Math. Program., Ser. A (2014) 147:171–206
DOI 10.1007/s10107-013-0716-2

FULL LENGTH PAPER

Trust-region problems with linear inequality constraints: exact SDP relaxation, global optimality and robust optimization

V. Jeyakumar · G. Y. Li

Received: 25 November 2011 / Accepted: 17 September 2013 / Published online: 26 September 2013

Abstract The trust-region problem, which minimizes a nonconvex quadratic function over a ball, is a key subproblem in trust-region methods for solving nonlinear optimization problems. It enjoys many attractive properties such as an exact semi-definite linear programming relaxation (SDP-relaxation) and strong duality. Unfortunately, such properties do not, in general, hold for an extended trust-region problem having extra linear constraints. This paper shows that two useful and powerful features of the classical trust-region problem continue to hold for an extended trust-region problem with linear inequality constraints under a new dimension condition. First, we establish that the class of extended trust-region problems has an exact SDP-relaxation, which holds without the Slater constraint qualification. This is achieved by proving that a system of quadratic and affine functions involved in the model satisfies a range-convexity whenever the dimension condition is fulfilled. Second, we show that the dimension condition together with the Slater condition ensures that a set of combined first and second-order Lagrange multiplier conditions is necessary and sufficient for global optimality of the extended trust-region problem and consequently for strong duality. Through simple examples we also provide an insightful account of our development from SDP-relaxation to strong duality. Finally, we show that the dimension condition

The authors are grateful to the referees for their valuable suggestions and helpful comments which have contributed to the final preparation of the paper. Research was partially supported by a grant from the Australian Research Council.

V. Jeyakumar (✉) · G. Y. Li
Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia
e-mail: v.jeyakumar@unsw.edu.au

G. Y. Li
e-mail: g.li@unsw.edu.au

is easily satisfied for the extended trust-region model that arises from the reformulation of a robust least squares problem (LSP) as well as a robust second order cone programming model problem (SOCP) as an equivalent semi-definite linear programming problem. This leads us to conclude that, under mild assumptions, solving a robust LSP or SOCP under matrix-norm uncertainty or polyhedral uncertainty is equivalent to solving a semi-definite linear programming problem and so, their solutions can be validated in polynomial time.

Mathematics Subject Classification 90C20 · 90C22 · 90C26 · 49N30

1 Introduction

Consider the extended trust-region model problem with linear inequality constraints

$$(P)\quad \min_{x\in\mathbb{R}^n} \quad x^T Ax + a^T x$$
$$\text{s.t.} \quad \|x - x_0\|^2 \le \alpha,$$
$$b_i^T x \le \beta_i,\ i = 1, \ldots, m,$$

where A is a symmetric $(n \times n)$ matrix, $a, b_i, x_0 \in \mathbb{R}^n$ and $\alpha, \beta_i \in \mathbb{R}, \alpha > 0,\ i = 1, \ldots, m$. Model problems of this form arise from the application of the trust region method for solving constrained optimization problems [10], such as nonlinear programming problems with linear inequality constraints, nonlinear optimization problems with discrete variables [2,22] (see Sect. 2) and robust optimization problems [6,8] under matrix norm [9] or polyhedral uncertainty [6,15,18] (see Sect. 5). The model (P) with a single linear inequality constraint, where $m = 1$ and $x_0 = 0$, has recently been examined in the literature (see [2,3] and other references therein).

In the special case of (P) where $(b_i, \beta_i) = (0, 0)$, it is the well-known trust-region model, and it has been extensively studied from both theoretical and algorithmic points of view [21,26,27,31]. The classical trust-region problem enjoys exact semi-definite programming relaxation (SDP-relaxation) and admits strong duality. Moreover, its solution can be found by solving a dual Lagrangian system. Unfortunately, these results are, in general, no longer true for our extended trust-region model (P). Indeed, even in the simplest case of (P) with a single linear inequality constraint, it has been shown that the SDP-relaxation is not exact (see [3,28] and other references therein). However, in the case of single inequality constraint, exact SDP-relaxation and strong duality hold under a dimension condition (see [3] and Corollary 4.2 in Sect. 4).

In this paper, we make the following contributions which extend the attractive features of the classical trust-region model to our extended trust-region model (P) under a new dimension condition:

(i) Exploiting a hidden convexity property of the extended trust-region system of (P), we establish that the SDP-relaxation of our extended trust-region problems (P) is exact whenever a dimension condition is fulfilled.
The dimension condition requires that the number of inequalities must be strictly less than the multiplicity of the minimum eigenvalue of the matrix A. It guarantees a joint-range-convexity for the extended trust-region system of (P).

The exact SDP-relaxation is derived without the standard Slater condition. For related exact relaxation result for problems involving uniform quadratic systems (see [4] and other reference therein).

(ii) We present a necessary and sufficient condition for global optimality for our model problem (P). Consequently, we derive strong duality between (P) and its Lagrangian dual problem under the Slater condition. Also, we obtain two forms of S-lemma for extended trust-region systems. In the case of (P) with two linear (bound) constraints our result provides a more general dimension condition than the corresponding condition, given recently in [3].

(iii) Under suitable, but commonly used, uncertainty sets of robust optimization, we show that the dimension condition is easily satisfied for our extended trust-region model that arises from the reformulation of a robust least squares model problem (LSP) as well as a second order cone programming model problem (SOCP) as a semi-definite linear programming problem. As a result, we establish a complete characterization of the solution of a robust (LSP) and a robust (SOCP) in terms of the solution of a semi-definite linear programming problem.

The significance of our contributions is that:

(i) Our dimension condition, expressed in terms of original data, not only reveals a hidden convexity of the extended trust-region problems but also allows direct applications to solving robust optimization problems such as the robust (LSP) and (SOCP) models. These models are increasingly becoming the models of choice for efficiently solving many classes of hard problems by relaxation or reformulation techniques [1,5,6].

(ii) Our results show that a worst-case solution of a least-squares problems or a second-order cone programming problem in the face of data uncertainty, especially in the case of a matrix-norm uncertainty or a polyhedral uncertainty, can be found by solving a semi-definite linear programming problem.

(iii) Our approach suggests further extensions of global optimality, strong duality and exact SDP-relaxation results to broad classes of extended trust-region models with (uniformly) convex quadratic constraints by way of examining joint-range convexity properties of the corresponding systems (see Sect. 6).

The outline of the paper is as follows. In Sect. 2, we introduce the dimension condition and establish a joint-range convexity property. In Sect. 3 we derive exact relaxation results for (P) and illustrate the results with numerical examples. In Sect. 4, we show that the dimension condition together with the Slater condition ensures that a combined first and second-order Lagrange multiplier condition is necessary and sufficient for global optimality of (P) and guarantees strong duality between (P) and its Lagrangian dual. In Sect. 5, we present an application of strong duality to S-lemma and consequently to robust optimization problems [6]. In Sect. 6, we show how our dimension condition can be extended to obtain corresponding exact relaxation and strong duality results for trust-regions problems with certain convex quadratic inequalities. Finally, in "Appendix", for the sake of self-containment, we describe some useful technical results that are related to non-convex quadratic systems and robust optimization.

2 Hidden convexity of extended trust regions

In this section, we derive an important hidden convexity property of extended trust-region quadratic systems which will play a key role in our study of exact relaxation and strong duality later on.

We begin by fixing the notation and definitions that will be used later in the paper. The real line is denoted by $\mathbb{R}$ and the n-dimensional real Euclidean space is denoted by $\mathbb{R}^n$. The set of all non-negative vectors of $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$. The space of all $(n \times n)$ symmetric real matrices is denoted by $S^{n\times n}$. The $(n \times n)$ identity matrix is denoted by I_n. The notation $A \succeq B$ means that the matrix $A - B$ is positive semi-definite. Moreover, the notation $A \succ B$ means the matrix $A - B$ is positive definite. The set consists of all $n \times n$ positive semidefinite matrices is denoted by S^n_+. Let $A, B \in S^{n\times n}$. The (trace) inner product of A and B is defined by $A \cdot B = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji}$, where a_{ij} is the (i, j) element of A and b_{ji} is the (j, i) element of B. A useful fact about the trace inner product is that $A \cdot (xx^T) = x^T Ax$ for all $x \in \mathbb{R}^n$ and $A \in S^{n\times n}$. For a matrix $A \in S^{n\times n}$, $\mathrm{Ker}(A) := \{d \in \mathbb{R}^n : Ad = 0\}$. For a subspace L, we use dim L to denote the dimension of L.

As in [3,4], the study of exact relaxation and strong duality requires the examination of topological and geometrical properties of the set

$$U(f, g_0, g_1, \ldots, g_m) := \{(f(x), g_0(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\} + \mathbb{R}^{m+2}_+,$$

where $f(x) = x^T Ax + a^T x + \gamma$, $g_0(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = b_i^T x - \beta_i$, $i = 1, \ldots, m$, $A \in S^{n\times n}$, $a, x_0, b_i \in \mathbb{R}^n$ and $\gamma, \alpha, \beta_i \in \mathbb{R}$, $i = 1, \ldots, m$.

We note that the range set $U(f, g_0, g_1, \ldots, g_m)$ is the sum of the nonnegative orthant and the image of the quadratic mapping $\{(f(x), g_0(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\}$. Hence, the range set $U(f, g_0, g_1, \ldots, g_m)$ is convex whenever the image of the quadratic mapping is convex. It is known that the joint-range convexity of quadratic mappings has a close relationship with strong duality of an associated optimization problem. For example, Fradkov and Yakubovich [12,29] used convexity of the joint-range $\{(f(x), g_0(x)) : x \in \mathbb{R}^n\}$ in the case of homogeneous (not necessarily convex) quadratic functions f, g_0 (cf. [11]) to show that strong duality holds for quadratic optimization problem with single quadratic constraint, under the Slater condition.

Recently, Polyak [25] established a strong duality result for homogenous nonconvex quadratic problems involving two quadratic constraints by showing that the joint-range of three homogenous quadratic functions is convex under a positive definiteness condition. On the other hand, the image of three nonhomogeneous quadratic function is, in general, not convex. See [4,24,25] for more detailed discussion for joint-range convexity of quadratic functions.

We begin by showing that the set $U(f, g_0, g_1, \ldots, g_m)$ is always closed.

Proposition 2.1 *Let $f(x) = x^T Ax + a^T x + \gamma$, $g_0(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = b_i^T x - \beta_i$, $i = 1, \ldots, m$, $A \in S^{n\times n}$, $a, x_0, b_i \in \mathbb{R}^n$ and $\gamma, \alpha, \beta_i \in \mathbb{R}$, $i = 1, \ldots, m$. Then $U(f, g_0, g_1, \ldots, g_m)$ is closed.*

Proof Let $\left(r^k, s_0^k, s_1^k, \ldots, s_m^k\right) \in U\,(f, g_0, g_1, \ldots, g_m)$ with

$$\left(r^k, s_0^k, s_1^k, \ldots, s_m^k\right) \to (r, s_0, s_1, \ldots, s_m)\,.$$

By the definition, for each k, there exists $x^k \in \mathbb{R}^n$

$$f(x^k) \le r^k,\ \|x^k - x_0\|^2 \le \alpha + s_0^k,\ b_1^T x^k \le \beta_1 + s_1^k\,, \ldots,\ b_m^T x^k \le \beta_m + s_m^k. \tag{2.1}$$

This implies that x^k is bounded, and so, by passing to subsequences, we may assume that $x_k \to x$. Then, passing limits in (2.1), we have

$$f(x) \le r, \|x - x_0\|^2 \le \alpha + s_0, b_1^T x \le \beta_1 + s_1, \ldots, b_m^T x \le \beta_m + s_m.$$

That is to say, $(r, s_0, s_1, \ldots, s_m) \in U(f, g_0, g_1, \ldots, g_m)$. So, $U(f, g_0, g_1, \ldots, g_m)$ is closed. □

The following simple one-dimensional example shows that the set $U(f, g_0, g_1, \ldots, g_m)$ is, in general, not a convex set.

Example 2.1 (Nonconvexity of $U(f, g_0, g_1, \ldots, g_m)$) For (P), let $n = 1, m = 1,\ f(x) = x - x^2,\ g_0(x) = x^2 - 1$ and $g_1(x) = -x$. Then, $f(x) = x^T Ax + a^T x + r$ with $A = -1,\ a = 1$ and $r = 0,\ g_0(x) = \|x - x_0\|^2 - \alpha$ with $x_0 = 0,\ \alpha = 1$ and $g_1(x) = b_1^T x - \beta_1$ with $b_1 = 1$ and $\beta_1 = 0$.

