SIAM J. OPTIM.
Vol. 17, No. 3, pp. 844–860

STRONG DUALITY IN NONCONVEX QUADRATIC OPTIMIZATION WITH TWO QUADRATIC CONSTRAINTS*

AMIR BECK† AND YONINA C. ELDAR‡

Abstract. We consider the problem of minimizing an indefinite quadratic function subject to two quadratic inequality constraints. When the problem is defined over the complex plane we show that strong duality holds and obtain necessary and sufficient optimality conditions. We then develop a connection between the image of the real and complex spaces under a quadratic mapping, which together with the results in the complex case lead to a condition that ensures strong duality in the *real* setting. Preliminary numerical simulations suggest that for random instances of the extended trust region subproblem, the sufficient condition is satisfied with a high probability. Furthermore, we show that the sufficient condition is always satisfied in two classes of nonconvex quadratic problems. Finally, we discuss an application of our results to robust least squares problems.

Key words. quadratic programming, nonconvex optimization, strong duality, quadratic mappings

AMS subject classifications. 90C20, 90C26, 90C46

DOI. 10.1137/050644471

1. Introduction. In this paper we consider quadratic minimization problems with two quadratic constraints both in the real and the complex domain:

$$(QP_{\mathbb{C}}) \quad \min_{\boldsymbol{z}\in\mathbb{C}^n}\{f_0(\boldsymbol{z}) : f_1(\boldsymbol{z}) \geq 0, f_2(\boldsymbol{z}) \geq 0\}, \tag{1}$$

$$(QP_{\mathbb{R}}) \quad \min_{\boldsymbol{x}\in\mathbb{R}^n}\{f_0(\boldsymbol{x}) : f_1(\boldsymbol{x}) \geq 0, f_2(\boldsymbol{x}) \geq 0\}. \tag{2}$$

In the real case each function $f_j : \mathbb{R}^n \to \mathbb{R}$ is defined by $f_j(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A}_j\boldsymbol{x} + 2\boldsymbol{b}_j^T\boldsymbol{x} + c_j$ with $\boldsymbol{A}_j = \boldsymbol{A}_j^T \in \mathbb{R}^{n\times n}, \boldsymbol{b}_j \in \mathbb{R}^n$, and $c_j \in \mathbb{R}$. In the complex setting, $f_j : \mathbb{C}^n \to \mathbb{R}$ is given by $f_j(\boldsymbol{z}) = \boldsymbol{z}^*\boldsymbol{A}_j\boldsymbol{z} + 2\Re(\boldsymbol{b}_j^*\boldsymbol{z}) + c_j$, where $\boldsymbol{A}_j = \boldsymbol{A}_j^*$ are Hermitian matrices, $\boldsymbol{b}_j \in \mathbb{C}^n$, and $c_j \in \mathbb{R}$. The problem $(QP_{\mathbb{R}})$ appears as a subproblem in some trust region algorithms for constrained optimization [6, 10, 26] where the original problem is to minimize a general nonlinear function subject to equality constraints. The subproblem, often referred to as the *two trust region problem* [1] or the *extended trust region problem* [35], has the form

$$(TTRS) \quad \min_{\boldsymbol{x}\in\mathbb{R}^n}\left\{\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} + 2\boldsymbol{g}^T\boldsymbol{x} : \|\boldsymbol{x}\| \leq \Delta, \|\boldsymbol{A}^T\boldsymbol{x} + \boldsymbol{c}\| \leq \xi\right\}. \tag{3}$$

More details on trust region algorithms can be found in [8, 23, 36, 37, 10]. A simpler (nonconvex) quadratic problem than (TTRS) is the *trust region subproblem*, which appears in trust region algorithms for *unconstrained* optimization:

$$(TR) \quad \min_{\boldsymbol{x}\in\mathbb{R}^n}\{\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} + 2\boldsymbol{g}^T\boldsymbol{x} : \|\boldsymbol{x}\|^2 \leq \delta\}. \tag{4}$$

*Received by the editors November 7, 2005; accepted for publication (in revised form) April 25, 2006; published electronically October 16, 2006.
http://www.siam.org/journals/siopt/17-3/64447.html

†Department of Industrial Engineering and Management, Technion—Israel Institute of Technology, Haifa 32000, Israel (becka@ie.technion.ac.il).
‡Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (yonina@ee.technion.ac.il).

Problem (TR) has been studied extensively in the literature; see, e.g., [5, 12, 20, 21, 22, 30, 31] and references therein; it enjoys many useful and attractive properties. In particular, it is known that (TR) admits no duality gap and that the semidefinite relaxation (SDR) of (TR) is tight. Moreover, the solution of (TR) can be extracted from the dual solution. A necessary and sufficient condition for $\bar{\boldsymbol{x}}$ to be optimal for (TR) is that there exists $\bar{\alpha} \geq 0$ such that [15, 30]

$$(\boldsymbol{B} + \bar{\alpha}\boldsymbol{I})\bar{\boldsymbol{x}} + \boldsymbol{g} = \boldsymbol{0}, \tag{5}$$

$$\|\bar{\boldsymbol{x}}\|^2 \leq \delta, \tag{6}$$

$$\bar{\alpha}(\|\bar{\boldsymbol{x}}\|^2 - \delta) = 0, \tag{7}$$

$$\boldsymbol{B} + \bar{\alpha}\boldsymbol{I} \succeq \boldsymbol{0}. \tag{8}$$

Unfortunately, in general these results cannot be extended to the (TTRS) problem, or to $(QP_{\mathbb{R}})$. Indeed, it is known that the SDR of $(QP_{\mathbb{R}})$ is not necessarily tight [35, 36]. An exception is when the functions f_0, f_1, f_2 are homogeneous quadratic functions and there exists a positive definite linear combination of the matrices $\boldsymbol{A}_j$ [35]. Another interesting result obtained in [35], based on the dual cone representation approach [33], is that if f_1 is concave and f_2 is linear, then, although the SDR is *not necessarily tight*, $(QP_{\mathbb{R}})$ can be solved efficiently.

If the original nonlinear constrained problem has complex variables, then instead of $(QP_{\mathbb{R}})$ one should consider the complex variant $(QP_{\mathbb{C}})$. Optimization problems with complex variables appear naturally in many engineering applications. For example, if the estimation problem is posed in the Fourier domain, then typically the parameters to be estimated will be complex [24, 28]. In the context of digital communications, many signal constellations are modelled as complex valued. Another area where complex variables naturally arise is narrowband array processing [9].

Of course, every complex quadratic problem of dimension n can be written as a *real* quadratic problem of dimension $2n$ by decomposing the complex vector $\boldsymbol{z}$ as $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y}$, where $\boldsymbol{x} = \Re(\boldsymbol{z})$ and $\boldsymbol{y} = \Im(\boldsymbol{z})$ are real. Then $f_j(\boldsymbol{z})$ can be written as $f_j(\boldsymbol{z}) = \boldsymbol{w}^T\mathbf{Q}_j\boldsymbol{w} + 2\boldsymbol{d}_j^T\boldsymbol{w} + c_j$, with

$$\boldsymbol{w} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \in \mathbb{R}^{2n}, \mathbf{Q}_j = \begin{pmatrix} \Re(\boldsymbol{A}_j) & -\Im(\boldsymbol{A}_j) \\ \Im(\boldsymbol{A}_j) & \Re(\boldsymbol{A}_j) \end{pmatrix}, \boldsymbol{d} = \begin{pmatrix} \Re(\boldsymbol{b}_j) \\ \Im(\boldsymbol{b}_j) \end{pmatrix}.$$

However, the opposite claim is false: not every real quadratic problem of dimension $2n$ can be formulated as an n-dimensional complex quadratic problem. Evidently, the family of complex quadratic problems is a special case of real quadratic problems. *Why then consider the complex setting separately?* The answer to this question is that, as we shall see, there are stronger results for complex problems than for their real counterparts (cf. section 2).

In this paper we discuss both the complex and real settings. Our interest in the complex case is two-fold: First, as noted above, in certain applications we naturally deal with complex variables. Second, our derivations in the complex setting will serve as a basis for the results in the real case. In section 2, we use an extended version of the S-lemma [13] to show that under some mild conditions strong duality holds for the complex valued problem $(QP_{\mathbb{C}})$ and that the SDR is tight. We then develop optimality conditions similar to those known for the TR problem (4), and present a method for calculating the optimal solution of $(QP_{\mathbb{C}})$ from the dual solution. Thus, all the results known for (TR) can essentially be extended to $(QP_{\mathbb{C}})$. Section 3 treats the real setting. After a discussion of the complex relaxation of $(QP_{\mathbb{R}})$, which is an alternative lifting procedure to the popular SDP relaxation, we present a sufficient condition that ensures zero duality gap (and tightness of the SDR) for $(QP_{\mathbb{R}})$. Our result is based on the connection between the image of the real and complex spaces under a quadratic mapping. The advantage of our condition is that it is expressed via the dual optimal solution and therefore can be validated in polynomial-time. Furthermore, this condition can be used to establish strong duality in some general classes of problems. As we show, an example where a problem of this form arises naturally is in robust least squares design where the uncertainty set is described by two norm constraints. In addition, preliminary numerical experiments suggest that for random instances of the TTRS problem (3), our condition is often satisfied.

Throughout the paper, the following notation is used: For simplicity, instead of inf/sup we use min/max; however, this does not mean that we assume that the optimum is attained and/or finite. Vectors are denoted by boldface lowercase letters; e.g., $\boldsymbol{y}$, and matrices by boldface uppercase letters; e.g., $\boldsymbol{A}$. For two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, $\boldsymbol{A} \succ \boldsymbol{B}$ ($\boldsymbol{A} \succeq \boldsymbol{B}$) means that $\boldsymbol{A} - \boldsymbol{B}$ is positive definite (semidefinite). $\mathcal{S}_+^n = \{\boldsymbol{A} \in \mathbb{R}^{n\times n} : \boldsymbol{A} \succeq \boldsymbol{0}\}$ is the set all real valued $n \times n$ symmetric positive semidefinite matrices and $\mathcal{H}_n^+ = \{\boldsymbol{A} \in \mathbb{C}^{n\times n} : \boldsymbol{A} \succeq \boldsymbol{0}\}$ is the set of all complex valued $n \times n$ Hermitian positive semidefinite matrices. $\boldsymbol{I}_n$ is the identity matrix of order n. The real and imaginary part of scalars, vectors, or matrices are denoted by $\Re(\cdot)$ and $\Im(\cdot)$. The value of the optimal objective function of an optimization problem

$$(\mathrm{P}) : \min / \max\{f(\boldsymbol{x}) : \boldsymbol{x} \in C\}$$

is denoted by val(P). We use some standard abbreviations such as SDP (semidefinite programming), SDR (semidefinite relaxation), and LMI (linear matrix inequalities).

