-13.920	9	-14.669	1.99	1.66	-74.783	-14.967	0.00	$-\infty$	-5.574	62.76
1437	5e+0	1	-7.693	8	-7.789	0.00	6.15	-9.622	-7.789	0.00	-15.251	-7.789	0.00
1451	2e+1	2	-81.371	18	-87.555	0.02	9.02	-466.322	-87.040	0.61	$-\infty$	–	–
1493	2e+1	1	-38.699	20	-43.147	0.03	7.20	-49.673	-43.160	0.00	-84.615	-43.160	0.00
1535	5e+0	1	-9.603	32	-11.536	0.43	14.97	-65.995	-11.454	1.14	-26.499	-11.398	1.62
1661	5e+0	1	-15.691	9	-15.945	0.06	12.88	-19.610	-15.955	0.00	-29.410	-15.954	0.01
1675	2e+1	1	-71.995	7	-75.580	0.12	5.03	-99.009	-75.669	0.00	-147.830	-75.668	0.00
1703	5e+1	1	-116.411	15	-132.512	0.22	5.81	$-\infty$	–	–	$-\infty$	–	–
1745	2e+1	1	-70.078	10	-72.282	0.13	3.16	$-\infty$	–	–	$-\infty$	–	–
1773	5e+0	1	-12.240	18	-14.560	0.56	26.31	-21.705	-14.642	0.00	-36.844	-14.635	0.05

Table 6: Sequential SDP initiated from unpenalized SDP+RLT.

Inst	Sequential SDP initiated from SDP+RLT							BARON			COUENNE		
	η	i^{tight}	q_0^{tight}	i^{stop}	q_0^{stop}	GAP(%)	t(s)	LB	UB	GAP(%)	LB	UP	GAP(%)
1143	1e+1	1	-54.575	10	-57.214	0.06	0.82	-383.133	-57.247	0.00	$-\infty$	-2.874	94.98
1423	1e+0	1	-14.508	8	-14.848	0.80	0.61	$-\infty$	–	–	$-\infty$	-4.949	66.93
1437	1e+0	1	-7.759	3	-7.789	0.00	0.13	-83.849	-7.709	1.03	-15.267	–	–
1451	1e+1	1	-81.530	12	-87.561	0.02	2.66	$-\infty$	–	–	$-\infty$	–	–
1493	1e+1	1	-39.964	12	-43.160	0.00	0.56	-393.301	-43.160	0.00	$-\infty$	-12.785	70.38
1535	2e+0	1	-10.054	19	-11.537	0.42	6.42	$-\infty$	–	–	$-\infty$	–	–
1661	1e+0	1	-15.848	4	-15.948	0.04	0.23	$-\infty$	–	–	-29.424	–	–
1675	1e+1	1	-74.018	5	-75.589	0.11	0.30	-431.719	-75.669	0.00	-147.860	–	–
1703	1e+1	1	-128.930	7	-132.591	0.16	1.40	$-\infty$	–	–	-305.190	–	–
1745	1e+1	1	-70.803	7	-72.292	0.12	0.98	$-\infty$	–	–	$-\infty$	–	–
1773	2e+0	1	-13.035	11	-14.560	0.56	0.69	-115.845	-14.642	0.00	-36.856	–	–

Table 7: Sequential SDP+RLT initiated from unpenalized SDP+RLT.

Inst	Sequential SDP+RLT initiated from SDP+RLT							BARON			COUENNE		
	η	i^{tight}	q_0^{tight}	i^{stop}	q_0^{stop}	GAP(%)	t(s)	LB	UB	GAP(%)	LB	UP	GAP(%)
1143	1e+0	1	-57.083	3	-57.219	0.05	1.51	-383.133	-57.247	0.00	$-\infty$	-2.874	94.98
1423	2e-1	1	-14.929	2	-14.929	0.25	1.11	-74.783	-14.967	0.00	$-\infty$	-5.574	62.76
1437	1e-1	1	-7.789	2	-7.789	0.00	2.06	-9.622	-7.789	0.00	-15.267	-7.476	4.02
1451	2e+0	2	-86.905	4	-87.574	0.00	11.1	-466.322	-87.576	0.00	-190.514	-86.859	0.82
1493	1e+0	1	-43.160	2	-43.160	0.00	0.79	-393.302	-43.160	0.00	$-\infty$	-12.785	70.38
1535	5e-1	1	-10.932	8	-11.541	0.39	21.0	-65.995	-11.454	1.14	-26.500	-11.398	1.62
1661	1e-1	2	-15.955	2	-15.955	0.00	5.95	-19.610	-15.955	0.00	-29.424	-15.955	0.00
1675	1e+0	1	-75.662	2	-75.664	0.01	5.05	-99.009	-75.669	0.00	-147.837	-75.669	0.00
1703	2e+0	2	-132.589	3	-132.621	0.14	7.46	-926.764	-132.38	0.32	$-\infty$	–	–
1745	2e+0	1	-72.061	3	-72.313	0.09	3.74	$-\infty$	-71.070	1.80	$-\infty$	–	–
1773	2e-1	3	-14.623	4	-14.626	0.11	9.64	-21.705	-14.642	0.00	-36.844	-14.636	0.04