Then, the set $U(f, g_0, g_1)$ is not a convex set. To see this, note that $f(0) = 0,\ g_0(0) = -1$ and $g_1(0) = 0$, and $f(1) = 0,\ g_0(1) = 0$ and $g_1(1) = -1$. So, $(0, -1, 0) \in U(f, g_0, g_1)$ and $(0, 0, -1) \in U(f, g_0, g_1)$. However, the mid point $(0, -\frac{1}{2}, -\frac{1}{2}) \notin U(f, g_0, g_1)$. Otherwise, there exists $x \in \mathbb{R}$ such that

$$x - x^2 \le 0, x^2 - 1 \le -\frac{1}{2} and - x \le -\frac{1}{2}.$$

It is easy to check that the above inequality system has no solution. This is a contradiction, and hence $\left(0, -\frac{1}{2}, -\frac{1}{2}\right) \notin U(f, g_0, g_1)$. Thus, $U(f, g_0, g_1)$ is not convex.

The following dimension condition plays a key role in the rest of the paper. Recall that, for a matrix $A \in S^n,\ \lambda_{\min}(A)$ denotes the smallest eigenvalue of A.

Definition 2.1 (Dimension condition) Consider the system of functions $f(x) = x^T Ax + a^T x + \gamma,\ g_0(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = b_i^T x - \beta_i,\ i = 1, \ldots, m$, where $A \in S^{n\times n},\ a, x_0, b_i \in \mathbb{R}^n$ and $\gamma, \alpha, \beta_i \in \mathbb{R}$. Let dim span$\{b_1, \ldots, b_m\} = s,\ s \le n$. Then, we say that the dimension condition holds for the system whenever

$$\dim \mathrm{Ker}(A - \lambda_{\min}(A) I_n) \ge s + 1. \tag{2.2}$$

In other words, the dimension condition states that the multiplicity of the minimum eigenvalue of A is at least $s+1$.

Recall that the optimal value function $h : \mathbb{R}^{m+1} \to \mathbb{R} \cup \{+\infty\}$ of (P) is given by

$$h(r, s_1, \ldots, s_m) = \begin{cases} \min\limits_{x \in \mathbb{R}^n} \left\{ f(x) : \|x - x_0\|^2 \le \alpha + r,\ b_i^T x \le \beta + s_i,\ i = 1, \ldots, m \right\}, & (r, s_1, \ldots, s_m) \in D, \\ +\infty, & \text{otherwise}, \end{cases}$$

where $D = \{(r, s_1, \ldots, s_m) : \|x - x_0\|^2 \le \alpha + r, b_i^T x \le \beta_i + s_i \text{ for some } x \in \mathbb{R}^n\}$.

Theorem 2.1 (Dimension condition revealing hidden convexity) *Let $f(x) = x^T Ax + a^T x + \gamma$, $g_0(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = b_i^T x - \beta_i$, $i = 1, \ldots, m$, $A \in S^{n \times n}$, $a, x_0, b_i \in \mathbb{R}^n$ and $\gamma, \alpha, \beta_i \in \mathbb{R}$. Suppose that the dimension condition (2.2) is satisfied.*

Then,

$$U(f, g_0, g_1, \ldots, g_m) := \{(f(x), g_0(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\} + \mathbb{R}_+^{m+2}$$

is a convex set.

Proof We first note that, if A is positive semidefinite, then $f, g_i,\ i = 0, 1, \ldots, m$, are all convex functions. So, $U(f, g_0, g_1, \ldots, g_m)$ is always convex in this case. Therefore, we may assume that A is not positive semidefinite and hence $\lambda_{\min}(A) < 0$.

[$U(f, g_0, g_1, \ldots, g_m) = epih$]. Let $D = \{(r, s_1, \ldots, s_m) : \|x - x_0\|^2 \le \alpha + r, b_i^T x \le \beta_i + s_i \text{ for some } x \in \mathbb{R}^n\}$. Clearly, D is a convex set. Then, by the definition, we have $U(f, g_0, g_1, \ldots, g_m) = \text{epi}h$.

[**Convexity of the value function** h]. To see this, we claim that, for each $(r, s_1, \ldots, s_m) \in D$, the minimization problem

$$\min_{x \in \mathbb{R}^n} \{f(x) - \lambda_{\min}(A)\|x - x_0\|^2 : \|x - x_0\|^2 \le \alpha + r, b_i^T x \le \beta_i + s_i\}$$

attains its minimum at some $\overline{x} \in \mathbb{R}^n$ with $\|\overline{x} - x_0\|^2 = \alpha + r$ and $b_i^T \overline{x} \le \beta_i + s_i$. Granting this, we have

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} \{f(x) - \lambda_{\min}(A)\|x - x_0\|^2 : \|x - x_0\|^2 \le \alpha + r,\ b_i^T x \le \beta_i + s_i\} \\ &= f(\overline{x}) - \lambda_{\min}(A)(\alpha + r) \\ &\ge \min_{x \in \mathbb{R}^n} \{f(x) : \|x - x_0\|^2 \le \alpha + r,\ b_i^T x \le \beta + s_i\} - \lambda_{\min}(A)(\alpha + r) \\ &= \min_{x \in \mathbb{R}^n} \{f(x) - \lambda_{\min}(A)(\alpha + r) : \|x - x_0\|^2 \le \alpha + r,\ b_i^T x \le \beta + s_i\} \\ &\ge \min_{x \in \mathbb{R}^n} \{f(x) - \lambda_{\min}(A)\|x - x_0\|^2 : \|x - x_0\|^2 \le \alpha + r,\ b_i^T x \le \beta_i + s_i\}, \end{aligned}$$

where the last inequality follows by $\lambda_{\min}(A) < 0$. This yields that

$$\begin{aligned}&\min_{x\in\mathbb{R}^n}\{f(x):\|x-x_0\|^2\le\alpha+r,\ b_i^Tx\le\beta+s_i,\ i=1,\dots,m\}\\&\quad=\min_{x\in\mathbb{R}^n}\{f(x)-\lambda_{\min}(A)\|x-x_0\|^2:\|x-x_0\|^2\le\alpha+r,\ b_i^Tx\le\beta_i+s_i\}\\&\qquad+\lambda_{\min}(A)(\alpha+r).\end{aligned}$$

Note that

$$\begin{aligned}F(x)&:=f(x)-\lambda_{\min}(A)\|x-x_0\|^2\\&=x^T(A-\lambda_{\min}(A)I_n)x+(a+2\lambda_{\min}(A)x_0)^Tx+(\gamma-\lambda_{\min}(A)\|x_0\|^2)\end{aligned}$$

is a convex function, and so, $(r,s_1,\dots,s_m)\mapsto\min_{x\in\mathbb{R}^n}\{F(x):\|x-x_0\|^2\le\alpha+r,b_i^Tx\le\beta_i+s_i\}$ is also convex. It follows that

$$(r,s_1,\dots,s_m)\mapsto\min_{x\in\mathbb{R}^n}\{f(x):\|x-x_0\|^2\le\alpha+r,\ b_i^Tx\le\beta+s_i,\ i=1,\dots,m\}$$

is convex. Therefore, h is convex, and so, $U(f,g_0,g_1,\dots,g_m)=\mathrm{epi}h$ is a convex set.

[**Attainment of minimizer on the sphere**] To see the claim, we proceed by the method of contradiction and suppose that any minimizer x^* of

$$\min_{x\in\mathbb{R}^n}\{F(x):\|x-x_0\|^2\le\alpha+r,b_i^Tx\le\beta_i+s_i\}$$

satisfy $\|x^*-x_0\|^2<\alpha+r$ and $b_i^Tx^*\le\beta_i+s_i$. We note that there exists $v\in\mathbb{R}^n\backslash\{0\}$ such that

$$v\in\left(\bigcap_{i=1}^m b_i^\perp\right)\cap\mathrm{Ker}(A-\lambda_{\min}(A)I_n).\tag{2.3}$$

[Otherwise, $\left(\bigcap_{i=1}^m b_i^\perp\right)\cap\mathrm{Ker}(A-\lambda_{\min}(A)I_n)=\{0\}$. Recall from our dimension condition that $\dim\mathrm{Ker}(A-\lambda_{\min}(A)I_n)\ge s+1$ where s is the dimension of $\mathrm{span}\{b_1,\dots,b_m\}$. Then, it follows from the dimension theorem that

$$\begin{aligned}n+1&=(s+1)+(n-s)\\&\le\dim\mathrm{Ker}(A-\lambda_{\min}(A)I_n)+\dim\left(\bigcap_{i=1}^m b_i^\perp\right)\\&=\dim\left(\mathrm{Ker}(A-\lambda_{\min}(A)I_n)+\bigcap_{i=1}^m b_i^\perp\right)+\dim\left(\bigcap_{i=1}^m b_i^\perp\cap\mathrm{Ker}(A-\lambda_{\min}(A)I_n)\right)\\&\le n,\end{aligned}$$

which is impossible, and hence (2.3) holds.] Fix an arbitrary minimizer x^* of $\min_{x\in\mathbb{R}^n}\{F(x):\|x-x_0\|^2\le\alpha+r,b_i^Tx\le\beta_i+s_i\}$. We now split the discussion into two cases: Case 1, $(a+2\lambda_{\min}(A)x_0)^Tv=0$; Case 2, $(a+2\lambda_{\min}(A)x_0)^Tv\ne0$.

Suppose that case 1 holds, i.e., $(a+2\lambda_{\min}(A)x_0)^T v = 0$. Consider $x(t) = x^* + tv$. As $\|x^* - x_0\|^2 < \alpha + r$, there exists $t_0 > 0$ such that $\|x(t_0) - x_0\|^2 = \alpha + r$. Note that $b_i^T x(t_0) = b_i^T (x^* + t_0 v) = b_i^T x^* \le \beta_i + s_i$ and

$$\begin{aligned} F(x(t_0)) &= (x^* + t_0 v)^T (A - \lambda_{\min}(A) I_n)(x^* + t_0 v) \\ &\quad + (a + 2\lambda_{\min}(A)x_0)^T (x^* + t_0 v) + (\gamma - \lambda_{\min}(A)\|x_0\|^2) \\ &= (x^*)^T (A + \lambda_{\min}(A) I_n) x^* + (a+2\lambda_{\min}(A)x_0)^T x^* + (\gamma - \lambda_{\min}(A)\|x_0\|^2) \\ &= F(x^*). \end{aligned}$$

This contradicts our assumption that any minimizer x^* of $\min_{x\in\mathbb{R}^n}\{F(x) : \|x-x_0\|^2 \le \alpha + r, b_i^T x \le \beta_i + s_i\}$ satisfy $\|x^* - x_0\|^2 < \alpha + r$.

Suppose that case 2 holds, i.e., $(a+2\lambda_{\min}(A)x_0)^T v \neq 0$. By replacing v with $-v$ if necessary, we may assume without loss of generality that $(a + 2\lambda_{\min}(A)^T x_0)^T v < 0$. As $\|x^* - x_0\|^2 < \alpha + r$, there exists $t_0 > 0$ such that $\|x(t) - x_0\|^2 \le \alpha + r$ for all $t \in (0, t_0]$. Note that $b_i^T x(t_0) = b_i^T (x^* + t_0 v) = b_i^T x^* \le \beta_i + s_i$ and

$$\begin{aligned} F(x(t_0)) &= (x^* + t_0 v)^T (A - \lambda_{\min}(A) I_n)(x^* + t_0 v) \\ &\quad + (a + 2\lambda_{\min}(A)x_0)^T (x^* + t_0 v) + (\gamma - \lambda_{\min}(A)\|x_0\|^2) \\ &< (x^*)^T (A - \lambda_{\min}(A) I_n) x^* + (a+2\lambda_{\min}(A)x_0)^T x^* + (\gamma - \lambda_{\min}(A)\|x_0\|^2) \\ &= F(x^*). \end{aligned}$$

This contradicts our assumption that x^* is a minimizer. □

As a consequence, we deduce the hidden convexity of the well-known trust region system.