2. The complex case.

We begin by treating the complex valued problem $(QP_{\mathbb{C}})$. Using an extended version of the S-lemma we prove a strong duality result, and then develop necessary and sufficient optimality conditions, similar to those known for the TR problem (4) (conditions (5)–(8)). Finally, we discuss how to extract a solution for $(QP_{\mathbb{C}})$, given a dual optimal point.

2.1. Strong duality for $(QP_{\mathbb{C}})$.

The fact that strong duality in (nonconvex) quadratic optimization problems is equivalent in some sense to the existence of a corresponding S-lemma has already been exhibited by several authors [13, 25]. For example, strong duality for quadratic problems with a single constraint can be shown to follow from the nonhomogeneous S-lemma [13], which states that if there exists $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ such that $\bar{\boldsymbol{x}}^T\boldsymbol{A}_2\bar{\boldsymbol{x}} + 2\boldsymbol{b}_2^T\bar{\boldsymbol{x}} + c_2 > 0$, then the following two conditions are equivalent:

1. $\boldsymbol{x}^T\boldsymbol{A}_1\boldsymbol{x} + 2\boldsymbol{b}_1^T\boldsymbol{x} + c_1 \geq 0$ for every $\boldsymbol{x} \in \mathbb{R}^n$ such that $\boldsymbol{x}^T\boldsymbol{A}_2\boldsymbol{x} + 2\boldsymbol{b}_2^T\boldsymbol{x} + c_2 \geq 0$.
2. There exists $\lambda \geq 0$ such that

$$\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{b}_1 \\ \boldsymbol{b}_1^T & c_1 \end{pmatrix} \succeq \lambda \begin{pmatrix} \boldsymbol{A}_2 & \boldsymbol{b}_2 \\ \boldsymbol{b}_2^T & c_2 \end{pmatrix}.$$

Generalizations of the S-lemma in the real case are in general not true. For example, the natural extension to the case of two quadratic inequalities that imply a third quadratic inequality does not hold in general (see the example in [4]). However, the following theorem of Fradkov and Yakubovich [13, Theorem 2.2] extends the S-lemma to the complex case. This result will be the key ingredient in proving strong duality.

THEOREM 2.1 (extended S-lemma [13]). *Let*

$$f_j(\boldsymbol{z}) = \boldsymbol{z}^*\boldsymbol{A}_j\boldsymbol{z} + 2\Re(\boldsymbol{b}_j^*\boldsymbol{z}) + c_j, \quad \boldsymbol{z} \in \mathbb{C}^n, j = 0, 1, 2,$$

where $\boldsymbol{A}_j$ *are* $n \times n$ *Hermitian matrices,* $\boldsymbol{b}_j \in \mathbb{C}^n$*, and* $c_j \in \mathbb{R}$*. Suppose that there exists* $\tilde{\boldsymbol{z}} \in \mathbb{C}^n$ *such that* $f_1(\tilde{\boldsymbol{z}}) > 0, f_2(\tilde{\boldsymbol{z}}) > 0$*. Then the following two claims are equivalent:*

1. $f_0(\boldsymbol{z}) \geq 0$ *for every* $\boldsymbol{z} \in \mathbb{C}^n$ *such that* $f_1(\boldsymbol{z}) \geq 0$ *and* $f_2(\boldsymbol{z}) \geq 0$*.*
2. *There exists* $\alpha, \beta \geq 0$ *such that*

$$\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{b}_0 \\ \boldsymbol{b}_0^* & c_0 \end{pmatrix} \succeq \alpha \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{b}_1 \\ \boldsymbol{b}_1^* & c_1 \end{pmatrix} + \beta \begin{pmatrix} \boldsymbol{A}_2 & \boldsymbol{b}_2 \\ \boldsymbol{b}_2^* & c_2 \end{pmatrix}.$$

The Lagrangian dual of $(QP_{\mathbb{C}})$ can be shown to have the following form:[1]

$$(9) \quad (D_{\mathbb{C}}) \qquad \max_{\alpha \geq 0, \beta \geq 0, \lambda} \left\{ \lambda \;\middle|\; \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{b}_0 \\ \boldsymbol{b}_0^* & c_0 - \lambda \end{pmatrix} \succeq \alpha \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{b}_1 \\ \boldsymbol{b}_1^* & c_1 \end{pmatrix} + \beta \begin{pmatrix} \boldsymbol{A}_2 & \boldsymbol{b}_2 \\ \boldsymbol{b}_2^* & c_2 \end{pmatrix} \right\}.$$

Problem $(D_{\mathbb{C}})$ is sometimes called Shor's relaxation [29]. Theorem 2.2 states that if problem $(QP_{\mathbb{C}})$ is finite and strictly feasible, then $\mathrm{val}(QP_{\mathbb{C}}) = \mathrm{val}(D_{\mathbb{C}})$.

THEOREM 2.2 (strong duality for complex valued quadratic problems). *Suppose that problem* $(QP_{\mathbb{C}})$ *is strictly feasible, i.e., there exists* $\tilde{\boldsymbol{z}} \in \mathbb{C}^n$ *such that* $f_1(\tilde{\boldsymbol{z}}) > 0, f_2(\tilde{\boldsymbol{z}}) > 0$*. If val*$(QP_{\mathbb{C}})$ *is finite, then the maximum of problem* $(D_{\mathbb{C}})$ *is attained and val*$(QP_{\mathbb{C}})$ = *val*$(D_{\mathbb{C}})$*.*

Proof. Since $\mathrm{val}(QP_{\mathbb{C}})$ is finite then clearly

$$(10) \qquad \mathrm{val}(QP_{\mathbb{C}}) = \max_{\lambda} \{\lambda : \mathrm{val}(QP_{\mathbb{C}}) \geq \lambda\}.$$

Now, the statement $\mathrm{val}(QP_{\mathbb{C}}) \geq \lambda$ holds true if and only if the implication

$$f_1(\boldsymbol{z}) \geq 0, f_2(\boldsymbol{z}) \geq 0 \Rightarrow f_0(\boldsymbol{z}) \geq \lambda$$

is valid. By Theorem 2.1 this is equivalent to

$$(11) \qquad \exists \alpha, \beta \geq 0 \quad \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{b}_0 \\ \boldsymbol{b}_0^* & c_0 - \lambda \end{pmatrix} \succeq \alpha \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{b}_1 \\ \boldsymbol{b}_1^* & c_1 \end{pmatrix} + \beta \begin{pmatrix} \boldsymbol{A}_2 & \boldsymbol{b}_2 \\ \boldsymbol{b}_2^* & c_2 \end{pmatrix}.$$

Therefore, by replacing the constraint in (10) with the LMI (11), we obtain that $\mathrm{val}(QP_{\mathbb{C}}) = \mathrm{val}(D_{\mathbb{C}})$. The maximum of $(D_{\mathbb{C}})$ is attained at $(\bar{\lambda}, \bar{\alpha}, \bar{\beta})$, where $\bar{\lambda}$ is the (finite) value $\mathrm{val}(QP_{\mathbb{C}})$ and $\bar{\alpha}, \bar{\beta}$ are the corresponding nonnegative constants that satisfy the LMI (11) for $\lambda = \bar{\lambda}$. ∎

One referee pointed us to a recent related paper [18] from June 2005, which was posted to a web site after we submitted our paper. In [18], the strong duality result of Theorem 2.2 is derived by using an interesting new rank-one decomposition, while our proof is a direct consequence of the classical extended S-lemma of Fradkov and Yakubovich.

It is interesting to note that the dual problem to $(D_{\mathbb{C}})$ is the so-called SDR of $(QP_{\mathbb{C}})$:

$$(12) \quad (SDR_{\mathbb{C}}) \quad \min_{\boldsymbol{Z}} \{\mathrm{Tr}(\boldsymbol{Z}\boldsymbol{M}_0) : \mathrm{Tr}(\boldsymbol{Z}\boldsymbol{M}_1) \geq 0, \mathrm{Tr}(\boldsymbol{Z}\boldsymbol{M}_2) \geq 0, Z_{n+1,n+1} = 1, \boldsymbol{Z} \in \mathcal{H}_{n+1}^{+}\},$$

where

$$\boldsymbol{M}_j = \begin{pmatrix} \boldsymbol{A}_j & \boldsymbol{b}_j \\ \boldsymbol{b}_j^* & c_j \end{pmatrix}.$$

[1] This formulation can be found in [34].

By the conic duality theorem (see, e.g., [4]), it follows that if both problems ($QP_{\mathbb{C}}$) and ($D_{\mathbb{C}}$) are strictly feasible, then they attain their solutions and $\mathrm{val}(QP_{\mathbb{C}}) = \mathrm{val}(D_{\mathbb{C}}) = \mathrm{val}(SDR_{\mathbb{C}})$. Finally, we note that strict feasibility of the dual problem ($D_{\mathbb{C}}$) is equivalent to saying that there exist $\tilde{\alpha} \geq 0, \tilde{\beta} \geq 0$ such that $\boldsymbol{A}_0 \succ \tilde{\alpha}\boldsymbol{A}_1 + \tilde{\beta}\boldsymbol{A}_2$. This condition is automatically satisfied when at least one of the constraints or the objective function is strictly convex (see also [35, Proposition 2.1]), an assumption that is true in many practical scenarios, for example in the TTRS problem (3).

2.2. Optimality conditions. Theorem 2.3 will be very useful in section 2.3, where a method for extracting the optimal solution of ($QP_{\mathbb{C}}$) from the optimal dual solution of ($D_{\mathbb{C}}$) will be described.

THEOREM 2.3. *Suppose that both problems* ($QP_{\mathbb{C}}$) *and* ($D_{\mathbb{C}}$) *are strictly feasible, and let* $(\bar{\alpha}, \bar{\beta}, \bar{\lambda})$ *be an optimal solution of* ($D_{\mathbb{C}}$). *Then* $\bar{\boldsymbol{z}}$ *is an optimal solution of* ($QP_{\mathbb{C}}$) *if and only if*

$$(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)\bar{\boldsymbol{z}} + \boldsymbol{b}_0 - \bar{\alpha}\boldsymbol{b}_1 - \bar{\beta}\boldsymbol{b}_2 = \boldsymbol{0}, \tag{13}$$

$$f_1(\bar{\boldsymbol{z}}), f_2(\bar{\boldsymbol{z}}) \geq 0, \tag{14}$$

$$\bar{\alpha} f_1(\bar{\boldsymbol{z}}) = \bar{\beta} f_2(\bar{\boldsymbol{z}}) = 0. \tag{15}$$

Proof. The proof follows from the strong duality result (Theorem 2.2) and from saddle point optimality conditions (see, e.g., [2, Theorem 6.2.5]). ∎

Note that a direct consequence of Theorem 2.3 is that the linear system (13) is consistent.