5.3. Sequential Penalization with SDP+RLT initialization. In this experiment, we examine a better initialization of Algorithm 2.1. Let $(\overset{*}{\boldsymbol{x}}^{\text{SDP+RLT}}, \overset{*}{\boldsymbol{X}}^{\text{SDP+RLT}})$ denote the optimal solution of the unpenalized SDP+RLT relaxation (2.14a)-(2.14e). We use the point $\hat{\boldsymbol{x}} = \overset{*}{\boldsymbol{x}}^{\text{SDP+RLT}}$ 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 gap (See Table 1). Tables 5, 6 and 7 report the results of sequential penalized SOCP, SDP and SDP+RLT relaxations. For cases 1493 and 1773, better initialization of sequential SOCP relaxation results in 3.11% and 4.39% optimality gap improvements, respectively. For the remaining cases initialization does not have any major impact and hence, the proposed framework is not very sensitive to the choice of initial point. Figures 1 and 2 show the convergence behavior of Algorithm 2.1 for cases 1451 and 1745 given different initial points. The choice of η for all diagrams are taken from the

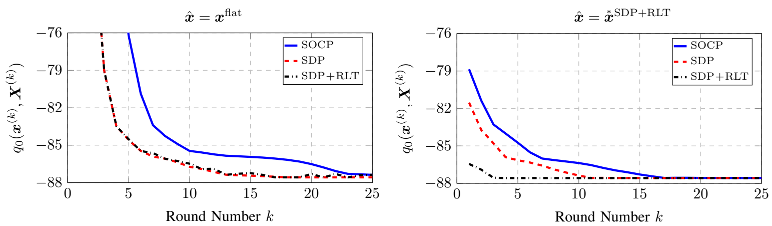

Fig. 1: Convergence behavior of the sequential SOCP, SDP, and SDP+RLT relaxations for the case 1451 with the initial points $\hat{x} = x^{\text{flat}}$ and $\hat{x} = \overset{*}{x}^{\text{ SDP+RLT}}$, respectively.

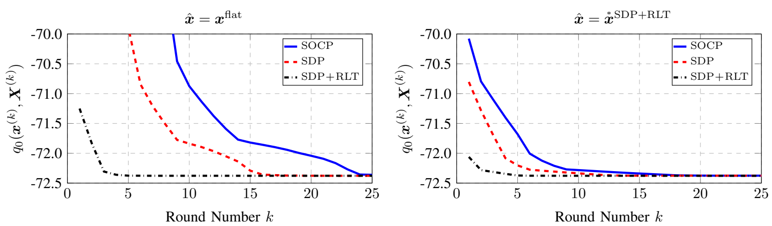

Fig. 2: Convergence behavior of the sequential SOCP, SDP, and SDP+RLT relaxations for the case 1745 with the initial points $\hat{x} = x^{\text{flat}}$ and $\hat{x} = \overset{*}{x}^{\text{ SDP+RLT}}$, respectively.

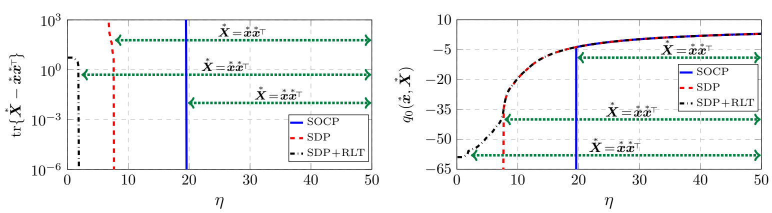

Fig. 3: Sensitivity of the penalized SOCP, SDP, and SDP+RLT relaxations to the choice of penalty parameter η for the QCQP benchmark case 1143 and $\hat{x} = x^{\text{flat}}$.