Corollary 2.1 (Polyak [25, Theorem 2.2]) *Let $f(x) = x^T Ax + a^T x + \gamma$ and $g_0(x) = \|x - x_0\|^2 - \alpha$ where $A \in S^{n\times n}$, $a, x_0 \in \mathbb{R}^n$ and $\gamma, \alpha \in \mathbb{R}$. Then, $U(f, g_0)$ is convex.*

Proof Let $b_i = 0$, $i = 1, \ldots, m$ (so, $\dim \text{span}\{b_1, \ldots, b_m\} = 0$). Then the dimension condition (2.2) reduces to $\dim\text{Ker}(A - \lambda_{\min}(A) I_n) \ge 1$ which is always satisfied. So, Theorem 2.1 shows the $U(f, g_0)$ is always convex. □

Remark 2.1 (Observations on the Dimension Condition) We observe that the dimension condition (2.2) in the case of quadratic programs with one linear inequality constraint, i.e. $m = 1$ in (2.2), has been used to establish strong duality in [3]. This conclusion is deduced in Corollary 4.2 of Sect. 4.

Moreover, a close inspection of the proof of Theorem 2.1 suggests that, for a general system of quadratic functions $f(x) = x^T Ax + a^T x + \gamma$, $g_0(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = \|Bx\|^2 + b_i^T x - \beta_i$, $i = 1, \ldots, m$ with $B \in \mathbb{R}^{l\times n}$ for some $l \in \mathbb{N}$, $U(f, g_0, g_1, \ldots, g_m)$ can be shown to be convex under a modified dimension condition. This will be given later in Sect. 6.

3 Exact SDP relaxations

In this section, we establish that a semi-definite relaxation of the model problem (P)

$$\begin{array}{rll}(P) & \min\limits_{x\in\mathbb{R}^n} & x^TAx + a^Tx \\ & \text{s.t.} & \|x - x_0\|^2 \le \alpha, \\ & & b_i^Tx \le \beta_i, \quad i = 1, \ldots, m,\end{array}$$

is exact under the dimension condition. Importantly, it holds without the Slater condition. To formulate a SDP relaxation of (P), let us introduce the following $(n+1)\times(n+1)$ matrices: $M = \left(\begin{smallmatrix} A & a/2 \\ a^T/2 & 0\end{smallmatrix}\right)$,

$$H_0 = \begin{pmatrix} I_n & -x_0 \\ -x_0^T & \|x_0\|^2 - \alpha\end{pmatrix} \quad \text{and} \quad H_i = \begin{pmatrix} 0 & b_i/2 \\ b_i^T/2 & -\beta_i\end{pmatrix}, \quad i = 1, \ldots, m. \quad (3.1)$$

Note that $x^TAx + a^Tx = \mathrm{Tr}(MX)$, $\|x - x_0\|^2 - \alpha = \mathrm{Tr}(H_0X)$ and $b_i^Tx - \beta_i = \mathrm{Tr}(H_iX)$ where $X = \tilde{x}\tilde{x}^T$ with $\tilde{x} = (x^T, 1)^T$. Thus, the model problem can be equivalently rewritten as

$$\begin{array}{ll}\min\limits_{X\in S_+^{n+1}} & \mathrm{Tr}(MX) \\ \text{s.t.} & \mathrm{Tr}(H_0X) \le 0, \\ & \mathrm{Tr}(H_iX) \le 0, \quad i = 1, \ldots, m \\ & X_{n+1,n+1} = 1, \mathrm{rank}(X) = 1,\end{array}$$

where $\mathrm{rank}(X)$ denotes the rank of the matrix X and $X_{n+1,n+1}$ is the element of X that lies in the $n+1^{\text{th}}$ row and $n+1^{\text{th}}$ column. By removing the rank one constraint, we obtain the following semi-definite relaxation of (P)

$$\begin{array}{rll}(SDRP) & \min\limits_{X\in S_+^{n+1}} & \mathrm{Tr}(MX) \\ & \text{s.t.} & \mathrm{Tr}(H_0X) \le 0, \\ & & \mathrm{Tr}(H_iX) \le 0, \quad i = 1, \ldots, m \\ & & X_{n+1,n+1} = 1.\end{array}$$

The semi-definite relaxation problem (SDRP) is a convex program over a matrix space. Its convex dual problem can be stated as follows

$$\begin{aligned}(D) & \max_{\mu\in\mathbb{R},\ \lambda_i\ge 0, i=0,\ldots,m} \left\{\mu : M + \sum_{i=0}^m \lambda_i H_i \succeq \begin{pmatrix} 0 & 0 \\ 0 & \mu\end{pmatrix}\right\} \\ & = \max_{\substack{\lambda_i\ge 0, \\ i=0,\ldots,m}} \min_{x\in\mathbb{R}^n} \left\{x^TAx + a^Tx + \lambda_0\left(\|x - x_0\|^2 - \alpha\right) + \sum_{i=1}^m \lambda_i(b_i^Tx - \beta_i)\right\},\end{aligned}$$

which coincides with the Lagrangian dual problem of (P). Clearly, (SDRP) and (D) are semi-definite linear programming problems and hence can be solved efficiently,

whereas the original problem (P) which is a non-convex quadratic program with multiple constraints, is, in general, a computationally hard problem. Therefore, it is of interest to study when the semi-definite relaxation is exact in the sense that $\min(P) = \min(SDRP)$. For related most recent results on exact SDP relaxations, see [16].

If A is positive semidefinite, then the problem (P) is a convex quadratic optimization problem which is known to enjoy nice properties such as strong duality and exact relaxation. *Therefore, from now on, we assume that A is not positive semidefinite and so, has at least one negative eigenvalue.*

Theorem 3.1 (Exact SDP-relaxation) *Suppose that the dimension condition* (2.2) *is satisfied. Then, the semi-definite relaxation is exact, i.e.,* $\min(P) = \min(SDRP)$.

Proof $[\min(P) = \max(D) < +\infty]$. We first prove that there is no duality gap between (P) and (D) under the dimension condition. It is known that this will follow if we show that the optimal value function of (P)

$$v(s_0, s_1, \ldots, s_m) := \inf_{x \in \mathbb{R}^n} \{x^T Ax + a^T x : \|x - x_0\|^2 \leq \alpha + s_0, \\ b_i^T x \leq \beta_i + s_i,\ i = 1, \ldots, m\},$$

is lower semicontinuous and convex function on $\mathbb{R}^{m+1}$ (See, for instance [20] for details). To see this, we first note that $\text{epi}v = U(f, g_0, g_1, \ldots, g_m)$ where $f(x) = x^T Ax + a^T x$, $g_0(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = b_i^T x - \beta_i$, $i = 1, \ldots, m$. So, by Proposition 2.1, $\text{epi}v$ is a convex set, and so, v is a convex function. The lower semicontinuity of v will follow from Proposition 2.1 as $U(f, g_0, g_1, \ldots, g_m)$ is a closed set.

[$\min(\mathbf{P}) = \min(\mathbf{SDRP})$]. By the construction of the SDP relaxation problem (SDRP) and the dual (D), it is easy see that

$$\min(P) \geq \min(SDRP) \geq \max(D).$$

As there is no duality gap between (P) and (D), we obtain that $\min(P) = \min(SDRP)$.

[**Attainment of Minimum of (SDRP)**] We now show that the minimum in (SDRP) is attained. To see this, we only need to show the feasible set of (SDRP) is bounded. If not, then there exist $X^k \in S_+^{n+1}$ with

$$X^k = \begin{pmatrix} Y^k & y^k \\ y^k & 1 \end{pmatrix}$$

such that $\|X^k\|_F := \sqrt{\text{Tr}(X^k X^k)} \to +\infty$, $\text{Tr}(H_i X^k) \leq 0$, $i = 0, 1, \ldots, m$ where H_i, $i = 0, 1, \ldots, m$ is defined as in (3.1). This implies that

$$0 \leq \text{Tr}(Y^k) \leq -\|x_0\|^2 + \alpha + 2(y^k)^T x_0 \text{ and } b_i^T y^k \leq \beta_i.$$

As $X^k \succeq 0$, we have $Y^k - y^k (y^k)^T \succeq 0$. So,

$$\|y^k\|^2 = \text{Tr}\big(y^k (y^k)^T\big) \leq \text{Tr}(Y^k) \leq -\|x_0\|^2 + \alpha + 2(y^k)^T x_0.$$

So, y^k is bounded, and so $\mathrm{Tr}(Y^k)$ is also a bounded sequence. Thus, both Y^k and y^k are bounded. It follows that X_k is bounded which contradicts the fact that $\|X^k\|_F \to +\infty$. □

It should be noted that convexity of the set $U(f, g_0, g_1, \ldots, g_m)$ plays an important role in establishing the exact SDP relaxation of (P). However, as we see in the following example, the convexity does not imply that problem (P) is equivalent to a convex optimization problem in the sense that they have the same minimizers.

Example 3.1 Consider $f(x) = x^2$, $g_0(x) = x^2 - 1$ and $g_1(x) = -x^2 + 1$. It can be checked that

$$U(f, g_0, g_1) = \{(x^2, x^2 - 1, -x^2 + 1) : x \in \mathbb{R}\} = \{(z, z - 1, -z + 1) : z \geq 0\},$$

which is a closed and convex set. On the other hand, the corresponding optimization problem $\min_{x\in\mathbb{R}}\{x^2 : x^2 - 1 \leq 0, -x^2 + 1 \leq 0\}$ cannot be equivalent to a convex optimization problem as its solution set is $\{-1, 1\}$ which is not a convex set.

One interesting feature of our SDP relaxation result is that its exactness is independent of the Slater condition. The following example illustrates that our SDP relaxation may be exact while the Slater condition fails.

Example 3.2 (Exact SDP-relaxation without the Slater condition) Consider the three dimensional quadratic optimization problem with two linear inequalities:

$$\begin{array}{lll} (EP) & \min\limits_{(x_1,x_2,x_3)\in\mathbb{R}^3} & -x_1^2 - x_2^2 - x_3^2 + 3x_1 + 2x_2 + 2x_3 \\ & \text{s.t.} & (x_1 - 1)^2 + x_2^2 + x_3^2 \leq 1, \\ & & x_1 \leq 0, \\ & & x_1 + x_2 + x_3 \leq 0. \end{array}$$

This can be written as our model problem where $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $a = (3, 2, 2)^T$, $x_0 = (1, 0, 0)^T$, $\alpha = 1$, $b_1 = (1, 0, 0)^T$, $b_2 = (1, 1, 1)$ and $\beta_1 = \beta_2 = 0$. Clearly, the only feasible point is $(0, 0, 0)$ and so, $\min(EP) = 0$. We also note that the Slater condition fails. Let $s = \dim \mathrm{span}\{b_1, b_2\} = 2$. We see that

$$\dim\mathrm{Ker}(A - \lambda_{\min}(A) I_n) = 3 = s + 1.$$

So, the dimension condition is satisfied.

On the other hand, the SDP-relaxation of (EP) is given by

$$\begin{array}{lll} (SDRP_E) & \min\limits_{X\in S^4} & -z_1 + 3z_4 - z_5 + 2z_7 - z_8 + 2z_9 \\ & \text{s.t.} & z_1 - 2z_4 + z_5 + z_8 \leq 0 \\ & & z_4 \leq 0 \\ & & z_4 + z_7 + z_9 \leq 0 \\ & & X = \begin{pmatrix} z_1 & z_2 & z_3 & z_4 \\ z_2 & z_5 & z_6 & z_7 \\ z_3 & z_6 & z_8 & z_9 \\ z_4 & z_7 & z_9 & 1 \end{pmatrix} \succeq 0. \end{array}$$

Since $z_1 = z_2 = \ldots = z_9 = 0$ is feasible for $(SDRP_E)$, $\min(SDRP_E) \leq 0$. Moreover, for each feasible $X = \begin{pmatrix} z_1 & z_2 & z_3 & z_4 \\ z_2 & z_5 & z_6 & z_7 \\ z_3 & z_6 & z_8 & z_9 \\ z_4 & z_7 & z_9 & 1 \end{pmatrix} \succeq 0$, we have $z_1 \geq 0,\ z_5 \geq 0,$

$$z_8 \geq z_9^2 \geq 0 \quad \textit{and} \quad z_5 \geq z_7^2 \geq 0. \tag{3.2}$$

This gives us that

$$-2z_4 \leq z_1 - 2z_4 + z_5 + z_8 \leq 0$$

and so, $z_4 \geq 0$. As $z_4 \leq 0$, we have $z_4 = 0$ and so, $z_1 + z_5 + z_8 \leq 0$. Hence, $z_1 = z_5 = z_8 = 0$ and $z_7 = z_9 = 0$ (by (3.2)). Thus, $\min(SDRP_E) = 0 = \min(EP)$.