We now develop necessary and sufficient optimality conditions for ($QP_{\mathbb{C}}$) assuming strict feasibility, which are a natural generalization of the optimality conditions (5)–(8) for the trust region subproblem. Notice that for the complex version of the (TTRS), strict feasibility of ($D_{\mathbb{C}}$) is always satisfied since the norm constraint is strictly convex.

THEOREM 2.4. *Suppose that both problems* ($QP_{\mathbb{C}}$) *and* ($D_{\mathbb{C}}$) *are strictly feasible. Then* $\bar{\boldsymbol{z}}$ *is an optimal solution of* ($QP_{\mathbb{C}}$) *if and only if there exist* $\alpha, \beta \geq 0$ *such that*

(i) $(\boldsymbol{A}_0 - \alpha\boldsymbol{A}_1 - \beta\boldsymbol{A}_2)\bar{\boldsymbol{z}} + \boldsymbol{b}_0 - \alpha\boldsymbol{b}_1 - \beta\boldsymbol{b}_2 = 0$;
(ii) $f_1(\bar{\boldsymbol{z}}), f_2(\bar{\boldsymbol{z}}) \geq 0$;
(iii) $\alpha f_1(\bar{\boldsymbol{z}}) = \beta f_2(\bar{\boldsymbol{z}}) = 0$;
(iv) $\boldsymbol{A}_0 - \alpha\boldsymbol{A}_1 - \beta\boldsymbol{A}_2 \succeq 0$.

Proof. The necessary part is trivial since $\bar{\boldsymbol{z}}$, $\bar{\alpha}$, and $\bar{\beta}$ of Theorem 2.3 satisfy conditions (i)–(iv). Suppose now that conditions (i)–(iv) are satisfied. Then by (ii), $\bar{\boldsymbol{z}}$ is feasible and therefore $f_0(\bar{\boldsymbol{z}}) \geq \mathrm{val}(QP_{\mathbb{C}})$. To prove the reverse inequality ($f_0(\bar{\boldsymbol{z}}) \leq \mathrm{val}(QP_{\mathbb{C}})$), consider the unconstrained minimization problem:

$$\min_{\boldsymbol{z} \in \mathbb{C}^n} \{f_0(\boldsymbol{z}) - \bar{\alpha} f_1(\boldsymbol{z}) - \bar{\beta} f_2(\boldsymbol{z})\}. \tag{16}$$

We have

$$\begin{aligned} \mathrm{val}((16)) &\leq \min_{\boldsymbol{z} \in \mathbb{C}^n} \{f_0(\boldsymbol{z}) - \bar{\alpha} f_1(\boldsymbol{z}) - \bar{\beta} f_2(\boldsymbol{z}) : f_1(\boldsymbol{z}) \geq 0, f_2(\boldsymbol{z}) \geq 0\} \\ &\leq \min_{\boldsymbol{z} \in \mathbb{C}^n} \{f_0(\boldsymbol{z}) : f_1(\boldsymbol{z}) \geq 0, f_2(\boldsymbol{z}) \geq 0\} = \mathrm{val}(QP_{\mathbb{C}}). \end{aligned} \tag{17}$$

Conditions (i) and (iv) imply that $\bar{\boldsymbol{z}}$ is an optimal solution of (16) so that

$$f_0(\bar{\boldsymbol{z}}) - \bar{\alpha} f_1(\bar{\boldsymbol{z}}) - \bar{\beta} f_2(\bar{\boldsymbol{z}}) = \mathrm{val}((16)) \leq \mathrm{val}(QP_{\mathbb{C}}), \tag{18}$$

where the latter inequality follows from (17). By condition (iii) we have that $f_0(\bar{\boldsymbol{z}}) = f_0(\bar{\boldsymbol{z}}) - \bar{\alpha} f_1(\bar{\boldsymbol{z}}) - \bar{\beta} f_2(\bar{\boldsymbol{z}})$. Combining this with (18) we conclude that $f_0(\bar{\boldsymbol{z}}) \leq \mathrm{val}(QP_{\mathbb{C}})$. ∎

2.3. Finding an explicit solution of ($QP_{\mathbb{C}}$). Theorem 2.3 can be used to find an explicit solution to ($QP_{\mathbb{C}}$) from the solution of the dual ($D_{\mathbb{C}}$). Specifically, in section 2.3.1 we show that given the optimal dual solution, ($QP_{\mathbb{C}}$) reduces to a quadratic feasibility problem, whose solution is described in section 2.3.2.

2.3.1. Reduction to a quadratic feasibility problem. Suppose that both ($QP_{\mathbb{C}}$) and ($D_{\mathbb{C}}$) are strictly feasible. From Theorem 2.3, $\bar{\boldsymbol{z}}$ is an optimal solution if it satisfies (13), (14), and (15). If $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2 \succ \boldsymbol{0}$, then the (unique) solution to the primal problem ($QP_{\mathbb{C}}$) is given by

$$\bar{\boldsymbol{z}} = -(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)^{-1}(\boldsymbol{b}_0 - \bar{\alpha}\boldsymbol{b}_1 - \bar{\beta}\boldsymbol{b}_2).$$

Next, suppose that $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2$ is positive semidefinite but not positive definite. In this case (13) can be written as $\boldsymbol{z} = \boldsymbol{B}\boldsymbol{w} + \boldsymbol{a}$, where the columns of $\boldsymbol{B}$ form a basis for the null space of $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2$ and $\boldsymbol{a} = -(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)^{\dagger}(\boldsymbol{b}_0 - \bar{\alpha}\boldsymbol{b}_1 - \bar{\beta}\boldsymbol{b}_2)$ is a solution of (13). It follows that $\bar{\boldsymbol{z}} = \boldsymbol{B}\bar{\boldsymbol{w}} + \boldsymbol{a}$ is an optimal solution to ($QP_{\mathbb{C}}$) if and only if conditions (14) and (15) of Theorem 2.3 are satisfied, i.e.,

$$g_1(\bar{\boldsymbol{w}}) \geq 0, g_2(\bar{\boldsymbol{w}}) \geq 0, \bar{\alpha}g_1(\bar{\boldsymbol{w}}) = 0, \bar{\beta}g_2(\bar{\boldsymbol{w}}) = 0, \quad (g_j(\boldsymbol{w}) \equiv f_j(\boldsymbol{B}\boldsymbol{w} + \boldsymbol{a})). \tag{19}$$

We are left with the problem of finding a vector which is a solution of a system of two quadratic equalities or inequalities as described in Table 1. This problem will be called the *quadratic feasibility problem*.

TABLE 1
Cases of the quadratic feasibility problem.

No.	Case	Feasibility problem
I	$\bar{\alpha} = 0, \bar{\beta} = 0$	$g_1(\boldsymbol{w}) \geq 0$ and $g_2(\boldsymbol{w}) \geq 0$
II	$\bar{\alpha} > 0, \bar{\beta} = 0$	$g_1(\boldsymbol{w}) = 0$ and $g_2(\boldsymbol{w}) \geq 0$
III	$\bar{\alpha} = 0, \bar{\beta} > 0$	$g_1(\boldsymbol{w}) \geq 0$ and $g_2(\boldsymbol{w}) = 0$
IV	$\bar{\alpha} > 0, \bar{\beta} > 0$	$g_1(\boldsymbol{w}) = 0$ and $g_2(\boldsymbol{w}) = 0$

Note that since $(\bar{\lambda}, \bar{\alpha}, \bar{\beta})$ is an optimal solution of the dual problem ($D_{\mathbb{C}}$), we must have $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2 \succeq \boldsymbol{0}$. Thus, the first case is possible only when $\boldsymbol{A}_0 \succeq \boldsymbol{0}$.

We summarize the above discussion in the following theorem.

THEOREM 2.5. *Suppose that both problems* ($QP_{\mathbb{C}}$) *and* ($D_{\mathbb{C}}$) *are strictly feasible and let* $(\bar{\alpha}, \bar{\beta}, \bar{\lambda})$ *be an optimal solution of problem* ($D_{\mathbb{C}}$). *Then*

1. *if* $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2 \succ \boldsymbol{0}$, *then the (unique) optimal solution of* ($QP_{\mathbb{C}}$) *is given by*

$$\bar{\boldsymbol{z}} = -(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)^{-1}(\boldsymbol{b}_0 - \bar{\alpha}\boldsymbol{b}_1 - \bar{\beta}\boldsymbol{b}_2),$$

2. *if* $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2 \succeq \boldsymbol{0}$ *but not positive definite, then the solutions of* ($QP_{\mathbb{C}}$) *are* $\boldsymbol{z} = \boldsymbol{B}\boldsymbol{w} + \boldsymbol{a}$, *where the columns of* $\boldsymbol{B} \in \mathbb{C}^{n \times d}$ *form a basis for* $\mathcal{N}(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)$, $\boldsymbol{a}$ *is a solution of* (13), *and* $\boldsymbol{w} \in \mathbb{C}^d$ *(*$d = \dim\left(\mathcal{N}(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)\right)$*) is any solution of* (19).

2.3.2. Solving the quadratic feasibility problem. We now develop a method for solving all cases of the quadratic feasibility problem described in Table 1, under the condition that f_1 is strictly concave, i.e., $\boldsymbol{A}_1 \prec \boldsymbol{0}$ (so that the corresponding constraint is strictly convex).[2] The strict concavity of $g_1(\boldsymbol{w}) = f_1(\boldsymbol{B}\boldsymbol{w} + \boldsymbol{a})$ follows

[2]Note that this assumption readily implies that problem ($D_{\mathbb{C}}$) is strictly feasible.

immediately. By applying an appropriate linear transformation on g_1, we can assume without loss of generality that $g_1(\boldsymbol{w}) = \gamma - \|\boldsymbol{w}\|^2$ ($\gamma \geq 0$). Our approach will be to use solutions of at most two (TR) problems.

We split our analysis according to the different cases.