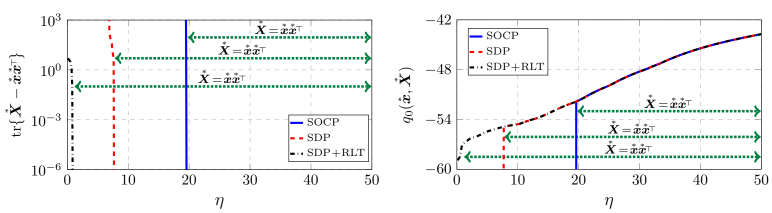

Fig. 4: Sensitivity of the penalized SOCP, SDP, and SDP+RLT relaxations to the choice of penalty parameter η for the QCQP benchmark case 1143 and $\hat{x} = \overset{*}{x}^{\text{ SDP+RLT}}$.

corresponding rows of the Tables 2, 3, 4, 5, 6 and 7.

5.4. 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 (2.16a)-(2.16e) is solved for a wide range of η values. The benchmark case 1143 is used for this experiment. Figures 3 and 4 show the results for $\hat{x} = x^{\text{flat}}$ and $\hat{x} = \overset{*}{x}^{\text{SDP+RLT}}$, respectively. If η is small, none of the proposed penalized

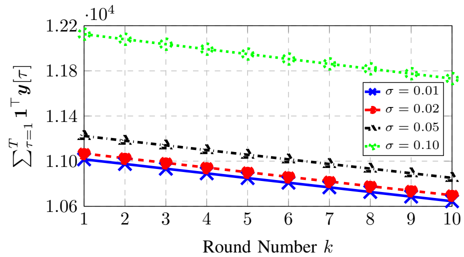

Fig. 5: Convergence behavior of the proposed sequential SOCP relaxation for large-scale system identification with different disturbance levels.

relaxations are tight for the case 1143. As the value of η increases, the feasibility violation $\mathrm{tr}\{\mathring{\boldsymbol{X}} - \mathring{\boldsymbol{x}}\mathring{\boldsymbol{x}}^\top\}$ abruptly vanishes once crossing a certain threshold. According to the figures, all three relaxations produce feasible points for a wide range of η values. Remarkably, if $\mathring{\boldsymbol{x}}^{\mathrm{SDP+RLT}}$ is used as the initial point and $\eta \simeq 1$, then the penalized SDP+RLT relaxation (2.16a)-(2.16e) produces a feasible point for the benchmark case 1143 whose objective value is within %0.2 of the reported optimal cost $q_0(\boldsymbol{x}^{\mathrm{QPLIB}})$.

6. Large-scale System Identification Problems. Following [25], 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:

$$\text{(6.1a)} \qquad \boldsymbol{z}[\tau+1] = \boldsymbol{A}\boldsymbol{z}[\tau] + \boldsymbol{B}\boldsymbol{u}[\tau] + \boldsymbol{w}[\tau] \qquad \tau = 1, 2, \ldots, T-1$$

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:

$$\text{(6.2a)} \qquad \underset{\substack{\{\boldsymbol{y}[\tau]\in\mathbb{R}^n\}_{\tau=1}^{T-1}\\ \{\boldsymbol{z}[\tau]\in\mathbb{R}^n\}_{\tau=1}^{T}\\ \boldsymbol{A}\in\mathbb{R}^{n\times n}\\ \boldsymbol{B}\in\mathbb{R}^{n\times m}}}{\text{minimize}} \quad \sum_{\tau=1}^{T-1} \mathbf{1}_n^\top \boldsymbol{y}[\tau]$$

$$\text{(6.2b)} \qquad \text{subject to} \quad \boldsymbol{y}[\tau] \geq +\boldsymbol{z}[\tau+1] - \boldsymbol{A}\boldsymbol{z}[\tau] - \boldsymbol{B}\bar{\boldsymbol{u}}[\tau] \qquad \tau \in \{1, 2, \ldots, T-1\},$$

$$\text{(6.2c)} \qquad \boldsymbol{y}[\tau] \geq -\boldsymbol{z}[\tau+1] + \boldsymbol{A}\boldsymbol{z}[\tau] + \boldsymbol{B}\bar{\boldsymbol{u}}[\tau] \qquad \tau \in \{1, 2, \ldots, T-1\},$$

$$\text{(6.2d)} \qquad \boldsymbol{z}[\tau] = \bar{\boldsymbol{z}}[\tau] \qquad \tau \in \{\tau_1, \ldots, \tau_o\}.$$

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 (6.2b) and (6.2c).