In the following, we use a simple one-dimensional quadratic optimization problem to show that the SDP relaxation may not be exact if our sufficient dimension condition (2.2) is not satisfied.

Example 3.3 (Importance of sufficient dimension condition) Consider the minimization problem

$$(EP_1)\ \min_{x\in\mathbb{R}}\{f(x) : g_0(x) \leq 0, g_1(x) \leq 0\},$$

where $f(x) = x - x^2$, $g_0(x) = x^2 - 1$, $g_1(x) = -x$, $n = 1$ and $m = 1$. Then, $f(x) = x^T Ax + a^T x + r$ with $A = -1$, $a = 1$ and $r = 0$, $g_0(x) = \|x - x_0\|^2 - \alpha$ with $x_0 = 0$, $\alpha = 1$ and $g_1(x) = b_1^T x - \beta_1$ with $b_1 = 1$ and $\beta_1 = 0$. Clearly, $\dim\mathrm{Ker}(A - \lambda_{\min}(A)I_n) = 1 < 2 = \dim \mathrm{span}\{b_1\} + 1$.

The SDP relaxation of (EP_1) is given by

$$\begin{array}{lll} (SDRP_{E1}) & \min\limits_{X\in S^2} & -z_1 + z_2 \\ & \text{s.t.} & z_1 - 1 \leq 0 \\ & & -z_2 \leq 0 \\ & & X = \begin{pmatrix} z_1 & z_2 \\ z_2 & 1 \end{pmatrix} \succeq 0. \end{array}$$

It can be easily verified that $\min(EP_1) = 0$ and $\min(SDRP_{E1}) = -1$. Thus, the SDP relaxation of (EP_1) is not exact.

Consider the quadratic optimization problem with one norm constraint and a rank-one quadratic inequality constraint:

$$\begin{array}{lll} (P_0) & \min\limits_{x\in\mathbb{R}^n} & x^T Ax + a^T x \\ & \text{s.t.} & \|x - x_0\|^2 \leq \alpha, \\ & & (b^T x)^2 \leq r, \end{array}$$

where $A \in S^{n\times n}$, $a, x_0, b \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ and$r \geq 0$.

Model problems of this form arise from the application of the trust-region method for the minimization of a nonlinear function with a discrete constraint. For instance, consider the trust-region approximation problem

$$\begin{array}{ll}\min\limits_{x\in\mathbb{R}^n} & x^TAx + a^Tx \\ \text{s.t.} & \|x - x_0\|^2 \le \alpha, \\ & b^Tx \in \{1, -1\}.\end{array}$$

The continuous relaxation of this problem becomes

$$\begin{array}{ll}\min\limits_{x\in\mathbb{R}^n} & x^TAx + a^Tx \\ \text{s.t.} & \|x - x_0\|^2 \le \alpha, \\ & -1 \le b^Tx \le 1,\end{array}$$

which is, in turn, equivalent to (P_0) with $r = 1$.

The SDP-relaxation of (P_0) is given by

$$\begin{array}{lll}(SDRP_0) & \min\limits_{X\in S_+^{n+1}} & \mathrm{Tr}(\tilde{M}X) \\ & \text{s.t.} & \mathrm{Tr}(\tilde{H}_0X) \le 0 \\ & & \mathrm{Tr}(\tilde{H}_iX) \le 0, \quad i = 1, 2 \\ & & X_{n+1,n+1} = 1.\end{array}$$

where

$$\tilde{M} = \begin{pmatrix} A & a/2 \\ a^T/2 & 0 \end{pmatrix}, \tilde{H}_0 = \begin{pmatrix} I_n & -x_0 \\ -x_0 & \|x_0\|^2 - \alpha \end{pmatrix}$$
$$H_1 = \begin{pmatrix} 0 & b/2 \\ b^T/2 & -\sqrt{r} \end{pmatrix} \text{ and } H_2 = \begin{pmatrix} 0 & -b/2 \\ -b^T/2 & \sqrt{r} \end{pmatrix}.$$

We now obtain the following exact SDP-relaxation result for the problem (P_0) under a dimension condition.

Corollary 3.1 (Trust-region model with rank-one constraint) *Suppose that* $\dim\mathrm{Ker}(A - \lambda_{\min}(A)I_n) \ge 2$. *Then, the semi-definite relaxation is exact for* (P_0), *i.e.,* $\min(P_0) = \min(SDRP_0)$.

Proof Note that $(b^Tx)^2 \le r$ is equivalent to $-\sqrt{r} \le b^Tx \le \sqrt{r}$. In this case the dimension condition of Theorem 3.1 reduces to the assumption that

$$\dim\mathrm{Ker}(A - \lambda_{\min}(A)I_n) \ge \dim \mathrm{span}\{b, -b\} + 1.$$

The conclusion follows from Theorem 3.1 and the fact that $\dim \mathrm{span}\{b, -b\} \le 1$. □

Remark 3.1 (Approximate S-lemma and SDP Relaxations) For a general homogeneous quadratic optimization problem with multiple convex quadratic constraints, an

estimate for the ratio between the optimal value of the underlying quadratic optimization and its associated SDP relaxation problem has been given in [7] (see "Appendix"). This result is known as an approximate S-lemma as it provides the approximate ratio from the SDP relaxation to the underlying problem. Clearly, Theorem 3.1 shows that the ratio between the optimal value of the underlying quadratic optimization problem and its associated SDP relaxation problem is one for the extended trust region problem (P), under the dimension condition. For other nonconvex quadratic optimization problems where the corresponding ratio also equals one, see [30].

Consider the quadratic optimization problem with the constraint set described by the intersection of an Euclidean ball and a box:

$$\begin{aligned}(P_1)\ \min_{x\in\mathbb{R}^n}\ & x^TAx + a^Tx\\ \text{s.t.}\ & \|x\|^2 \le 1,\\ & -l_i \le x_i \le l_i, \quad i = 1,\dots,n,\end{aligned}$$

where $l_i > 0$. This class of nonconvex quadratic problems is known to be NP-hard. Indeed, when $l_i \le \frac{1}{2}$ and A is negative definite, the norm constraint, $\|x\|^2 \le 1$, becomes superfluous and so, the problem (P_1) reduces to the quadratic concave minimization problem with bounded constraints which is an NP-hard problem (cf. [7]). Using the approximate S-lemma of [7] and a semidefinite programming relaxation, one can find an estimate for the value of the nonconvex quadratic problem (P_1).

We note that our dimension condition fails for (P_1). To see this, take $m = 2n$, $b_i = e_i$, $i = 1,\dots,n$ and $b_i = -e_i$, $i = n+1,\dots,2n$. Then, we see that $\dim \operatorname{span}\{b_1,\dots,b_n\} = n$ in this case, and so, the dimension condition reduces to $\dim\operatorname{Ker}(A - \lambda_{\min}(A)I_n) \ge n+1$ which is impossible.

On the other hand, consider the following semi-definite relaxation of (P_1) (see [7])

$$\begin{aligned}(SDRP_1)\ \min_{X\in S_+^{n+1}}\ & \operatorname{Tr}(MX)\\ \text{s.t.}\ & \operatorname{Tr}(H_0X) \le 1,\\ & \operatorname{Tr}(H_iX) \le 1, \quad i = 1,\dots,n\\ & \operatorname{Tr}(H_{2n+1}X) \le 1,\end{aligned}$$

where

$$M = \begin{pmatrix} A & a/2 \\ a^T/2 & 0 \end{pmatrix},\ H_0 = \begin{pmatrix} I_n & 0 \\ 0 & 0 \end{pmatrix} \tag{3.3}$$

$$H_i = \begin{pmatrix} \frac{1}{l_i^2}\operatorname{diag}(e_i) & 0 \\ 0 & 0 \end{pmatrix},\ i = 1,\dots,n, \quad \text{and} \quad H_{n+1} = \begin{pmatrix} 0_{n\times n} & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.4}$$

Following [7], we can get

$$2\log(6n+6)\min(P_1) \le \min(SDRP_1) \le \min(P_1) \le 0.$$

To see this, we first note that $\min(P_1)$ equals the optimal value of the following optimization problem

$$\begin{aligned} \min_{(x,t)\in\mathbb{R}^n\times\mathbb{R}} \quad & x^T Ax + ta^T x \\ \text{s.t.} \quad & \|x\|^2 \le 1, \\ & \tfrac{1}{l_i^2} x_i^2 \le 1, \quad i = 1, \ldots, n, \\ & t^2 \le 1. \end{aligned}$$

which is, in turn, equal to the negative of the optimal value of the following quadratically constrained quadratic problem

$$(QCQ_1) \quad \max_{y=(x^T,t)^T\in\mathbb{R}^n\times\mathbb{R}} \left\{ -y^T My : y^T H_0 y \le 1, \, y^T H_i y \le 1, \quad i = 1, \ldots, n+1 \right\}$$

where M and H_i, $i = 0, 1, \ldots, n+1$ are defined as in (3.3) and (3.4). Note that $\operatorname{rank} H_i = 1$, $i = 1, 2, \ldots, n+1$ and $\sum_{i=0}^{n+1} H_i \succ 0$. So, [7, Lemma A.6, Approximate S-lemma] implies that

$$\begin{aligned} \max(QCQ_1) \le \min(SDP_1) &\le 2\log\left(6\sum_{i=1}^{n+1} \operatorname{rank} H_i\right) \max(QCQ_1) \\ &= 2\log(6n+6)\max(QCQ_1), \end{aligned} \tag{3.5}$$

where (SDP_1) is given by

$$(SDP_1) \quad \min_{\mu_0,\ldots,\mu_{n+1}\ge 0} \left\{ \sum_{i=0}^{n+1} \mu_i : M + \sum_{i=0}^{n+1} \mu_k H_k \succeq 0 \right\}.$$

It can be verified that (SDP_1) is the Lagrange dual problem of the semi-definite problem

$$\begin{aligned} (SP_1) \quad \max_{X\in S_+^{n+1}} \quad & \operatorname{Tr}(-MX) \\ \text{s.t.} \quad & \operatorname{Tr}(H_0 X) \le 1, \\ & \operatorname{Tr}(H_i X) \le 1, \quad i = 1, \ldots, n \\ & \operatorname{Tr}(H_{n+1} X) \le 1, \end{aligned}$$

and Slater condition holds for (SP_1). So, $\min(SDP_1) = \max(SP_1)$. Finally, the conclusion follows from (3.5) by noting that $\max(SP_1) = -\min(SDRP_1)$.

4 Global optimality and strong duality

In this section, we present a necessary and sufficient condition for global optimality of (P) and consequently, obtain strong duality between (P) and (D) whenever the dimension condition is satisfied and Slater's condition holds for (P). Related global optimality and duality results for nonconvex quadratic optimization can be found in [14,17,19,23].

Theorem 4.1 (Necessary and sufficient global optimality condition) *For (P), suppose that there exists* $\overline{x} \in \mathbb{R}^n$ *with* $\|\overline{x} - x_0\|^2 < \alpha$ *and* $b_i^T \overline{x} < \beta_i$, $i = 1, \ldots, m$, *and that the dimension condition (2.2) is satisfied. Let* x^* *be a feasible point of (P). Then,* x^* *is a global minimizer of (P) if and only if there exists* $(\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}_+$ *such that the following condition holds:*

$$\begin{cases} 2\big((A + \lambda_0 I_n)x^*\big) = -\big(a + 2\lambda_0(x^* - x_0) \\ \quad + \sum_{i=1}^m \lambda_i b_i\big), & \textit{(KKT Condition)} \\ \lambda_0(\|x^* - x_0\|^2 - \alpha) = 0 \textit{ and } \lambda_i(b_i^T x^* - \beta_i) = 0, \\ \quad i = 1, \ldots, m, & \textit{(Complementary Slackness)} \\ A + \lambda_0 I_n \succeq 0, & \textit{(Second Order Condition).} \end{cases}$$

Proof [**Necessary condition for optimality**]. Let x^* be a global minimizer of (P). Then, the following inequality system has no solution:

$$\|x - x_0\|^2 \leq \alpha,\ b_i^T x \leq \beta_i,\ i = 1, \ldots, m,\ x^T A x + a^T x < (x^*)^T A x^* + a^T x^*.$$

In particular, letting $\gamma = -((x^*)^T A x^* + a^T x^*)$, the following inequality system also has no solution:

$$\|x - x_0\|^2 < \alpha,\ b_i^T x < \beta_i,\ i = 1, \ldots, m,\ x^T A x + a^T x + \gamma < 0.$$

Then, $0 \notin \text{int}U(f, g_0, g_1, \ldots, g_m)$, where

$$U(f, g_0, g_1, \ldots, g_m) := \{(f(x), g_0(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\} + \mathbb{R}^{m+2}_+$$

is a convex set by proposition 2.1. Moreover, as f, g_i are all continuous, we see that

$$\{(f(x), g_0(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\} + \text{int}\mathbb{R}^{m+2}_+ = \text{int}U(f, g_0, g_1, \ldots, g_m)$$

is also convex.