Case I+II. A solution to the feasibility problem in Case I (II) is any solution to the problem $\max\{g_2(\boldsymbol{w}) : \|\boldsymbol{w}\|^2 \leq \gamma\}$ ($\max\{g_2(\boldsymbol{w}) : \|\boldsymbol{w}\|^2 = \gamma\}$)

Case III. We first calculate $\boldsymbol{w}^0, \boldsymbol{w}^1$ given by

$$\boldsymbol{w}^0 \in \operatorname{argmin}\{g_2(\boldsymbol{w}) : \|\boldsymbol{w}\|^2 \leq \gamma\}, \quad \boldsymbol{w}^1 \in \operatorname{argmax}\{g_2(\boldsymbol{w}) : \|\boldsymbol{w}\|^2 \leq \gamma\}.$$

A solution to the feasibility problem is then given by $\bar{\boldsymbol{w}} = \boldsymbol{w}^0 + \eta(\boldsymbol{w}^1 - \boldsymbol{w}^0)$, where η is a solution to the scalar quadratic problem $g_2(\boldsymbol{w}^0 + \eta(\boldsymbol{w}^1 - \boldsymbol{w}^0)) = 0$ with $\eta \in [0, 1]$.

Case IV. Let $\boldsymbol{w}^0$ and $\boldsymbol{w}^1$ be defined by

$$\boldsymbol{w}^0 \in \operatorname{argmin}\{g_2(\boldsymbol{w}) : \|\boldsymbol{w}\|^2 = \gamma\}, \quad \boldsymbol{w}^1 \in \operatorname{argmax}\{g_2(\boldsymbol{w}) : \|\boldsymbol{w}\|^2 = \gamma\}.$$

The case in which $\boldsymbol{w}^0$ and $\boldsymbol{w}^1$ are linearly dependent can be analyzed in the same way as Case III. If $\boldsymbol{w}^0$ and $\boldsymbol{w}^1$ are linearly independent we can define

$$\boldsymbol{u}(\eta) = \boldsymbol{w}^0 + \eta(\boldsymbol{w}^1 - \boldsymbol{w}^0), \quad \boldsymbol{w}(\eta) = \sqrt{\gamma}\frac{\boldsymbol{u}(\eta)}{\|\boldsymbol{u}(\eta)\|}, \quad \eta \in [0, 1].$$

A solution to the feasibility problem is given by $w(\eta)$, where η is any root of the scalar equation $g_2(\boldsymbol{w}(\eta)) = 0$, $\eta \in [0, 1]$. The latter equation can be written (after some elementary algebraic manipulation) as the following *quartic* scalar equation:

$$\left(\gamma\boldsymbol{u}(\eta)^*\boldsymbol{A}_2\boldsymbol{u}(\eta) + c_2\|\boldsymbol{u}(\eta)\|^2\right)^2 = 4\gamma\|\boldsymbol{u}(\eta)\|^2(\Re(\boldsymbol{b}_2^*\boldsymbol{u}(\eta)))^2. \tag{20}$$

Notice that (20) has at most four solutions, which have explicit algebraic expressions.

An alternative procedure for finding an explicit solution of $(QP_{\mathbb{C}})$ is described in [18]. The dominant computational effort in both methods is the solution of the SDP $(SDR_{\mathbb{C}})$ or its dual $(D_{\mathbb{C}})$, which can be solved by a primal dual interior point method that requires $O(n^{3.5})$ operations per accuracy digit (see, e.g., [4, section 6.6.1]).

3. The real case. We now treat the problem $(QP_{\mathbb{R}})$ in which the data and variables are assumed to be real valued. The dual problem to $(QP_{\mathbb{R}})$ is

(21)

$$(D_{\mathbb{R}}) \quad \max_{\alpha\geq 0,\beta\geq 0,\lambda}\left\{\lambda \;\middle|\; \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{b}_0 \\ \boldsymbol{b}_0^T & c_0 - \lambda \end{pmatrix} \succeq \alpha\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{b}_1 \\ \boldsymbol{b}_1^T & c_1 \end{pmatrix} + \beta\begin{pmatrix} \boldsymbol{A}_2 & \boldsymbol{b}_2 \\ \boldsymbol{b}_2^T & c_2 \end{pmatrix}\right\}.$$

Note that this is *exactly the same* as problem $(D_{\mathbb{C}})$ (problem (9)), where here we use the fact that the data is real and therefore $\boldsymbol{b}_j^* = \boldsymbol{b}_j^T$. The SDR in this case is given by

(22)

$$(SDR_{\mathbb{R}}) \quad \min_{\boldsymbol{X}}\{\operatorname{Tr}(\boldsymbol{X}\boldsymbol{M}_0) : \operatorname{Tr}(\boldsymbol{X}\boldsymbol{M}_1) \geq 0, \operatorname{Tr}(\boldsymbol{X}\boldsymbol{M}_2) \geq 0, X_{n+1,n+1} = 1, \boldsymbol{X} \in \mathcal{S}_+^{n+1}\}.$$

In contrast to the complex case, strong duality is generally not true for $(QP_{\mathbb{R}})$. Nonetheless, in this section we use the results obtained for $(QP_{\mathbb{C}})$ in order to establish several results on $(QP_{\mathbb{R}})$. In section 3.1 we show that if the constraints of $(QP_{\mathbb{R}})$ are convex, then $(QP_{\mathbb{C}})$, considered as a relaxation of $(QP_{\mathbb{R}})$, can produce an approximate solution. In section 3.2 we relate the image of the real and complex space under a quadratic mapping, which will enable us to bridge between the real and complex case. Using the latter result, a sufficient condition for zero duality gap is proved in section 3.3. The condition is expressed via the optimal dual variables and thus can be verified in polynomial time. Preliminary numerical results suggest that for the TTRS problem (3) this condition is often satisfied. Moreover, we identify two general classes of problems with zero duality gap, based on this condition. As we show in section 3.4, these results can be applied to the robust least-squares problem in order to obtain a polynomial time algorithm in the presence of uncertainty sets described by two norm constraints.

3.1. The complex relaxation. As already mentioned, val($QP_{\mathbb{R}}$) is not necessarily equal to val($D_{\mathbb{R}}$). However, the *complex counterpart* ($QP_{\mathbb{C}}$) does satisfy val($QP_{\mathbb{C}}$) = val($D_{\mathbb{R}}$) and we can always find a *complex valued* solution to ($QP_{\mathbb{C}}$) that attains the bound val($D_{\mathbb{R}}$). Therefore, we can consider ($QP_{\mathbb{C}}$) as a tractable relaxation (*the complex relaxation*) of the real valued problem ($QP_{\mathbb{R}}$). The following example, whose data is taken from Yuan [36, p. 59], illustrates this fact.

Example. Consider the following real valued quadratic optimization problem:

$$\min_{x_1,x_2\in\mathbb{R}}\{-2x_1^2+2x_2^2+4x_1 : x_1^2+x_2^2-4\le 0, x_1^2+x_2^2-4x_1+3\le 0\}, \tag{23}$$

which is a special case of ($QP_{\mathbb{R}}$) with

$$\boldsymbol{A}_0=\begin{pmatrix}-2 & 0\\ 0 & 2\end{pmatrix}, \boldsymbol{A}_1=\boldsymbol{A}_2=-\boldsymbol{I}, \boldsymbol{b}_0(2;0), \boldsymbol{b}_1=\boldsymbol{0}, \boldsymbol{b}_2=(2;0), c_0=0, c_1=4, c_2=-3.$$

The solution to the dual problem is given by $\bar{\alpha}=1, \bar{\beta}=1$, and $\bar{\lambda}=-1$. It is easy to see that the optimal solution to ($QP_{\mathbb{R}}$) is given by $x_1=2, x_2=0$ and its corresponding optimal solution is 0. The duality gap is thus 1. By the strong duality result of Theorem 2.2, we can find a complex valued solution to the complex counterpart

$$\min_{z_1,z_2\in\mathbb{C}}\{-2|z_1|^2+2|z_2|^2+4\Re(z_1) : |z_1|^2+|z_2|^2-4\le 0, |z_1|^2+|z_2|^2-4\Re(z_1)+3\le 0\}. \tag{24}$$

with value equal to that of the dual problem (that is, equal to -1). Using the techniques described in section 2.3 we obtain that the solution of problem (24) is $z_1=7/4+\sqrt{15/16}i, z_2=0$ with function value -1.

The following theorem states that if the constraints of ($QP_{\mathbb{C}}$) are convex (as in the two trust region problem), then we can extract an *approximate real solution* that is feasible from the optimal complex solution $\bar{\boldsymbol{z}}$ by taking $\bar{\boldsymbol{x}}=\Re(\bar{\boldsymbol{z}})$.

THEOREM 3.1. *Suppose that both* ($QP_{\mathbb{R}}$) *and* ($D_{\mathbb{R}}$) *are strictly feasible. Let* $\boldsymbol{A}_1,\boldsymbol{A}_2\in\mathbb{R}^{n\times n}$ *be negative definite matrices,* $\boldsymbol{A}_0=\boldsymbol{A}_0^T\in\mathbb{R}^{n\times n}$, $\boldsymbol{b}_j\in\mathbb{R}^n$, *and* $c_j\in\mathbb{R}$. *Let* $\bar{\boldsymbol{z}}$ *be an optimal complex valued solution of* ($QP_{\mathbb{C}}$) *and let* $\bar{\boldsymbol{x}}=\Re(\bar{\boldsymbol{z}})$. *Then* $\bar{\boldsymbol{x}}$ *is a feasible solution of* ($QP_{\mathbb{R}}$) *and*

$$f_0(\bar{\boldsymbol{x}})-val(QP_{\mathbb{R}})\le -\Im(\bar{\boldsymbol{z}})^T\boldsymbol{A}_0\Im(\bar{\boldsymbol{z}}).$$

Proof. To show that $\bar{\boldsymbol{x}}$ is a feasible solution of ($QP_{\mathbb{R}}$) note that for $\boldsymbol{z}\in\mathbb{C}^n, j=1,2$ one has

$$\begin{aligned}0\le f_j(\boldsymbol{z})&=\boldsymbol{z}^*\boldsymbol{A}_j\boldsymbol{z}+2\Re(\boldsymbol{b}_j^*\boldsymbol{z})+c_j\\&=\Re(\boldsymbol{z})^T\boldsymbol{A}_j\Re(\boldsymbol{z})+\Im(\boldsymbol{z})^T\boldsymbol{A}_j\Im(\boldsymbol{z})+2\boldsymbol{b}_j^T\Re(\boldsymbol{z})+c_j\\&\le\Re(\boldsymbol{z})^T\boldsymbol{A}_j\Re(\boldsymbol{z})+2\boldsymbol{b}_j^T\Re(\boldsymbol{z})+c_j=f_j(\Re(\boldsymbol{z})),\end{aligned}$$

where the last inequality follows from $\boldsymbol{A}_j \prec \boldsymbol{0}$. Thus, since $\bar{\boldsymbol{z}}$ is feasible, so is $\Re(\bar{\boldsymbol{z}})$. Finally,

$$f_0(\Re(\bar{\boldsymbol{z}}))-\mathrm{val}(QP_{\mathbb{R}}) \le f_0(\Re(\bar{\boldsymbol{z}}))-\mathrm{val}(QP_{\mathbb{C}}) = f_0(\Re(\bar{\boldsymbol{z}}))-f_0(\bar{\boldsymbol{z}}) = -\Im(\bar{\boldsymbol{z}})^T\boldsymbol{A}_0\Im(\bar{\boldsymbol{z}}). \qquad \square$$

In our example the approximate solution is $(7/4, 0)$ and its function value is equal to 0.875.