The problem (6.2a)-(6.2d), can be cast in the form of (2.1a)-(2.1c), with respect to the vector

$$\text{(6.3)} \qquad \boldsymbol{x} \triangleq [\boldsymbol{z}[1]^\top, \ldots, \boldsymbol{z}[T]^\top, \mathrm{vec}\{\boldsymbol{A}\}^\top, \alpha\boldsymbol{y}[1]^\top, \ldots, \alpha\boldsymbol{y}[T-1]^\top, \alpha\mathrm{vec}\{\boldsymbol{B}\}^\top],$$

where α is a preconditioning constant. To solve the resulting problem, we use the sequential Algorithm 2.1 equipped with the SOCP relaxation and the initial point $\hat{\boldsymbol{x}} = \mathbf{0}$.

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 21005-dimensional. Due to sparsity of the QCQP (6.2a)-(6.2d) 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 [5] 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 2.1 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 2.1. Figure 5 illustrates the convergence behavior of the objective functions for one of the trials for each disturbance level.

7. Conclusions.

This paper investigates the design of a global optimization technique for the class of quadratically-constrained quadratic programming (QCQP) problems. A penalization method 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. 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] T. Aittomaki and V. Koivunen, *Beampattern optimization by minimization of quartic polynomial*, in 2009 IEEE/SP 15th Workshop on Statistical Signal Processing, IEEE, 2009, pp. 437–440.

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

[4] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor, *Quadratic forms on graphs*, Inventiones Mathematicae, 163 (2006), pp. 499–522.

[5] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 8.1.*, 2017, http://docs.mosek.com/8.1/toolbox/index.html.
[6] A. Aubry, A. De Maio, B. Jiang, and S. Zhang, *Ambiguity function shaping for cognitive radar via complex quartic optimization*, IEEE Transactions on Signal Processing, 61 (2013), pp. 5603–5619.
[7] A. S. Bandeira, N. Boumal, and A. Singer, *Tightness of the maximum likelihood semidefinite relaxation for angular synchronization*, arXiv preprint arXiv:1411.3272, (2014).
[8] 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.
[9] P. Belotti, *COUENNE: A users manual*, tech. report, Technical report, Lehigh University, 2013.
[10] D. Bienstock and G. Munoz, *LP formulations for polynomial optimization problems*, SIAM Journal on Optimization, 28 (2018), pp. 1121–1150.
[11] J. Briet, F. Vallentin, et al., *Grothendieck inequalities for semidefinite programs with rank constraint*, arXiv preprint arXiv:1011.1754, (2010).
[12] 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.
[13] 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).
[14] S. Burgdorf, M. Laurent, and T. Piovesan, *On the closure of the completely positive semidefinite cone and linear approximations to quantum colorings*, arXiv preprint arXiv:1502.02842, (2015).
[15] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, *Phase retrieval via matrix completion*, SIAM Review, 57 (2015), pp. 225–251.
[16] E. J. Candès and B. Recht, *Exact matrix completion via convex optimization*, Foundations of Computational Mathematics, 9 (2009), pp. 717–772.
[17] E. J. Candes, T. Strohmer, and V. Voroninski, *Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming*, Communications on Pure and Applied Mathematics, 66 (2013), pp. 1241–1274.
[18] 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.
[19] C.-Y. Chen and P. Vaidyanathan, *Mimo radar waveform optimization with prior information of the extended target and clutter*, IEEE Transactions on Signal Processing, 57 (2009), pp. 3533–3544.
[20] J. Chen and S. Burer, *Globally solving nonconvex quadratic programming problems via completely positive programming*, Mathematical Programming Computation, 4 (2012), pp. 33–52.
[21] C. Cid, S. Murphy, and M. Robshaw, *Computational and algebraic aspects of the advanced encryption standard*, in Proceedings of the Seventh International Workshop on Computer Algebra in Scientific Computing-CASC, vol. 2004, 2004.
[22] C. Cid, S. Murphy, and M. J. Robshaw, *Small scale variants of the aes*, in International Workshop on Fast Software Encryption, Springer, 2005, pp. 145–162.
[23] N. T. Courtois and J. Pieprzyk, *Cryptanalysis of block ciphers with overdefined systems of equations*, in International Conference on the Theory and Application of Cryptology and Information Security, Springer, 2002, pp. 267–287.
[24] M. Deza and M. Laurent, *Applications of cut polyhedrałii*, Journal of Computational and Applied Mathematics, 55 (1994), pp. 217–247.
[25] S. Fattahi and S. Sojoudi, *Data-driven sparse system identification*, in 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, 2018.
[26] G. Fazelnia, R. Madani, A. Kalbat, and J. Lavaei, *Convex relaxation for optimal distributed control problems*, IEEE Transactions on Automatic Control, 62 (2017), pp. 206–221.
[27] F. Fogel, I. Waldspurger, and A. dAspremont, *Phase retrieval for imaging problems*, Mathematical Programming Computation, (2013), pp. 1–25.
[28] 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.
[29] GAMS Development Corporation, *General Algebraic Modeling System (GAMS) Release 24.2.1*. Washington, DC, USA, 2013, http://www.gams.com/.
[30] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, *Convex optimization-based beamforming: From receive to transmit and network designs*, IEEE Signal Process. Mag, 27 (2010), pp. 62–75.
[31] 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.
[32] M. X. Goemans and D. P. Williamson, *Approximation algorithms for MAX-3-CUT and other problems*