Now, by the convex separation theorem, there exists $(\mu, \tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_m) \in \mathbb{R}^{m+2}_+ \backslash \{0\}$ such that, for all $x \in \mathbb{R}^n$,

$$\mu\left(x^T A x + a^T x + \gamma\right) + \tilde{\lambda}_0(\|x - x_0\|^2 - \alpha) + \sum_{i=1}^m \tilde{\lambda}_i(b_i^T x - \beta_i) \geq 0.$$

By the strict feasibility condition, we see that $\mu \neq 0$. Thus, for all $x \in \mathbb{R}^n$

$$x^T A x + a^T x + \gamma + \lambda_0\left(\|x - x_0\|^2 - \alpha\right) + \sum_{i=1}^m \lambda_i(b_i^T x - \beta_i) \geq 0.$$

where $\lambda_i = \frac{\tilde{\lambda}_i}{\mu}$, $i = 0, 1, \ldots, m$. Letting $x = x^*$, we see that

$$\lambda_0 \left(\|x^* - x_0\|^2 - \alpha \right) + \sum_{i=1}^{m} \lambda_i \left(b_i^T x^* - \beta_i \right) \geq 0.$$

As x^* is feasible for (P), it follows that

$$\lambda_0 \left(\|x^* - x_0\|^2 - \alpha \right) = 0 \text{ and } \lambda_i \left(b_i^T x^* - \beta_i \right) = 0, \quad i = 1, \ldots, m.$$

Let $h(x) := x^T Ax + a^T x + \lambda_0 \left(\|x - x_0\|^2 - \alpha \right) + \sum_{i=1}^{m} \lambda_i (b_i^T x - \beta_i)$. Then, we see that x^* is a global minimizer of h, and so, $\nabla h(x^*) = 0$ and $\nabla^2 h(x^*) \succeq 0$. That is to say,

$$2 \left(A + \lambda_0 I_n \right) x^* + \left(a + 2\lambda_0 (x^* - x_0) + \sum_{i=1}^{m} \lambda_i b_i \right) = 0 \quad \text{and} \quad A + \lambda_0 I_n \succeq 0.$$

[**Sufficient condition for optimality**] Conversely, if the optimality condition holds, then we see that $h(x) := x^T Ax + a^T x + \lambda_0 (\|x - x_0\|^2 - \alpha) + \sum_{i=1}^{m} \lambda_i (b_i^T x - \beta_i)$ is convex with $\nabla h(x^*) = 0$ and $\nabla^2 h(x^*) \succeq 0$. So, x^* is a global minimizer of h, and hence, for all feasible point $x \in \mathbb{R}^n$ of (P),

$$\begin{aligned} x^T Ax + a^T x &\geq x^T Ax + a^T x + \lambda_0 \left(\|x - x_0\|^2 - \alpha \right) + \sum_{i=1}^{m} \lambda_i \left(b_i^T x - \beta_i \right) \\ &\geq \left(x^* \right)^T Ax^* + a^T x^* + \lambda_0 \left(\|x^* - x_0\|^2 - \alpha \right) + \sum_{i=1}^{m} \lambda_i \left(b_i^T x^* - \beta_i \right) \\ &= \left(x^* \right)^T Ax^* + a^T x^*, \end{aligned}$$

where the last equality follows by the complementary condition. Thus, x^* is a global minimizer of (P). □

Consider the Lagrangian dual problem of (P):

$$(D) \quad \max_{\lambda_i \geq 0} \min_{x \in \mathbb{R}^n} \left\{ x^T Ax + a^T x + \lambda_0 \left(\|x - x_0\|^2 - \alpha \right) + \sum_{i=1}^{m} \lambda_i (b_i^T x - \beta_i) \right\}.$$

We now show that the strong duality holds under the dimension condition together with the Slater condition.

Corollary 4.1 (Strong Duality) *Suppose that there exists* $\overline{x} \in \mathbb{R}^n$ *with* $\|\overline{x} - x_0\|^2 < \alpha$ *and* $b_i^T \overline{x} < \beta_i$, $i = 1, \ldots, m$, *and that the dimension condition* (2.2) *is satisfied.*

Then, strong duality holds, i.e.,

$$\min_{x\in\mathbb{R}^n}\left\{x^TAx+a^Tx : \|x-x_0\|^2\le\alpha,\ b_i^Tx\le\beta_i,\ i=1,\ldots,m\right\}$$
$$=\max_{\lambda_i\ge 0}\min_{x\in\mathbb{R}^n}\left\{x^TAx+a^Tx+\lambda_0(\|x-x_0\|^2-\alpha)+\sum_{i=1}^m\lambda_i(b_i^Tx-\beta_i)\right\}. \quad (4.1)$$

where the maximum in (4.1) is attained.

Proof First of all, we note that the following weak duality always holds:

$$\min_{x\in\mathbb{R}^n}\left\{x^TAx+a^Tx : \|x-x_0\|^2\le\alpha,\ b_i^Tx\le\beta_i,\ i=1,\ldots,m\right\}$$
$$\ge\max_{\lambda_i\ge 0}\min_{x\in\mathbb{R}^n}\left\{x^TAx+a^Tx+\lambda_0(\|x-x_0\|^2-\alpha)+\sum_{i=1}^m\lambda_i(b_i^Tx-\beta_i)\right\}.$$

To see the reverse inequality, let x^* be a minimizer of $\min_{x\in\mathbb{R}^n}\{x^TAx+a^Tx : \|x-x_0\|^2\le\alpha,\ b_i^Tx\le\beta_i,\ i=1,\ldots,m\}$. Then, by Theorem 4.1, there exists $(\lambda_0,\lambda_1,\ldots,\lambda_m)\in\mathbb{R}^{m+1}_+$ such that the following condition holds:

$$\begin{cases} 2(A+\lambda_0I_n)x^*=-\big(a+2\lambda_0(x^*-x_0)+\sum_{i=1}^m\lambda_ib_i\big),\\ \lambda_0(\|x^*-x_0\|^2-\alpha)=0 \text{ and } \lambda_i(b_i^Tx^*-\beta_i)=0,\ i=1,\ldots,m,\\ A+\lambda_0I_n\succeq 0 \end{cases}.$$

Then we see that $h(x):=x^TAx+a^Tx+\lambda_0(\|x-x_0\|^2-\alpha)+\sum_{i=1}^m\lambda_i(b_i^Tx-\beta_i)$ is convex with $\nabla h(x^*)=0$ and $\nabla^2h(x^*)\succeq 0$. So, x^* is a global minimizer of h, and hence, for all $x\in\mathbb{R}^n$

$$x^TAx+a^Tx+\lambda_0\left(\|x-x_0\|^2-\alpha\right)+\sum_{i=1}^m\lambda_i\left(b_i^Tx-\beta_i\right)$$
$$\ge\left(x^*\right)^TAx^*+a^Tx^*+\lambda_0\left(\|x^*-x_0\|^2-\alpha\right)+\sum_{i=1}^m\lambda_i\left(b_i^Tx^*-\beta_i\right)$$
$$=\left(x^*\right)^TAx^*+a^Tx^*.$$

Thus, the reverse inequality is true and the maximum in (4.1) is attained. So, the conclusion follows. □

It is easy to see that, for the extended trust-region model problem with linear inequality constraints, our Corollary 4.1 shows that the ratio between the optimal value of the underlying problem and its associated SDP relaxation problem is one whenever the dimension condition is satisfied. For other quadratic optimization problems where the approximate ratio, is one see [30].

Consider the following nonconvex quadratic optimization problem subject to a norm constraint and a linear constraint:

$$(P_2)\ \min_{x\in\mathbb{R}^n}\ x^TAx + a^Tx$$
$$\text{s.t.}\ \ \|x - x_0\|^2 \le \alpha,$$
$$b_1^Tx \le \beta_1,$$

where $b_1 \in \mathbb{R}^n$ and $\beta_1 \in \mathbb{R}$.

As a corollary of Theorem 4.1, we now establish strong duality for (P_1) which was established in [3].

Corollary 4.2 (Trust-region model with single linear constraint) [3, Theorem 3.6] *For problem* (P_1)*, suppose that* $\dim\big(\mathrm{Ker}(A - \lambda_{\min}(A)I_n)\big) \ge 2$ *and suppose that there exists* $\overline{x}$ *such that* $\|\overline{x} - x_0\|^2 < \alpha$ *and* $b_1^T\overline{x} < \beta_1$*. Then, strong duality holds for problem* (P_1)*, i.e.,*

$$\min_{x\in\mathbb{R}^n}\left\{x^TAx + a^Tx : \|x - x_0\|^2 \le \alpha,\ b_1^Tx \le \beta_1\right\}$$
$$= \max_{\lambda_0,\lambda_1\ge 0}\min_{x\in\mathbb{R}^n}\left\{x^TAx + a^Tx + \lambda_0(\|x - x_0\|^2 - \alpha) + \lambda_1(b_1^Tx \le \beta_1)\right\}, \qquad (4.2)$$

and the maximum in (4.2) *is attained.*

Proof The conclusion follows by letting $l = 1$ in Corollary 4.1 and noting that $s = \mathrm{span}\{b_1\} \le 1$. □

Let us note that, if the Slater condition is not satisfied, strong duality may fail while the SDP relaxation is exact. Indeed, the same problem discussed in Example 3.2 can be used to illustrate this situation.

Example 4.1 (Exact SDP-relaxation without Strong duality) Consider the same problem in Example 3.2:

$$(EP)\ \min_{(x_1,x_2,x_3)\in\mathbb{R}^3}\ -x_1^2 - x_2^2 - x_3^2 + 3x_1 + 2x_2 + 2x_3$$
$$\text{s.t.}\ \ (x_1 - 1)^2 + x_2^2 + x_3^2 \le 1,$$
$$x_1 \le 0,$$
$$x_1 + x_2 + x_3 \le 0.$$

We have already shown that $\min(EP) = 0$, the Slater condition fails for (EP) and the SDP relaxation of (EP) is exact. We now show that strong duality fails. The Lagrangian dual problem of (EP) is

$$\max_{\lambda_0,\lambda_1\ge 0}\min_{(x_1,x_2,x_3)\in\mathbb{R}^3}\Big\{-x_1^2 - x_2^2 - x_3^2 + 3x_1 + 2x_2 + 2x_3 + \lambda_0\big((x_1 - 1)^2 + x_2^2 + x_3^2 - 1\big) + \lambda_1x_1 + \lambda_2(x_1 + x_2 + x_3)\Big\}$$
$$= \max_{\lambda_0,\lambda_1\ge 0}\min_{(x_1,x_2,x_3)\in\mathbb{R}^3}\Big\{(\lambda_0 - 1)x_1^2 + (\lambda_1 + \lambda_2 - 2\lambda_0 + 3)x_1 + (\lambda_0 - 1)x_2^2 + (2 + \lambda_2)x_2 + (\lambda_0 - 1)x_3^2 + (2 + \lambda_2)x_3\Big\}.$$

For each $\lambda_0, \lambda_1 \geq 0$,

$$\begin{aligned}\min_{(x_1,x_2,x_3)\in\mathbb{R}^3} &\left\{(\lambda_0-1)x_1^2+(\lambda_1-2\lambda_0+3)x_1+(\lambda_0-1)x_2^2+(2+\lambda_2)x_2+(\lambda_0-1)x_3^2\right.\\ &\left.+(2+\lambda_2)x_3\right\}\\ &= \begin{cases} -\infty, & \text{if } \lambda_0 < 1,\\ -\infty, & \text{if } \lambda_0 = 1,\\ <0, & \text{if } \lambda_0 > 1.\end{cases}\end{aligned}$$

Hence, strong duality fails.