The extension from real to complex variables can be considered as *lifting*. A very popular lifting procedure is the SDR in which a nonconvex quadratic optimization problem defined over $\mathbb{R}^n$ is lifted to the corresponding SDR, which is defined over the space of $n \times n$ positive semidefinite matrices $\mathcal{S}_+^n$. This approach has been studied in various contexts such as approximation of combinatorial optimization problems (see [4] and references therein), polynomial inequalities [19], and more. The lifting procedure we suggest is relevant only in the context of quadratic optimization problems with *two* quadratic constraints. Our method is based on extending the real number field $\mathbb{R}$ into the complex number field $\mathbb{C}$. The *value* of the convex relaxation $\mathrm{val}(QP_{\mathbb{C}})$ is equal to the value of the SDR $\mathrm{val}(SDR_{\mathbb{R}})$. The main difference between the two strategies is in the "projection" stage onto $\mathbb{R}^n$. In our strategy, the projection is simple and natural: we take the real part of the vector. If the constraints are convex, then we have obtained a feasible point. In contrast, the choice of projection of the SDR solution, which is an $n \times n$ matrix, is not obvious. There are well established methods for specific instances (such as Max-Cut problems), but it is not clear how to extract a "good" approximate and feasible solution for general convex quadratic constraints. Another advantage to our method is that the procedure for finding a solution to $(QP_{\mathbb{C}})$ defined in section 2.3 can be manipulated so that it will output a real valued optimal solution in the case where strong duality indeed holds. In contrast, projection of the SDR solution may no longer be optimal, even in the case of strong duality.

3.2. The image of the complex and real space under a quadratic mapping. One of the key ingredients in proving the sufficient condition in section 3.3 is a result (Theorem 3.3) on the image of the spaces $\mathbb{C}^n$ and $\mathbb{R}^n$ under a quadratic mapping, composed from two nonhomogeneous quadratic functions. Results on the image of quadratic mappings play an important role in nonconvex quadratic optimization (see, e.g., [17, 25, 27] and references therein). We begin with the following theorem due to Polyak [25, Theorem 2.2], which is very relevant to our analysis.

THEOREM 3.2 (see [25]). *Let $\boldsymbol{A}_1, \boldsymbol{A}_2 \in \mathbb{R}^{n\times n}, (n \ge 2)$ be symmetric matrices for which the following condition is satisfied:*

$$\exists \alpha, \beta \in \mathbb{R} \text{ such that } \alpha\boldsymbol{A}_1 + \beta\boldsymbol{A}_2 \succ \boldsymbol{0}. \tag{25}$$

Let $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \mathbb{R}^n$ and $c_1, c_2 \in \mathbb{R}$, and define $f_j(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A}_j\boldsymbol{x} + 2\boldsymbol{b}_j^T\boldsymbol{x} + c_j$. Then the set

$$W = \{(f_1(\boldsymbol{x}), f_2(\boldsymbol{x})) : \boldsymbol{x} \in \mathbb{R}^n\}$$

is closed and convex.

The following theorem states that the images of $\mathbb{C}^n$ and $\mathbb{R}^n$ under the quadratic mapping defined in Theorem 3.2 are the same.

THEOREM 3.3. *Consider the setup of Theorem 3.2, and let $f_j(\boldsymbol{z}) = \boldsymbol{z}^*\boldsymbol{A}_j\boldsymbol{z} + 2\Re(\boldsymbol{b}_j^*\boldsymbol{z}) + c_j$. Then the sets*

$$F = \{(f_1(\boldsymbol{z}), f_2(\boldsymbol{z})) : \boldsymbol{z} \in \mathbb{C}^n\}, \quad W = \{(f_1(\boldsymbol{x}), f_2(\boldsymbol{x})) : \boldsymbol{x} \in \mathbb{R}^n\}$$

are equal. The proof of Theorem 3.3 relies on the following lemma.

LEMMA 3.4. *Let $\boldsymbol{A}$ be a real $n \times n$ symmetric matrix, $\boldsymbol{b} \in \mathbb{R}^n, c \in \mathbb{R}$, and $\beta \geq 0$. Then*

(26)
$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} + 2\boldsymbol{b}^T \boldsymbol{x} + c : \|\boldsymbol{x}\|^2 = \beta\} = \min_{\boldsymbol{z} \in \mathbb{C}^n} \{\boldsymbol{z}^* \boldsymbol{A} \boldsymbol{z} + 2\Re(\boldsymbol{b}^* \boldsymbol{z}) + c : \|\boldsymbol{z}\|^2 = \beta\}.$$

Proof. First note that (26) is obvious for $\beta = 0$. Suppose that $\beta > 0$. The value of the first problem in (26) is equal to

$$\max_{\mu} \{\mu : \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} + 2\boldsymbol{b}^T \boldsymbol{x} + c \geq \mu \text{ for every } \boldsymbol{x} \in \mathbb{R}^n \text{ such that } \|\boldsymbol{x}\|^2 = \beta\}. \tag{27}$$

Similarly, the value of the second problem is equal to

$$\max_{\mu} \{\mu : \boldsymbol{z}^* \boldsymbol{A} \boldsymbol{z} + 2\Re(\boldsymbol{b}^* \boldsymbol{z}) + c \geq \mu \text{ for every } \boldsymbol{z} \in \mathbb{C}^n \text{ such that } \|\boldsymbol{z}\|^2 = \beta\}. \tag{28}$$

By Theorem A.2 (note that condition (45) is satisfied for $f_1(\boldsymbol{x}) \equiv \|\boldsymbol{x}\|^2 - \beta$ with $\beta > 0$), the value of both problems is equal to the value of

$$\begin{array}{ll} \max_{\mu,\lambda} & \mu \\ \text{s.t.} & \begin{pmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{b}^T & c - \mu \end{pmatrix} \succeq \lambda \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\beta \end{pmatrix}, \end{array}$$

and therefore these values are the same. $\square$

We are now ready to prove Theorem 3.3.

Proof of Theorem 3.3. By Theorem 3.2 both W and F are convex. Obviously $W \subseteq F$. To prove the opposite, we first assume without loss of generality that $f_1(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$. The latter assumption is possible since (25) is satisfied. Suppose that $(a, b) \in F$, i.e., $a = \|\boldsymbol{z}\|^2, b = f_2(\boldsymbol{z})$ for some $\boldsymbol{z} \in \mathbb{C}^n$, and let

$$b_{min} = \min\{f_2(\boldsymbol{z}) : \|\boldsymbol{z}\|^2 = a\} \text{ and } b_{max} = \max\{f_2(\boldsymbol{z}) : \|\boldsymbol{z}\|^2 = a\}.$$

By Lemma 3.4, there must be two *real* vectors $\boldsymbol{x}^0, \boldsymbol{x}^1 \in \mathbb{R}^n$ such that $\|\boldsymbol{x}^0\|^2 = \|\boldsymbol{x}^1\|^2 = a$ and $f_2(\boldsymbol{x}^0) = b_{min} \leq b \leq b_{max} = f_2(\boldsymbol{x}^1)$. Therefore, $(a, b_{min}), (a, b_{max}) \in W$. Since W is convex we conclude that (a, b), being a convex combination of (a, b_{min}) and (a, b_{max}), also belongs to W. $\square$

3.3. A sufficient condition for zero duality gap of ($QP_{\mathbb{R}}$).

3.3.1. The condition. We now use the results on the complex valued problem ($QP_{\mathbb{C}}$) in order to find a sufficient condition for zero duality gap and tightness of the SDR of the real valued problem ($QP_{\mathbb{R}}$). Our derivation is based on the fact that if an optimal solution of ($QP_{\mathbb{C}}$) is *real valued*, then ($QP_{\mathbb{R}}$) admits no gap with its dual problem ($D_{\mathbb{R}}$).

THEOREM 3.5. *Suppose that both problems* ($QP_{\mathbb{R}}$) *and* ($D_{\mathbb{R}}$) *are strictly feasible and that*

$$\exists \hat{\alpha}, \hat{\beta} \in \mathbb{R} \text{ such that } \hat{\alpha}\boldsymbol{A}_1 + \hat{\beta}\boldsymbol{A}_2 \succ \boldsymbol{0}. \tag{29}$$

Let $(\bar{\lambda}, \bar{\alpha}, \bar{\beta})$ *be an optimal solution of the dual problem* ($D_{\mathbb{R}}$). *If*

$$d = \dim\left(\mathcal{N}(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)\right) \neq 1, \tag{30}$$

then val$(QP_{\mathbb{R}})$ = *val*$(D_{\mathbb{R}})$ = *val*$(SDR_{\mathbb{R}})$ *and there exists a real valued solution to* $(QP_{\mathbb{C}})$.

Proof. Since both $(SDR_{\mathbb{R}})$ and $(D_{\mathbb{R}})$ are strictly feasible, $\mathrm{val}(D_{\mathbb{R}}) = \mathrm{val}(SDR_{\mathbb{R}})$. Now, suppose that $d = 0$. Then by (13), a solution to $(QP_{\mathbb{C}})$ is given by the real valued vector

$$\bar{\boldsymbol{x}} = -(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)^{-1}(\boldsymbol{b}_0 - \bar{\alpha}\boldsymbol{b}_1 - \bar{\beta}\boldsymbol{b}_2).$$

Since $(QP_{\mathbb{C}})$ has a real valued solution it follows that $\mathrm{val}(QP_{\mathbb{R}}) = \mathrm{val}(QP_{\mathbb{C}}) = \mathrm{val}(D_{\mathbb{R}})$, where the last equality follows from Theorem 2.2.