via complex semidefinite programming, Journal of Computer and System Sciences, 68 (2004), pp. 422–470.

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

[34] 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.

[35] J. J. Hilling and A. Sudbery, *The geometric measure of multipartite entanglement and the singular values of a hypermatrix*, Journal of Mathematical Physics, 51 (2010), p. 072102.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

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

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

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

[44] M. Laurent and T. Piovesan, *Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone*, SIAM Journal on Optimization, 25 (2015), pp. 2461–2493.

[45] M. Laurent and A. Varvitsiotis, *Computing the Grothendieck constant of some graph classes*, Operations Research Letters, 39 (2011), pp. 452–456.

[46] Z. Li, S. He, and S. Zhang, *Approximation methods for polynomial optimization: Models, Algorithms, and Applications*, Springer Science & Business Media, 2012.

[47] 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.

[48] Z.-Q. Luo, W.-k. Ma, A. M.-C. So, Y. Ye, and S. Zhang, *Semidefinite relaxation of quadratic optimization problems*, IEEE Signal Processing Magazine, 27 (2010), p. 20.

[49] 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.

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

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

[52] R. Madani, J. Lavaei, and R. Baldick, *Convexification of power flow problem over arbitrary networks*, in 2015 54th IEEE Conference on Decision and Control (CDC), IEEE, 2015, pp. 1–8.

[53] R. Madani, S. Sojoudi, G. Fazelnia, and J. Lavaei, *Finding low-rank solutions of sparse linear matrix inequalities using convex optimization*, SIAM Journal on Optimization, 27 (2017), pp. 725–758.

[54] 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.

[55] 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.

[56] B. Mariere, Z.-Q. Luo, and T. N. Davidson, *Blind constant modulus equalization via convex optimization*, IEEE Transactions on Signal Processing, 51 (2003), pp. 805–818.

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

[58] C. Mu, Y. Zhang, J. Wright, and D. Goldfarb, *Scalable robust matrix recovery: Frank–wolfe meets proximal methods*, SIAM Journal on Scientific Computing, 38 (2016), pp. A3291–A3317.

[59] 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.

[60] S. Murphy and M. J. Robshaw, *Essential algebraic structure within the aes*, in Annual International Cryptology Conference, Springer, 2002, pp. 1–16.

[61] K. Natarajan, D. Shi, and K.-C. Toh, *A penalized quadratic convex reformulation method for random*

quadratic unconstrained binary optimization, Optimization Online, (2013).

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

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

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

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

[66] 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.

[67] 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.

[68] A. Singer, *Angular synchronization by eigenvectors and semidefinite programming*, Applied and Computational Harmonic Analysis, 30 (2011), pp. 20–36.

[69] 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.

[70] 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.

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

[72] O. Toker and H. Ozbay, *On the complexity of purely complex μ computation and related problems in multidimensional systems*, IEEE Transactions on Automatic Control, 43 (1998), pp. 409–414.

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

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

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

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

[77] 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.

[78] 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.