As a consequence of our strong duality theorem, we derive a dual characterization for the non-negativity of a nonconvex quadratic function over the extended trust-region constraints. This characterization can be regarded as a form of the celebrated S-lemma [5]. See "Appendix" for variants of S-lemma.

Corollary 4.3 (S-lemma for extended trust-regions) *Let $x_0, a, b_i \in \mathbb{R}^n$ and $\gamma, \beta_i, \alpha \in \mathbb{R}$, $i = 1, \ldots, m$. Suppose that there exists $\overline{x} \in \mathbb{R}^n$ with $\|\overline{x} - x_0\|^2 < \alpha$ and $b_i^T\overline{x} < \beta_i$, $i = 1, \ldots, m$, and that the dimension condition (2.2) is satisfied. Then, the following statements are equivalent:*

(1) $\|x - x_0\|^2 - \alpha \leq 0,\ b_i^T x - \beta_i \leq 0,\ i = 1, \ldots, m \ \Rightarrow\ x^T Ax + a^T x + \gamma \geq 0.$
(2) $(\exists \lambda_i \geq 0, i = 0, 1, \ldots, m)(\forall x \in \mathbb{R}^n)$

$$(x^T Ax + a^T x + \gamma) + \lambda_0(\|x - x_0\|^2 - \alpha) + \sum_{i=1}^m \lambda_i(b_i^T x - \beta_i) \geq 0.$$

Proof We only need to show (1) $\Rightarrow$ (2) as the converse implication always holds. To see this, suppose (1) holds. Then, the optimal value of the following optimization problem is greater than $-\gamma$

$$\min_{x\in\mathbb{R}^n}\left\{x^T Ax + a^T x : \|x - x_0\|^2 \leq \alpha, b_i^T x \leq \beta_i, \quad i = 1, \ldots, m\right\}.$$

Then, Corollary 4.1 implies that

$$\begin{aligned}&\min_{x\in\mathbb{R}^n}\left\{x^T Ax + a^T x : \|x - x_0\|^2 \leq \alpha,\ b_i^T x \leq \beta_i, \quad i = 1, \ldots, m\right\}\\ &= \max_{\lambda_i\geq 0}\min_{x\in\mathbb{R}^n}\left\{x^T Ax + a^T x + \lambda_0(\|x - x_0\|^2 - \alpha) + \sum_{i=1}^m \lambda_i(b_i^T x - \beta_i)\right\} \geq -\gamma,\end{aligned} \tag{4.3}$$

and the maximum in (4.3) is attained. So, (2) follows. □

Recall that the celebrated S-lemma states that, for two quadratic functions f, g, $[g(x) \leq 0 \Rightarrow f(x) \geq 0]$ is equivalent to the existence of $\lambda \geq 0$ such that $f + \lambda g$ is always nonnegative. Note that, in the case where $b_i = 0$ and $\beta_i = 1$, the dimension condition is always satisfied as $\dim\mathrm{Ker}(A - \lambda_{\min}(A)I_n) \geq 1$ and $\dim \mathrm{span}\{b_1, \ldots, b_m\} = 0$, and so, the above corollary reduces to the S-lemma in the case where $g = \|x - x_0\|^2 - \alpha$.

It is worth noting that in Corollary 4.3, the strict feasibility condition cannot be dropped even if the dimension condition is satisfied. To see this, consider the following one-dimensional quadratic functions $f(x) = x$ and $g_0(x) = x^2$. It can be verified that the dimension condition is satisfied and $[g_0(x) \leq 0 \Rightarrow f(x) \geq 0]$. On the other hand, for any $\lambda \geq 0$,

$$\inf_{x\in\mathbb{R}} \{f(x) + \lambda g(x)\} = \begin{cases} -\frac{1}{4\lambda} < 0, & \text{if } \lambda > 0, \\ -\infty, & \text{if } \lambda = 0. \end{cases}$$

Therefore, Corollary 4.3 can fail if the strict feasibility condition is not satisfied.

On the other hand, if the strict feasibility condition fails, we now show that a new form of asymptotic S-lemma still holds. For related asymptotic S-lemma of this form for general quadratic constraint without Slater condition see [17].

Corollary 4.4 (Asymptotic S-lemma) *Let $A \in S^n$, $x_0, a, b_i \in \mathbb{R}^n$ and $\gamma, \beta_i, \alpha \in \mathbb{R}$, $i = 1, \ldots, m$ with $\{x : \|x - x_0\|^2 \leq \alpha,\ b_i^T x \leq \beta_i,\ i = 1, \ldots, m\} \neq \emptyset$. Suppose that the dimension condition (2.2) is satisfied.*

Then, the following statements are equivalent:

(1) $\|x - x_0\|^2 - \alpha \leq 0,\ b_i^T x - \beta_i \leq 0,\ i = 1, \ldots, m \Rightarrow x^T Ax + a^T x + \gamma \geq 0.$

(2) $(\forall \epsilon > 0)(\exists \lambda_i \geq 0, i = 0, 1, \ldots, m)(\forall x \in \mathbb{R}^n)$

$$(x^T Ax + a^T x + \gamma) + \lambda_0(\|x - x_0\|^2 - \alpha) + \sum_{i=1}^{m} \lambda_i (b_i^T x - \beta_i) + \epsilon \geq 0.$$

Proof [(1) $\Rightarrow$ (2)] Suppose that (1) holds. Let $f(x) = x^T Ax + a^T x + \gamma$, $g_0(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = b_i^T x - \beta_i$, $i = 1, \ldots, m$. Then, for each $\epsilon > 0$, $(-\epsilon, 0, 0, \ldots, 0) \notin U(f, g_0, g_1, \ldots, g_m)$. As the dimension condition (2.2) holds, it follows from Proposition 2.1 and Theorem 2.1 that $U(f, g_0, g_1, \ldots, g_m)$ is a closed convex set. So, the strong separation theorem gives us that $(\mu, \overline{\lambda}_0, \overline{\lambda}_1, \ldots, \overline{\lambda}_m) \in \mathbb{R}_+^{m+2}\backslash\{0\}$ and $\delta \in \mathbb{R}$ such that

$$-\mu\epsilon < \delta \leq \mu f(x) + \sum_{i=0}^{m} \overline{\lambda}_i g_i(x) \quad \text{for all } x \in \mathbb{R}^n.$$

Then, $\mu > 0$. Otherwise, $\mu = 0$. Then, $\sum_{i=0}^m \lambda_i g_i(x) \geq \delta > 0$ for all $x \in \mathbb{R}^n$. This is impossible as $\sum_{i=0}^m \lambda_i g_i(a) \leq 0$ for all $a \in \{x : g_i(x) \leq 0, i = 0, 1, \ldots, m\}$. So, (2) follows with $\lambda_i = \frac{\overline{\lambda}_i}{\mu}$, $i = 0, 1, \ldots, m$.

[(2) $\Rightarrow$ (1)] For any x with $g_i(x) \leq 0$, then (2) implies that for each $\epsilon > 0$, there exist $\lambda_i \geq 0$ such that for all $x \in \mathbb{R}^n$,

$$0 \leq f(x) + \sum_{i=0}^{m} \lambda_i g_i(x) + \epsilon \leq f(x) + \epsilon.$$

Letting $\epsilon \to 0$, we see that $f(x) \geq 0$, and so, (1) follows. □

Before we end this section, let us use the preceding example to illustrate the new form of asymptotic S-lemma.

Example 4.2 (Example illustrating the asymptotic S-lemma) Consider the following one-dimensional quadratic functions $f(x) = x$ and $g_0(x) = x^2$. It can be easily checked that $[x^2 \leq 0 \Rightarrow x \geq 0]$ and the dimension condition is satisfied. Now, for each $\epsilon > 0$, $x + \frac{1}{4\epsilon}x^2 + \epsilon = \left(\frac{1}{2\sqrt{\epsilon}}x + \sqrt{\epsilon}\right)^2 \geq 0$. So, our form of asymptotic S-lemma holds.

5 Applications to robust optimization

In this section, we establish SDP characterizations of the solution of a robust least squares problem (LSP) as well as a robust second order cone programming problem (SOCP) where the uncertainty set is given by the intersection of the norm constraint and the polyhedral constraint. Consequently, we show that solving the robust (LSP) or a robust (SOCP) is equivalent to solving a semi-definite linear programming problem and so the solution can be validated in polynomial time.

Let us note first that, for a $(p \times q)$ matrix M, $\mathrm{vec}(M)$ denotes the vector in $\mathbb{R}^{pq}$ obtained by stacking the columns of M. The tensor product of I_n and a matrix $M \in \mathbb{R}^{p\times p}$ is defined by

$$I_n \otimes M := \begin{pmatrix} M & 0 & 0 & 0 & 0 \\ 0 & M & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & \ddots & M & 0 \\ 0 & 0 & \dots & 0 & M \end{pmatrix} \in \mathbb{R}^{np\times np}.$$

Consider the uncertainty set which is described by a matrix norm constraint and polyhedral constraints, i.e.,

$$\mathcal{U} = \{\tilde{A}^{(0)} + \Delta : \Delta \in \mathbb{R}^{k\times(n+1)}, \|\Delta - \overline{\Delta}\|_F \leq \rho,\ (w^j)^T \mathrm{vec}\Delta \leq \beta^j,\ j = 1, \dots, l\}, \tag{5.1}$$

where $\tilde{A}^{(0)} := (A^{(0)}, a^{(0)}) \in \mathbb{R}^{k\times n} \times \mathbb{R}^k = \mathbb{R}^{k\times(n+1)}$ is the data of a given model, examined in this Section (see Sects. 5.1, 5.2), and $\|M\|_F$ is the Frobenius norm defined by $\|M\|_F = \sqrt{\mathrm{Tr}(M^T M)}$. In the special case when $l = 2$, $w^2 = -w^1$ and $\beta^1 = -\beta^2 = 1$, this uncertainty set reduces to an intersection of two ellipsoids which was examined in [4].

We say $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}$ is robust feasible for the quadratic constraint of the form $\|Ax - a\|^2 \leq \lambda$ with respect to the uncertainty set $\mathcal{U}$ whenever $\max_{(A,a)\in\mathcal{U}} \|Ax - a\|^2 \leq \lambda$. This form of quadratic constraint arises in a robust least squares models as well as a second order cone programming models.

We now show that checking robust feasibility is equivalent to solving a SDP, under suitable conditions.

Lemma 5.1 (SDP reformulation of robust feasibility) *Let $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}$ and $\mathcal{U}$ be given as in (5.1). Suppose that $k \geq s + 1$, where k is the number of rows in the matrix data of $\mathcal{U}$ and $s = \dim \operatorname{span}\{w^1, \ldots, w^l\}$, and that $\{\Delta : \|\Delta - \overline{\Delta}\|_F < \rho,\ (w^j)^T \operatorname{vec}\Delta < \beta^j, j = 1, \ldots, l\} \neq \emptyset$.. Then, (x, λ) is robust feasible for the quadratic constraint $\|Ax - a\|^2 \leq \lambda$ with respect to the uncertainty set $\mathcal{U}$ if and only if there exist $\lambda^0, \ldots, \lambda^l \geq 0$ such that*

$$\begin{pmatrix} I_k & I_k \otimes \tilde{x} & A^{(0)}x - a^{(0)} \\ (I_k \otimes \tilde{x})^T & \lambda^0 I_{k(n+1)} & -\lambda^0 b + \frac{1}{2}\sum_{j=1}^l \lambda^j w^j \\ (A^{(0)}x - a^{(0)})^T & \left(-\lambda^0 b + \frac{1}{2}\sum_{j=1}^l \lambda^j w^j\right)^T & \lambda - \lambda^0(\gamma - \|b\|^2) - \sum_{j=1}^l \lambda^j \beta^j \end{pmatrix} \geq 0,$$

where $\tilde{x} = (x^T, -1)^T \in \mathbb{R}^{n+1}$, $b = \operatorname{vec}(\overline{\Delta})$ and $\gamma = \rho^2 - \operatorname{Tr}(\overline{\Delta}^T\overline{\Delta})$.