Next, suppose that $d \geq 2$. By Theorem 2.5, any optimal solution $\bar{\boldsymbol{z}}$ of $(QP_{\mathbb{C}})$ has the form $\bar{\boldsymbol{z}} = \boldsymbol{B}\bar{\boldsymbol{w}} + \boldsymbol{a}$, where $\bar{\boldsymbol{w}} \in \mathbb{C}^d$ is a solution of

$$g_1(\boldsymbol{w}) \geq 0, g_2(\boldsymbol{w}) \geq 0, \bar{\alpha}g_1(\boldsymbol{w}) = 0, \bar{\beta}g_2(\boldsymbol{w}) = 0, \quad (g_j(\boldsymbol{w}) \equiv f_j(\boldsymbol{B}\boldsymbol{w} + \boldsymbol{a})). \tag{31}$$

Both the matrix $\boldsymbol{B}$ and the vector $\boldsymbol{a}$ are chosen to be real valued; such a choice is possible since the columns of $\boldsymbol{B}$ form a basis for the null space of the *real valued* matrix $\boldsymbol{A}_1 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2$ and $\boldsymbol{a}$ is an arbitrary solution of a real valued linear system. Now, obviously $(g_1(\bar{\boldsymbol{w}}), g_2(\bar{\boldsymbol{w}})) \in S_1$, where $S_1 = \{(g_1(\boldsymbol{w}), g_2(\boldsymbol{w})) : \boldsymbol{w} \in \mathbb{C}^d\}$. Since $\boldsymbol{B}$ has full column rank, if (29) is satisfied, then

$$\hat{\alpha}\boldsymbol{B}^T\boldsymbol{A}_1\boldsymbol{B} + \hat{\beta}\boldsymbol{B}^T\boldsymbol{A}_2\boldsymbol{B} \succ \boldsymbol{0}. \tag{32}$$

The LMI (32) together with the fact that $d \geq 2$ imply that the conditions of Theorem 3.3 are satisfied and thus $S_1 = S_2$, where $S_2 = \{(g_1(\boldsymbol{x}), g_2(\boldsymbol{x})) : \boldsymbol{x} \in \mathbb{R}^d\}$. Therefore, there exists $\bar{\boldsymbol{x}} \in \mathbb{R}^d$ such that $g_j(\bar{\boldsymbol{w}}) = g_j(\bar{\boldsymbol{x}})$ and as a result, (31) has a real valued solution. To conclude, $\bar{\boldsymbol{z}} = \boldsymbol{B}\bar{\boldsymbol{x}} + \boldsymbol{a} \in \mathbb{R}^n$ is a real valued vector which is an optimal solution to $(QP_{\mathbb{C}})$. □

A more restrictive sufficient condition than (30) is

$$\dim(\mathcal{N}(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)) = 0.$$

This condition, as opposed to condition (30), can be directly derived from complementarity conditions of $(SDR_{\mathbb{R}})$ and its dual $(D_{\mathbb{R}})$.

We note that although a direct verification of the sufficient condition (30) requires the solution of the dual problem $(D_{\mathbb{R}})$, we will show that it is possible to use this condition in order to prove strong duality is *always* satisfied for certain classes of structured nonconvex quadratic problems (see section 3.3.3).

3.3.2. Numerical experiments. To demonstrate the fact that for the TTRS problem (3), the sufficient condition of Theorem 3.5 often holds for random problems, we considered different values of m and n (the number of constraints and the number of variables in the original nonlinear problem) and randomly generated 1000 instances of $\boldsymbol{B}, \boldsymbol{g}, \boldsymbol{A}$, and $\boldsymbol{c}$. We chose $\Delta = 0.1$ and $\xi = \|\boldsymbol{A}^T(-\alpha\boldsymbol{A}\boldsymbol{c}) + \boldsymbol{c}\|$, with

$$\alpha = \min\left\{\frac{\Delta}{\|\boldsymbol{A}\boldsymbol{c}\|}, \frac{\boldsymbol{c}^T(\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{c}}{\boldsymbol{c}^T(\boldsymbol{A}^T\boldsymbol{A})^2\boldsymbol{c}}\right\},$$

as suggested in the trust region algorithm of [6]. The SDP problems were solved by SeDuMi [32]. The results are given in Table 2.

In the table, *distribution* is the distribution from which the coefficients of $\boldsymbol{B}, \boldsymbol{g}, \boldsymbol{A}$, and $\boldsymbol{c}$ are generated. There are two possibilities: uniform distribution ($U[0,1]$) or

TABLE 2
Results for TTRS.

n	m	distribution	N_{suf}	mean	sd
10	1	Normal	997	5.50	2.34
10	1	Uniform	1000	1.61	0.62
10	10	Normal	1000	5.04	2.31
10	10	Uniform	1000	1.60	0.61
100	1	Normal	1000	13.15	2.65
100	1	Uniform	1000	3.75	0.64
100	100	Normal	1000	12.54	2.31
100	100	Uniform	1000	3.71	0.65

standard normal distribution ($N(0,1)$). N_{suf} is the number of problems satisfying the sufficient condition (30) out of 1000. *mean* and *sd* are the mean and standard deviation of the minimal eigenvalue of the matrix $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2$. Numerically, the dimension of the null space in condition (30) was determined by the number of eigenvalues of the matrix $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2$ whose absolute value was less than 10^{-8}. It is interesting to note that *almost all* the instances satisfied condition (30) except for 3 cases when $n = 10, m = 1$ with data generated from the normal distribution. Of course, these experiments reflect the situation in random problems and the results might be different (for better or for worse) if the data is generated differently.

3.3.3. Two classes of problems with zero duality gap. We will now present two classes of nonconvex quadratic problems for which the sufficient condition of Theorem 3.5 is always satisfied.

First class. Consider the problem of minimizing an indefinite quadratic function subject to a norm constraint and a linear inequality constraint:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \{\boldsymbol{x}^T\mathbf{Q}\boldsymbol{x} + 2\boldsymbol{b}^T\boldsymbol{x} : \|\boldsymbol{x}\|^2 \leq \delta, \boldsymbol{a}^T\boldsymbol{x} \leq \xi\}. \tag{33}$$

This problem was treated in [33, 35], where it was shown that the SDR is not always tight, although a polynomial-time algorithm for solving this problem was presented. We will find a condition on the data $(\mathbf{Q}, \boldsymbol{a}, \boldsymbol{b})$ that will be sufficient for zero duality gap.

THEOREM 3.6. *Suppose that problem* (33) *is strictly feasible and* $n \geq 2$. *If the dimension of* $\mathcal{N}(\mathbf{Q} - \lambda_{min}(\mathbf{Q})\boldsymbol{I}_n)$ *is at least* 2, *then strong duality holds for problem* (33).

Proof. Let $(\bar{\lambda}, \bar{\alpha}, \bar{\beta})$ be an optimal solution of the dual problem to (33). From the feasibility of the dual problem it follows that $\mathbf{Q}+\bar{\alpha}\boldsymbol{I}_n \succeq \mathbf{0}$. Now, either $\bar{\alpha} > -\lambda_{min}(\mathbf{Q})$ and in that case $\mathbf{Q} + \bar{\alpha}\boldsymbol{I}_n$ is nonsingular and thus the dimension of $\mathcal{N}(\mathbf{Q} + \bar{\alpha}\boldsymbol{I}_n)$ is 0 or $\bar{\alpha} = -\lambda_{min}(\mathbf{Q})$ and in this case $\mathcal{N}(\mathbf{Q} + \bar{\alpha}\boldsymbol{I}_n)$ is of dimension at least 2 by the assumptions. The result follows now from Theorem 3.5. □

Second class. Consider problem $(QP_{\mathbb{R}})$ with matrices $\boldsymbol{A}_i$ of the following form:

$$\boldsymbol{A}_i = \boldsymbol{I}_r \otimes \mathbf{Q}_i, \quad i = 0, 1, 2, \tag{34}$$

where $\mathbf{Q}_i = \mathbf{Q}_i^T \in \mathbb{R}^{m\times m}, r > 1$, and $n = rm$. Here $\otimes$ denotes the Kronecker product. In section 3.4 we will show that this class of problems naturally arises in unstructured robust least squares problems. The following theorem, which is a direct consequence of the sufficient condition (30), states that under some mild conditions (such as strict feasibility), strong duality holds.

THEOREM 3.7. *Suppose that both problems* $(QP_{\mathbb{R}})$ *and* $(D_{\mathbb{R}})$ *are strictly feasible and that* $\boldsymbol{A}_i$ *is given by* (34). *Moreover, suppose that there exist* $\hat{\alpha}$ *and* $\hat{\beta}$ *such that*

$$\hat{\alpha}\mathbf{Q}_1 + \hat{\beta}\mathbf{Q}_2 \succ \mathbf{0}. \tag{35}$$

Then strong duality holds for $(QP_{\mathbb{R}})$.

Proof. The validity of condition (35) readily implies that (29) holds true. Moreover, by the premise of the theorem, both problems $(QP_{\mathbb{R}})$ and $(D_{\mathbb{R}})$ are strictly feasible. We are thus left with the task of proving that condition (30) is satisfied. Indeed, let $(\bar{\lambda}, \bar{\alpha}, \bar{\beta})$ be an optimal solution of the dual problem $(D_{\mathbb{R}})$. Then the matrix $\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2$ is equal to $\boldsymbol{I}_r \otimes (\mathbf{Q}_0 - \bar{\alpha}\mathbf{Q}_1 - \bar{\beta}\mathbf{Q}_2)$. Using properties of eigenvalues of Kronecker products [14], we conclude that the multiplicities of the eigenvalues of the latter matrix must be multiplicities of r, i.e., $r, 2r, \ldots$. The dimension of $\mathcal{N}(\boldsymbol{A}_0 - \bar{\alpha}\boldsymbol{A}_1 - \bar{\beta}\boldsymbol{A}_2)$ is the multiplicity of the eigenvalue 0, which by the fact that $r > 1$, cannot be equal to 1. Hence, by Theorem 3.5, strong duality holds. ∎

3.4. Application to unstructured robust least squares. The *robust least squares* (RLS) problem was introduced and studied in [16, 7]. Consider a linear system $\boldsymbol{Ax} \approx \boldsymbol{b}$ where $\boldsymbol{A} \in \mathbb{R}^{r\times n}, \boldsymbol{b} \in \mathbb{R}^r$, and $\boldsymbol{x} \in \mathbb{R}^n$. Assume that the matrix and right-hand side vector $(\boldsymbol{A}, \boldsymbol{b})$ are not fixed but rather given by a family of matrices[3] $(\boldsymbol{A}, \boldsymbol{b}) + \boldsymbol{\Delta}^T$, where $(\boldsymbol{A}, \boldsymbol{b})$ is a known nominal value and $\boldsymbol{\Delta} \in \mathbb{R}^{(n+1)\times r}$ is an unknown perturbation matrix known to reside in a compact uncertainty set $\mathcal{U}$. The RLS approach to this problem is to seek a vector $\boldsymbol{x} \in \mathbb{R}^n$ that minimizes the worst case data error with respect to all possible values of $\boldsymbol{\Delta} \in \mathcal{U}$:

$$\min_{\boldsymbol{x}} \max_{\boldsymbol{\Delta}\in\mathcal{U}} \left\| \boldsymbol{Ax} - \boldsymbol{b} + \boldsymbol{\Delta}^T \begin{pmatrix} \boldsymbol{x} \\ -1 \end{pmatrix} \right\|^2. \tag{36}$$

In [16] the uncertainty set $\mathcal{U}$ in the unstructured case was chosen to contain all matrices $\boldsymbol{\Delta}$ satisfying a simple Frobenius norm constraint, i.e.,

$$\mathrm{Tr}(\boldsymbol{\Delta}^T\boldsymbol{\Delta}) \le \rho. \tag{37}$$

The RLS problem is considered difficult in the case when the uncertainty set $\mathcal{U}$ is given by an *intersection* of ellipsoids; see the related problem[4] of finding a robust counterpart of a conic quadratic problem [3]. Nonetheless, we will now show that a byproduct of our results is that as long as $r > 1$, the RLS problem with uncertainty set given by an intersection of two ellipsoids is tractable. Specifically, we consider an uncertainty set $\mathcal{U}$ given by two norm constraints:

$$\mathcal{U} = \{\boldsymbol{\Delta} \in \mathbb{R}^{(n+1)\times r} : \mathrm{Tr}(\boldsymbol{\Delta}^T\boldsymbol{B}_i\boldsymbol{\Delta}) \le \rho_i, i = 1, 2\}, \tag{38}$$

where $\boldsymbol{B}_i = \boldsymbol{B}_i^T \in \mathbb{R}^{(n+1)\times(n+1)}$ and $\rho_i > 0$. We also assume that

$$\exists\gamma_1 \ge 0, \gamma_2 \ge 0 \text{ such that } \gamma_1\boldsymbol{B}_1 + \gamma_2\boldsymbol{B}_2 \succ \mathbf{0}. \tag{39}$$

The above condition will ensure strict feasibility of the dual problem to the inner maximization problem of (36).

[3]The perturbation matrix appears in a transpose form for the sake of simplicity of notation.

[4]Note that finding a tractable formulation to the RLS problem is the key ingredient in deriving a robust counterpart of a conic quadratic constraint of the form $\|\boldsymbol{Ax} + \boldsymbol{b}\| \le \boldsymbol{c}^T\boldsymbol{x} + d$.

The form of the uncertainty set (38) is more general than the standard single-constraint form (37) and it can thus be used to describe more complicated scenarios of uncertainties. Using some simple algebraic manipulations the objective function in (36) can be written as

$$\|\boldsymbol{Ax}-\boldsymbol{b}+\boldsymbol{\Delta}^T\tilde{\boldsymbol{x}}\|^2 = \mathrm{Tr}(\boldsymbol{\Delta}^T\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T\boldsymbol{\Delta}) + 2\mathrm{Tr}((\boldsymbol{Ax}-\boldsymbol{b})\tilde{\boldsymbol{x}}^T\boldsymbol{\Delta}) + \mathrm{Tr}((\boldsymbol{Ax}-\boldsymbol{b})(\boldsymbol{Ax}-\boldsymbol{b})^T),$$

where we denoted

$$\tilde{\boldsymbol{x}} = \begin{pmatrix} \boldsymbol{x} \\ -1 \end{pmatrix}. \tag{40}$$

Relying on the identities

$$\mathrm{Tr}(\boldsymbol{A}^T\boldsymbol{B}\boldsymbol{A}) = \mathrm{vec}(\boldsymbol{A})^T(\boldsymbol{I}_r\otimes\boldsymbol{B})\,\mathrm{vec}(\boldsymbol{A}), \quad \mathrm{Tr}(\boldsymbol{A}^T\mathbf{C}) = \mathrm{vec}(\boldsymbol{A})^T\,\mathrm{vec}(\mathbf{C}) \tag{41}$$

for every $\boldsymbol{A},\mathbf{C}\in\mathbb{R}^{p\times r}\boldsymbol{B}\in\mathbb{R}^{p\times p}$, where for a matrix $\boldsymbol{M}$, $\mathrm{vec}(\boldsymbol{M})$ denotes the vector obtained by stacking the columns of $\boldsymbol{M}$, the inner maximization problem in (36) takes the following form:

$$\max\{\mathrm{vec}(\boldsymbol{\Delta})^T\mathbf{Q}\,\mathrm{vec}(\boldsymbol{\Delta}) + 2\boldsymbol{f}^T\,\mathrm{vec}(\boldsymbol{\Delta}) + c : \boldsymbol{\Delta}\in\mathcal{U}\}, \tag{42}$$

where $\mathbf{Q} = \boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T$, $\boldsymbol{f} = \mathrm{vec}(\tilde{\boldsymbol{x}}(\boldsymbol{Ax}-\boldsymbol{b})^T)$, and $c = \|\boldsymbol{Ax}-\boldsymbol{b}\|^2$. By the first identity of (41) it follows that $\mathcal{U}$ can be written as

$$\mathcal{U} = \{\boldsymbol{\Delta}\in\mathbb{R}^{(n+1)\times r} : \mathrm{vec}(\boldsymbol{\Delta})^T(\boldsymbol{I}_r\otimes\boldsymbol{B}_i)\,\mathrm{vec}(\boldsymbol{\Delta}) \le \rho_i, i=1,2\}.$$

Therefore, all the matrices in the inner maximization problem (42) are of the form $\boldsymbol{I}_r\otimes G$. Noting that all the other conditions of Theorem 3.7 are satisfied (strict feasibility of the primal and dual problems and (35)), we conclude that strong duality holds for (42) and its value is thus equal to the value of the dual problem given by

$$\min_{\alpha\ge0,\beta\ge0,\lambda}\left\{-\lambda\,\middle|\,\begin{pmatrix} -\mathbf{Q}+\boldsymbol{I}_r\otimes(\alpha\boldsymbol{B}_1+\beta\boldsymbol{B}_2) & -\boldsymbol{f} \\ -\boldsymbol{f}^T & -c-\lambda-\alpha\rho_1-\beta\rho_2 \end{pmatrix}\succeq\mathbf{0}\right\}.$$

Now, using the following identities (see [14]):

$$\begin{aligned} \mathbf{Q} &= \boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T = (\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}})(\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}})^T, \\ \boldsymbol{f} &= \mathrm{vec}(\tilde{\boldsymbol{x}}(\boldsymbol{Ax}-\boldsymbol{b})^T) = (\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}})(\boldsymbol{Ax}-\boldsymbol{b}) \end{aligned}$$

the dual problem is transformed to

$$\min_{\alpha\ge0,\beta\ge0,\lambda}$$
$$\left\{-\lambda\,\middle|\,\begin{pmatrix} -(\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}})(\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}}^T)+\boldsymbol{I}_r\otimes(\alpha\boldsymbol{B}_1+\beta\boldsymbol{B}_2) & -(\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}})(\boldsymbol{Ax}-\boldsymbol{b}) \\ -(\boldsymbol{Ax}-\boldsymbol{b})^T(\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}})^T & -\|\boldsymbol{Ax}-\boldsymbol{b}\|^2-\lambda-\alpha\rho_1-\beta\rho_2 \end{pmatrix}\succeq\mathbf{0}\right\},$$

which, by Schur complement, can be written as

$$\min_{\alpha\ge0,\beta\ge0,\lambda}\left\{-\lambda\,\middle|\,\begin{pmatrix} \boldsymbol{I}_r & (\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}})^T & \boldsymbol{Ax}-\boldsymbol{b} \\ \boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}} & \boldsymbol{I}_r\otimes(\alpha\boldsymbol{B}_1+\beta\boldsymbol{B}_2) & \mathbf{0} \\ (\boldsymbol{Ax}-\boldsymbol{b})^T & \mathbf{0} & -\lambda-\alpha\rho_1-\beta\rho_2 \end{pmatrix}\succeq\mathbf{0}\right\}.$$

Finally, we arrive at the following SDP formulation of the RLS problem (36):

$$\min_{\alpha\geq 0,\beta\geq 0,\lambda,x}\left\{-\lambda\,\middle|\begin{pmatrix}\boldsymbol{I}_r & (\boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}})^T & \boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\\ \boldsymbol{I}_r\otimes\tilde{\boldsymbol{x}} & \boldsymbol{I}_r\otimes(\alpha\boldsymbol{B}_1+\beta\boldsymbol{B}_2) & \boldsymbol{0}\\ (\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})^T & \boldsymbol{0} & -\lambda-\alpha\rho_1-\beta\rho_2\end{pmatrix}\succeq\boldsymbol{0}\right\}. \tag{43}$$

We summarize the discussion in this section in the following theorem.

THEOREM 3.8. *Consider the RLS problem* (36), *where the uncertainty set* $\mathcal{U}$ *is given by* (38), $r>1$, *and* $\boldsymbol{B}_1,\boldsymbol{B}_2$ *satisfy condition* (39). *Let* $(\alpha,\beta,\lambda,\boldsymbol{x})$ *be a solution to the SDP problem* (43), *where* $\tilde{\boldsymbol{x}}$ *is given in* (40). *Then* $\boldsymbol{x}$ *is the optimal solution of the RLS problem* (36).

Appendix. Extended Finsler's theorem.

THEOREM A.1 (Finsler's theorem [11, 21]). *Let* $\mathbb{F}$ *be one of the fields* $\mathbb{R}$ *or* $\mathbb{C}$ *and let* $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times n}$ *be symmetric matrices. Suppose that there exist* $\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathbb{F}^n$ *such that* $\boldsymbol{x}_1^*\boldsymbol{A}\boldsymbol{x}_1>0$ *and* $\boldsymbol{x}_2^*\boldsymbol{A}\boldsymbol{x}_2<0$. *Then*

$$\boldsymbol{z}^*\boldsymbol{B}\boldsymbol{z}\geq 0 \text{ for every } \boldsymbol{z}\in\mathbb{F}^n \text{ such that } \boldsymbol{z}^*\boldsymbol{A}\boldsymbol{z}=0$$

if and only if there exists $\alpha\in\mathbb{R}$ *such that* $\boldsymbol{B}-\alpha\boldsymbol{A}\succeq\boldsymbol{0}$.