Proof Let $\Delta = (\Delta A, \Delta a) \in \mathbb{R}^{k\times n} \times \mathbb{R}^k = \mathbb{R}^{k\times(n+1)}$. For $x \in \mathbb{R}^n$, denote $\tilde{x} = (x^T, -1)^T \in \mathbb{R}^{n+1}$. From the definition of $\mathcal{U}$, we note that $\max_{(A,a)\in\mathcal{U}} \|Ax - a\|^2 \leq \lambda$ if and only if

$$\|\Delta - \overline{\Delta}\|_F^2 \leq \rho^2,\ (w^j)^T \operatorname{vec}\Delta \leq \beta^j,\ j = 1, \ldots, l \ \Rightarrow\ \|A^{(0)}x - a^{(0)} + \Delta\tilde{x}\|^2 \leq \lambda,$$

which is equivalent to the following implication

$$\begin{aligned} &\operatorname{Tr}\big(\Delta^T\Delta - 2\overline{\Delta}^T\Delta + \overline{\Delta}^T\overline{\Delta}\big) \leq \rho^2,\ (w^j)^T\operatorname{vec}\Delta \leq \beta^j,\ j = 1, \ldots, l \\ &\quad\Rightarrow \operatorname{Tr}\big(\Delta\tilde{x}\,\tilde{x}^T\Delta^T + 2(A^{(0)}x - a^{(0)})\tilde{x}^T\Delta + (A^{(0)}x - a^{(0)})(A^{(0)}x - a^{(0)})^T\big) - \lambda \leq 0. \end{aligned}$$

Note that, for matrix $A, C \in \mathbb{R}^{p\times s}$ and $B \in \mathbb{R}^{p\times p}$,

$$\operatorname{Tr}(A^T B A) = \operatorname{vec}(A)^T(I_s \otimes B)\operatorname{vec}(A) \text{ and } \operatorname{Tr}(A^T C) = \operatorname{vec}(A)^T\operatorname{vec}(C). \quad (5.2)$$

Let $u = \operatorname{vec}(\Delta) \in \mathbb{R}^{k(n+1)}$. Then, using the identities in (5.2), we see that $\max_{(A,a)\in\mathcal{U}} \|Ax - a\|^2 \leq \lambda$ if and only if the following implication holds

$$\|u - b\|^2 \leq \gamma, (w^j)^T u \leq \beta^j,\ j = 1, \ldots, l \ \Rightarrow\ u^T Q u + a^T u + (r + \lambda) \geq 0$$

where $Q = -(I_k \otimes \tilde{x}\tilde{x}^T)$, $q = -\operatorname{vec}(2\tilde{x}(A^{(0)}x - a^{(0)})^T)$, $r = -\operatorname{Tr}((A^{(0)}x - a^{(0)})(A^{(0)}x - a^{(0)})^T)$, $b = \operatorname{vec}(\overline{\Delta})$ and $\gamma = \rho^2 - \operatorname{Tr}(\overline{\Delta}^T\overline{\Delta})$. As $Q = -(I_k \otimes \tilde{x}\tilde{x}^T)$, and so, $\dim\operatorname{Ker}(Q - \lambda_{\min}(Q)I_{k(n+1)}) \geq k \geq s + 1$.

$$\begin{aligned}&\dim\mathrm{Ker}(Q-\lambda_{\min}(Q)I_{k(n+1)})+\dim\left(\bigcap_{j=1}^{l}(w^j)^{\perp}\right)\\&\geq(s+1)+(k(n+1)-s)\geq k(n+1)+1,\end{aligned}$$

where $k(n+1)$ is the dimension of the given matrix data. Since $\mathcal{U}$ has a non-empty interior, by the extended version of S-lemma (Corollary 4.3), we see that $\max_{(A,a)\in\mathcal{U}}\|Ax-a\|^2\leq\lambda$ if and only if there exist $\lambda^0,\lambda^1,\ldots,\lambda^l\geq 0$ such that for all $u\in\mathbb{R}^{k(n+1)}$,

$$(u^TQu+q^Tu+r+\lambda)+\lambda^0(\|u-b\|^2-\gamma)+\sum_{j=1}^{l}\lambda^j((w^j)^Tu-\beta^j)\geq 0$$

which is equivalent to

$$\begin{pmatrix}Q+\lambda^0I_{k(n+1)} & \frac{1}{2}\left(q-2\lambda^0b+\sum_{j=1}^{l}\lambda^jw^j\right)\\ \frac{1}{2}(q-2\lambda^0b+\sum_{j=1}^{l}\lambda^jw^j)^T & r+\lambda-\lambda^0(\gamma-\|b\|^2)-\sum_{j=1}^{l}\lambda^j\beta^j\end{pmatrix}\succeq 0. \quad (5.3)$$

We now apply the method of Schur complement that, for $M_i\in S^n,\ i=1,2,3$ with $M_1\succ 0$, $\begin{pmatrix}M_1 & M_2\\ M_2^T & M_3\end{pmatrix}\succeq 0\Leftrightarrow M_3-M_2^TM_1^{-1}M_2\succeq 0$, to reformulate (5.3) into linear matrix inequalities. To see this, note that

$$\begin{aligned}Q&=-\left(I_k\otimes\tilde{x}\tilde{x}^T\right)=-\left(I_k\otimes\tilde{x}\right)\left(I_k\otimes\tilde{x}\right)^T\\q&=-\mathrm{vec}\left(2\tilde{x}\left(A^{(0)}x-a^{(0)}\right)^T\right)=-2\left(I_k\otimes\tilde{x}\right)\left(A^{(0)}x-a^{(0)}\right)\\r&=-\mathrm{Tr}\left(\left(A^{(0)}x-a^{(0)}\right)\left(A^{(0)}x-a^{(0)}\right)^T\right)=-\|A^{(0)}x-a^{(0)}\|^2,\end{aligned}$$

and let $M_1=I_k$, $M_2=(I_k\otimes\tilde{x},A^{(0)}x-a^{(0)})$ and

$$M_3=\begin{pmatrix}\lambda^0I_{k(n+1)} & -\lambda^0b+\frac{1}{2}\sum_{j=1}^{l}\lambda^jw^j\\ \left(-\lambda^0b+\frac{1}{2}\sum_{j=1}^{l}\lambda^jw^j\right)^T & \lambda-\lambda^0\left(\gamma-\|b\|^2\right)-\sum_{j=1}^{l}\lambda^j\beta^j\end{pmatrix}.$$

Then, $\max_{(A,a)\in\mathcal{U}}\|Ax-a\|^2\leq\lambda$ is equivalent to the following linear matrix inequality problem: there exist $\lambda^0,\ldots,\lambda^l\geq 0$ such that

$$\begin{pmatrix}I_k & I_k\otimes\tilde{x} & A^{(0)}x-a^{(0)}\\ (I_k\otimes\tilde{x})^T & \lambda^0I_{k(n+1)} & -\lambda^0b+\frac{1}{2}\sum_{j=1}^{l}\lambda^jw^j\\ (A^{(0)}x-a^{(0)})^T & (-\lambda^0b+\frac{1}{2}\sum_{j=1}^{l}\lambda^jw^j)^T & \lambda-\lambda^0(\gamma-\|b\|^2)-\sum_{j=1}^{l}\lambda^j\beta^j\end{pmatrix}\succeq 0.$$

□

Remark 5.1 (Key to SDP reformulation) The key to the SDP reformulation in Lemma 5.1 is that the robust feasibility of a given point can be equivalently rewritten as a quadratic optimization problem where the Hessian of the objective function is $-I_k \otimes \tilde{x}\tilde{x}^T$ (which has at least multiplicity k for each of its eigenvalues). So, the assumption that $k \geq s+1$ guarantees our dimension condition. This enables us to convert the robust problem into a SDP using our S-lemma. This technique has been exploited and used in robust optimization recently, see [3,4].

5.1 Robust least squares

Consider the least squares problem (LSP) under data uncertainty (see [13])

$$(LSP) \quad \min_{x \in \mathbb{R}^n} \|Ax - a\|^2$$

where the data $(A, a) \in \mathbb{R}^{k \times n} \times \mathbb{R}^k$ is uncertain and it belongs to the matrix uncertainty set $\mathcal{U}$. The robust counterpart of the uncertain least squares problem can be stated as follows:

$$(RLSP) \quad \min_{x \in \mathbb{R}^n} \max_{(A,a) \in \mathcal{U}} \|Ax - a\|^2,$$

which seeks a solution $x \in \mathbb{R}^n$ that minimizes the worst case data error with respect to all possible values of $(A, a) \in \mathcal{U}$.

The tractability of the robust problem (RSLP) strongly relies on the choice of the uncertainty set $\mathcal{U}$. For example, if the uncertainty set $\mathcal{U}$ is described by a single ellipsoid then (RSLP) can be reformulated as a semidefinite programming problem, and so, is tractable (see El Ghaoui and Lebretis [13]). Also, if $\mathcal{U}$ is given by an intersection of two ellipsoids, (RSLP) can be reformulated as a semidefinite programming problem under suitable regularity conditions (see [3]). However, if the uncertainty set $\mathcal{U}$ is given by an intersection of finitely many, but more than two, ellipsoids, then (RSLP) is generally not tractable (see [7]).

Here, we provide a new tractable case where the uncertainty is $\mathcal{U}$ is given by (5.1).

Theorem 5.1 (SDP characterization of (RSLP) solution) *Let $x \in \mathbb{R}^n$. For problem (RSLP) with $\mathcal{U}$ defined as in (5.1), assume that $k \geq s+1$, where k is the number of rows in the matrix data of $\mathcal{U}$ and $s = \dim \operatorname{span}\{w^1, \ldots, w^l\}$, and that $\{\Delta : \|\Delta - \overline{\Delta}\|_F < \rho,\ (w^j)^T \operatorname{vec}\Delta < \beta^j, j = 1, \ldots, l\} \neq \emptyset$.. Then x solves (RSLP) if and only if $(x, \lambda, \lambda^0, \ldots, \lambda^l) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+ \times \cdots \times \mathbb{R}_+$ solves the following linear semidefinite programming problem:*

$$\min_{(x,\lambda) \in \mathbb{R}^n \times \mathbb{R}, \lambda^0, \ldots, \lambda^l \geq 0} \lambda$$
$$\times \left\{ \lambda : \begin{pmatrix} I_k & I_k \otimes \tilde{x} & A^{(0)}x - a^{(0)} \\ (I_k \otimes \tilde{x})^T & \lambda^0 I_{k(n+1)} & -\lambda^0 b + \frac{1}{2}\sum_{j=1}^l \lambda^j w^j \\ (A^{(0)}x - a^{(0)})^T & (-\lambda^0 b + \frac{1}{2}\sum_{j=1}^l \lambda^j w^j)^T & \lambda - \lambda^0(\gamma - \|b\|^2) - \sum_{j=1}^l \lambda^j \beta^j \end{pmatrix} \succeq 0 \right\},$$

for some $\lambda \in \mathbb{R}$ *and* $\lambda^0, \ldots, \lambda^l \geq 0$, *where* $\tilde{x} = (x^T, -1)^T \in \mathbb{R}^{n+1}$, $b = \mathrm{vec}(\overline{\Delta})$ *and* $\gamma = \rho^2 - \mathrm{Tr}(\overline{\Delta}^T \overline{\Delta})$.

Proof Note that x is a solution of $\min_{x\in\mathbb{R}^n} \max_{(A,a)\in\mathcal{U}} \|Ax - a\|^2$ if and only if there exists $\lambda \in \mathbb{R}$ such that (x, λ) solves $\min_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R}}\{\lambda : \max_{(A,a)\in\mathcal{U}} \|Ax - a\|^2 \leq \lambda\}$. Then, by Lemma 5.1, we see that $x \in \mathbb{R}^n$ solves (RLSP) if and only if $(x, \lambda, \lambda^0, \ldots, \lambda^l) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+ \times \cdots \times \mathbb{R}_+$ solves the following linear semi-definite programming problem:

$$\min_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R},\lambda^0,\ldots,\lambda^l\geq 0} \times \left\{\lambda : \begin{pmatrix} I_k & I_k \otimes \tilde{x} & A^{(0)}x - a^{(0)} \\ (I_k \otimes \tilde{x})^T & \lambda^0 I_{k(n+1)} & -\lambda^0 b + \frac{1}{2}\sum_{j=1}^l \lambda^j w^j \\ (A^{(0)}x - a^{(0)})^T & (-\lambda^0 b + \frac{1}{2}\sum_{j=1}^l \lambda^j w^j)^T & \lambda - \lambda^0(\gamma - \|b\|^2) - \sum_{j=1}^l \lambda^j \beta^j \end{pmatrix} \succeq 0\right\}$$

for some $\lambda \in \mathbb{R}$ and $\lambda^0, \ldots, \lambda^l \geq 0$. □

Consider the special case of the uncertainty set, $\mathcal{U}$, in (5.1) where $l = 1$, $\overline{\Delta} = 0$, $w^1 = 0$ and $\beta^1 = 1$. In this case, the $\mathcal{U}$ reduces to the matrix norm uncertainty set of the form

$$\mathcal{U} = \{\tilde{A}^{(0)} + \Delta : \Delta \in \mathbb{R}^{k\times(n+1)}, \|\Delta\|_F \leq \rho\}, \tag{5.4}$$

and the tractability of robust least squares problem (RLSP) was established in El Ghoui and Lebret [13]. In the following Corollary we derive an SDP characterization of (RLSP) for the uncertainty set (5.4).