We note that the complex case can be reduced to the real case by using

$$\boldsymbol{z}^*\boldsymbol{A}\boldsymbol{z}=(\boldsymbol{x}^T\boldsymbol{y}^T)\begin{pmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}$$

for all $\boldsymbol{z}=\boldsymbol{z}+i\boldsymbol{y},\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$, where $\boldsymbol{A}\in\mathbb{R}^{n\times n}$ is symmetric.

While Finsler's theorem deals with *homogeneous* quadratic forms, the extended version considers *nonhomogeneous* quadratic functions.

THEOREM A.2 (extended Finsler's theorem). *Let* $\mathbb{F}$ *be one of the fields* $\mathbb{R}$ *or* $\mathbb{C}$ *and let* $\boldsymbol{A}_1,\boldsymbol{A}_2\in\mathbb{R}^{n\times n}$ *be symmetric matrices such that*

$$\boldsymbol{A}_2\succeq\eta\boldsymbol{A}_1 \text{ for some } \eta\in\mathbb{R}. \tag{44}$$

Let $f_j:\mathbb{F}^n\to\mathbb{R}, f_j(\boldsymbol{x})=\boldsymbol{x}^*\boldsymbol{A}_j\boldsymbol{x}+2\Re(\boldsymbol{b}_j^T\boldsymbol{x})+c_j$, *where* $\boldsymbol{b}_j\in\mathbb{R}^n$ *and* c_j *is a real scalar.*[5] *Suppose that*

$$\exists\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathbb{F}^n \text{ such that } f_1(\boldsymbol{x}_1)>0, f_1(\boldsymbol{x}_2)<0. \tag{45}$$

Then the following two statements are equivalent:

(i) $f_2(\boldsymbol{x})\geq 0$ *for every* $\boldsymbol{x}\in\mathbb{F}^n$ *such that* $f_1(\boldsymbol{x})=0$.

(ii) *There exists* $\lambda\in\mathbb{R}$ *such that*

$$\begin{pmatrix}\boldsymbol{A}_2 & \boldsymbol{b}_2\\ \boldsymbol{b}_2^T & c_2\end{pmatrix}\succeq\lambda\begin{pmatrix}\boldsymbol{A}_1 & \boldsymbol{b}_1\\ \boldsymbol{b}_1^T & c_1\end{pmatrix}.$$

Proof. (ii) $\Rightarrow$ (i) is a trivial implication. Now, suppose that (i) is satisfied. Making the change of variables $\boldsymbol{x}=(1/t)\boldsymbol{y}$ $(\boldsymbol{y}\in\mathbb{F}^n,t\neq 0)$ and multiplying f_1 and f_2 by $|t|^2$, (i) becomes

$$g_2(\boldsymbol{y},t)\geq 0 \text{ for every } \boldsymbol{y}\in\mathbb{F}^n,t\neq 0 \text{ such that } g_1(\boldsymbol{y},t)=0, \tag{46}$$

[5] In the case $\mathbb{F}=\mathbb{R}$, f_j can be written as $f_j(\boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{A}_j\boldsymbol{x}+2\boldsymbol{b}_j^T\boldsymbol{x}+c_j$.

where $g_j(\boldsymbol{y},t) = \boldsymbol{y}^*\boldsymbol{A}_j\boldsymbol{y} + 2\Re(\boldsymbol{b}_j^T\boldsymbol{y}t) + c_j|t|^2$. Notice that if t would not be restricted to be nonzero, then by Theorem A.1, statement (ii) is true (g_1 and g_2 are homogeneous quadratic functions). Thus, all is left to prove is that (46) is true for $t = 0$. However, by replacing $t \neq 0$ with $t = 0$, (46) reduces to

$$\boldsymbol{y}^*\boldsymbol{A}_2\boldsymbol{y} \geq 0 \text{ for every } \boldsymbol{y} \in \mathbb{F}^n \text{ such that } \boldsymbol{y}^*\boldsymbol{A}_1\boldsymbol{y} = 0,$$

which, by Theorem A.1, is equivalent to condition (44). $\square$

The condition in Theorem A.2 holds true, for instance, if $\boldsymbol{A}_2$ is positive definite or if $\boldsymbol{A}_1$ is definite. The case in which $\boldsymbol{A}_1$ is definite was already proven for the real case in [33, Corollary 6].

REFERENCES

[1] K. Anstreicher and H. Wolkowicz, *On Lagrangian relaxation of quadratic matrix constraints*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 41–55.

[2] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 2nd ed., John Wiley and Sons, New York, 1993.

[3] A. Ben-Tal and A. Nemirovski, *Robust convex optimization*, Math. Oper. Res., 23 (1998), pp. 769–805.

[4] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPS-SIAM Ser. Optim. 2, SIAM, Philadelphia, 2001.

[5] A. Ben-Tal and M. Teboulle, *Hidden convexity in some nonconvex quadratically constrained quadratic programming*, Math. Programming, 72 (1996), pp. 51–63.

[6] M. R. Celis, J. E. Dennis, and R. A. Tapia, *A trust region strategy for nonlinear equality constrained optimization*, in Numerical Optimization, 1984 (Boulder, CO, 1984), SIAM, Philadelphia, 1985, pp. 71–82.

[7] S. Chandrasekaran, G. H. Golub, M. Gu, and A. H. Sayed, *Parameter estimation in the presence of bounded data uncertainties*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 235–252.

[8] X. Chen and Y.-X. Yuan, *On local solutions of the Celis–Dennis–Tapia subproblem*, SIAM J. Optim., 10 (1999), pp. 359–383.

[9] P. Chevalier and B. Picinbono, *Complex linear-quadratic systems for detection and array processing*, IEEE Trans. Signal Process., 44 (1996), pp. 2631–2634.

[10] A. R. Conn, N. I. M. Gold, and P. L. Toint, *Trust-Region Methods*, MPS/SIAM Ser. Optim. 1, SIAM, Philadelphia, 2000.

[11] P. Finsler, *Uber das vorkommen definiter und semi-definiter formen in scharen quadratische formen*, Commentarii Mathematici Helvetici, 9 (1937), pp. 188–192.

[12] C. Fortin and H. Wolkowicz, *The trust region subproblem and semidefinite programming*, Optim. Methods Softw., 19 (2004), pp. 41–67.

[13] A. L. Fradkov and V. A. Yakubovich, *The S-procedure and the duality relation in convex quadratic programming problems*, Vestnik Leningrad. Univ., 1 (1973), pp. 81–87, 155.

[14] D. M. Gay, *Computing optimal locally constrained steps*, SIAM J. Sci. Statist. Comput., 2 (1981), pp. 186–197.

[15] D. M. Gay, *A trust-region approach for linearly constrained optimization*, in Proceedings of the Dundee Biennial Conference on Numerical Analysis, Dundee, Austalia, 1983, Lecture Notes in Math. 1066, Springer, Berlin, 1984, pp. 72–105.

[16] L. El Ghaoui and H. Lebret, *Robust solution to least-squares problems with uncertain data*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 1035–1064.

[17] J. B. Hiriart-Urruty and M. Torki, *Permanently going back and forth between the "quadratic world" and the "convexity world" in optimization*, Appl. Math. Optim., 45 (2002), pp. 169–184.

[18] Y. Huang and S. Zhang, *Complex Matrix Decomposition and Quadratic Programming*, Technical report SEEM2005-02, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong, 2005.

[19] J. B. Lasserre, *Semidefinite programming versus LP relaxations for polynomial programming*, Math. Oper. Res., 27 (2002), pp. 347–360.

[20] J. M. Martínez, *Local minimizers of quadratic functions on Euclidean balls and spheres*, SIAM J. Optim., 4 (1994), pp. 159–176.

[21] J. J. Moré, *Generalizations of the trust region subproblem*, Optim. Methods Softw. 2, (1993), pp. 189–209.
[22] J. J. Moré and D. C. Sorensen, *Computing a trust region step*, SIAM J. Sci. Statist. Comput., 4 (1983), pp. 553–572.
[23] J.-M. Peng and Y. Yuan, *Optimality conditions for the minimization of a quadratic with two quadratic constraints*, SIAM J. Optim., 7 (1997), pp. 579–594.
[24] B. Picinbono and P. Chevalier, *Widely linear estimation with complex data*, IEEE Trans. Signal Process., 43 (1995), pp. 2030–2033.
[25] B. T. Polyak, *Convexity of quadratic transformations and its use in control and optimization*, J. Optim. Theory Appl., 99 (1998), pp. 553–583.
[26] M. J. D. Powell and Y. Yuan, *A trust region algorithm for equality constrained optimization*, Math. Programming, 49 (1990/91), pp. 189–211.
[27] M. V. Ramana and A. J. Goldman, *Some geometric results in semidefinite programming*, J. Global Optim., 7 (1995), pp. 33–50.
[28] P. J. Schreier, L. L. Scharf, and C. T. Mullis, *Detection and estimation of improper complex random signals*, IEEE Trans. Inform. Theory, 51 (2005), pp. 306–312.
[29] N. Z. Shor, *Quadratic optimization problems*, Izv. Akad. Nauk SSSR Tekhn. Kibernet., 1 (1987), pp. 128–139, 222.
[30] D. C. Sorensen, *Newton's method with a model trust region modification*, SIAM J. Numer. Anal., 19 (1982), pp. 409–426.
[31] R. J. Stern and H. Wolkowicz, *Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations*, SIAM J. Optim., 5 (1995), pp. 286–313.
[32] J. F. Sturm, *Using SeDuMi* 1.02*, a MATLAB toolbox for optimization over symmetric cones*, Optim. Methods Softw., 11/12 (1999), pp. 625–653.
[33] J. F. Sturm and S. Zhang, *On cones of nonnegative quadratic functions*, Math. Oper. Res., 28 (2003), pp. 246–267.
[34] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 40–95.
[35] Y. Ye and S. Zhang, *New results on quadratic minimization*, SIAM J. Optim., 14 (2003), pp. 245–267.
[36] Y. Yuan, *On a subproblem of trust region algorithms for constrained optimization*, Math. Programming, 47 (1990), pp. 53–63.
[37] Y. Zhang, *Computing a Celis–Dennis–Tapia trust-region step for equality constrained optimization*, Math. Programming, 55 (1992), pp. 109–124.