Corollary 5.1 (Matrix norm uncertainty) *Let* $x \in \mathbb{R}^n$. *For problem (RSLP) with* $\mathcal{U}$ *defined as in* (5.4), *assume that* $\rho > 0$. *Then* x *solves (RLSP) if and only if* $(x, \lambda, \lambda^0, \lambda^1) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$ *solves the following linear semi-definite programming problem:*

$$\min_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R},\lambda^0,\lambda^1\geq 0} \times \left\{\lambda : \begin{pmatrix} I_k & I_k \otimes \tilde{x} & A^{(0)}x - a^{(0)} \\ (I_k \otimes \tilde{x})^T & \lambda^0 I_{k(n+1)} & 0 \\ (A^{(0)}x - a^{(0)})^T & 0 & \lambda - \lambda^0\rho^2 - \lambda^1 \end{pmatrix} \succeq 0\right\}.$$

for some $\lambda \in \mathbb{R}$ *and* $\lambda^0, \lambda^1 \geq 0$, *where* $\tilde{x} = (x^T, -1)^T \in \mathbb{R}^{n+1}$.

Proof Let $l = 1$, $\overline{\Delta} = 0$, $w^1 = 0$ and $\beta^1 = 1$. Then, $s = \dim\mathrm{span}\{w^1\} = 0$, and so, $k \geq 1 = s + 1$. Moreover, as $\rho > 0$, the strict feasibility condition is satisfied for $\Delta = 0$. Thus, the conclusion follows by the preceding theorem. □

Consider the special case of the uncertainty set, $\mathcal{U}$, in (5.1), where $l = 2$, $\overline{\Delta} = 0$, $w^2 = -w^1$ and $\beta^1 = -\beta^2 = 1$. In this case, $\mathcal{U}$ simplifies to case of an intersection of two ellipsoids of the form

$$\begin{aligned}\mathcal{U} &= \{\tilde{A}^{(0)} + \Delta : \Delta \in \mathbb{R}^{k\times(n+1)}, \|\Delta\|_F \leq \rho,\ -1 \leq (w^1)^T \mathrm{vec}\Delta \leq 1\} \\ &= \{\tilde{A}^{(0)} + \Delta : \Delta \in \mathbb{R}^{k\times(n+1)}, \mathrm{Tr}(\Delta^T\Delta) \leq \rho^2,\ \mathrm{Tr}(\Delta^T B\Delta) \leq 1\},\end{aligned} \tag{5.5}$$

where $B = (w^1)(w^1)^T$. In this case, an SDP characterization of robust solution was established in Beck and Eldar [3]. In this case we obtain the following corollary.

Corollary 5.2 (Intersection of two ellipsoids uncertainty) *Let $x \in \mathbb{R}^n$. For problem (RSLP) with $\mathcal{U}$ defined as in (5.5), assume that $k \geq 2$, where k is the number of rows in the matrix data of $\mathcal{U}$, and that $\rho > 0$. Then x solves (RLSP) if and only if $(x, \lambda, \lambda^0, \lambda^1, \lambda^2) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$ solves the following linear semi-definite programming problem:*

$$\min_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R},\lambda^0,\ldots,\lambda^l\geq 0} \times \left\{ \lambda : \begin{pmatrix} I_k & I_k \otimes \tilde{x} & A^{(0)}x - a^{(0)} \\ (I_k \otimes \tilde{x})^T & \lambda^0 I_{k(n+1)} & \frac{1}{2}(\lambda^1 w^1 - \lambda^2 w^1) \\ (A^{(0)}x - a^{(0)})^T & \frac{1}{2}(\lambda^1 w^1 - \lambda^2 w^1)^T & \lambda - \lambda^0\rho^2 - (\lambda^1 - \lambda^2) \end{pmatrix} \succeq 0 \right\},$$

for some $\lambda \in \mathbb{R}$ and $\lambda^0, \lambda^l, \lambda^2 \geq 0$, where $\tilde{x} = (x^T, -1)^T \in \mathbb{R}^{n+1}$.

Proof Let $l = 2$, $\overline{\Delta} = 0$, $w^2 = -w^1$ and $\beta^1 = -\beta^2 = 1$. Then, $s = \dim\text{span}\{w^1, w^2\} \leq 1$, and so, $k \geq 2 \geq s+1$. Moreover, as $\rho > 0$, the strict feasibility condition is satisfied for $\Delta = 0$. Thus, the conclusion follows from Theorem 5.1. □

Remark 5.2 (Tractability of (RLSP)) It follows easily from Theorem 5.1 that finding a solution of the robust least squares with the uncertainty set given by an intersection of a norm constraint and a polyhedral constraint is equivalent to solving a linear semi-definite programming problem. Note that a linear semi-definite programming problem can be solved in polynomial time and $s = \dim\text{span}\{w^1, \ldots, w^l\} \leq l$ (and so, $k \geq l+1$ implies that $k \geq s+1$). So, a solution of this robust least squares can be validated in polynomial time whenever $k \geq l+1$ where k is the number of rows in the matrix data A and l is the number of the linear inequalities that defines the uncertainty set.

5.2 Robust second order cone programming problems

Consider the linear second-order cone programs (SOCP) (cf. [1]) under constraint data uncertainty

$$(SOCP) \quad \min_{x\in\mathbb{R}^n} \ a^T x \quad \text{s.t.} \ \ \|B_i x - b_i\| \leq d_i, \ i = 1, \ldots, m,$$

where the data $\tilde{B}_i = (B_i, b_i) \in \mathbb{R}^{k_i \times n} \times \mathbb{R}^{k_i} = \mathbb{R}^{k_i \times (n+1)}$, $i = 1, \ldots, m$, is uncertain and it belongs to the matrix uncertainty set $\mathcal{U}_i$. The robust counterpart of the uncertain second-order cone problem can be stated as follows:

$$(RSOCP) \quad \min_{x\in\mathbb{R}^n} \ a^T x \quad \text{s.t.} \ \ \|B_i x - b_i\| \leq d_i, \ \forall (B_i, b_i) \in \mathcal{U}_i, \ i = 1, \ldots, m.$$

Note that, although the (RSOCP) is, in general, not tractable [7] when $\mathcal{U}_i$ is given by an intersection of finitely ellipsoids, recently, Beck [4] has identified an interesting

tractable subclass where $\mathcal{U}_i$ is described by at most k many homogeneous quadratic inequalities under a suitable regularity condition.

Here, we examine (RSOCP) in the case where the uncertainty set is given by an intersection of a matrix norm constraint and polyhedral constraints, i.e.,

$$\mathcal{U}_i = \Big\{ \tilde{B}_i^{(0)} + \Delta_i : \Delta_i \in \mathbb{R}^{k_i \times (n+1)}, \|\Delta_i - \overline{\Delta}_i\|_F \le \rho_i, (w_i^j)^T \mathrm{vec}\Delta_i \le \beta_i^j, j = 1, \ldots, l_i \Big\}, \tag{5.6}$$

with $\tilde{B}_i^{(0)} := (B_i^{(0)}, b_i^{(0)}) \in \mathbb{R}^{k_i \times n} \times \mathbb{R}^k = \mathbb{R}^{k_i \times (n+1)}$ and $\|M\|_F$ is the Frobenius norm defined by $\|M\|_F = \sqrt{\mathrm{Tr}(M^T M)}$. We denote $s_i = \dim \mathrm{span}\{w_i^1, \ldots, w_i^l\}$, $i = 1, \ldots, m$.

We note that our model differs from the model considered in [4] because a polyhedral set in $\mathbb{R}^{(n+1)\times(n+1)}$ cannot, in general, be described as the finite intersection of sets of the form $\{\Delta \in \mathbb{R}^{(n+1)\times(n+1)} : \|C_j \Delta^T\|_F^2 \le \rho_j^2\}$ in general.

Theorem 5.2 (SDP characterization of (RSOCP) solution) *For problem (RSOCP) with $\mathcal{U}_i$ defined as in (5.6). Assume that, for each $i = 1, \ldots, m$, $k_i \ge s_i + 1$, and that $\{\Delta_i : \|\Delta_i - \overline{\Delta}_i\|_F < \rho_i,\ (w_i^j)^T \mathrm{vec}\Delta_i < \beta_i^j, j = 1, \ldots, l_i\} \ne \emptyset$. A point $x \in \mathbb{R}^n$ solves (RSOCP) if and only if $(x, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^n \times \mathbb{R}_+^{l_1+1} \times \ldots \times \mathbb{R}_+^{l_m+1}$ solves the following linear semi-definite programming problem:*

$$\min_{\lambda_i^0, \ldots, \lambda_i^{l_i} \ge 0, x \in \mathbb{R}^n} \times \left\{ a^T x : \begin{pmatrix} I_{k_i} & I_{k_i} \otimes \tilde{x} & B_i^{(0)} x - b_i^{(0)} \\ (I_{k_i} \otimes \tilde{x})^T & \lambda_i^0 I_{k_i(n+1)} & -\lambda_i^0 \overline{b}_i + \frac{1}{2} \sum_{j=1}^{l_i} \lambda_i^j w_i^j \\ (B_i^{(0)} x - b_i^{(0)})^T & (-\lambda_i^0 \overline{b}_i + \frac{1}{2} \sum_{j=1}^{l_i} \lambda_i^j w_i^j)^T & d_i^2 - \lambda_i^0 (\overline{\gamma}_i - \|\overline{b}_i\|^2) - \sum_{j=1}^{l_i} \lambda_i^j \beta_i^j \end{pmatrix} \succeq 0 \right\},$$

for some $\lambda_i = (\lambda_i^0, \lambda_i^1, \ldots, \lambda_i^{l_i}) \in \mathbb{R}_+^{l_i+1}$, $i = 1, \ldots, m$, where $\tilde{x} = (x^T, -1)^T \in \mathbb{R}^{n+1}$ and $\overline{b}_i = \mathrm{vec}(\overline{\Delta}_i)$ and $\overline{\gamma}_i = \rho_i^2$.

Proof Note that a point x is robust feasible if, for all $i = 1, \ldots, m$,

$$\max_{(B_i, b_i) \in \mathcal{U}_i} \|B_i x - b_i\|^2 \le d_i^2.$$

So, Lemma 5.1 implies that the robust feasibility of x can be equivalently rewritten as the following linear matrix inequality problem: for all $i = 1, \ldots, m$, there exist $\lambda_i^0, \ldots, \lambda_i^{l_i} \ge 0$ such that

$$\begin{pmatrix} I_{k_i} & I_{k_i} \otimes \tilde{x} & B_i^{(0)} x - b_i^{(0)} \\ (I_{k_i} \otimes \tilde{x})^T & \lambda_i^0 I_{k_i(n+1)} & -\lambda_i^0 \overline{b}_i + \frac{1}{2} \sum_{j=1}^{l_i} \lambda_i^j w_i^j \\ (B_i^{(0)} x - b_i^{(0)})^T & (-\lambda_i^0 \overline{b}_i + \frac{1}{2} \sum_{j=1}^{l_i} \lambda_i^j w_i^j)^T & d_i^2 - \lambda_i^0 (\overline{\gamma}_i - \|\overline{b}_i\|^2) - \sum_{j=1}^{l_i} \lambda_i^j \beta_i^j \end{pmatrix} \succeq 0.$$

Thus, the conclusion follows. □

Consider the special case of the uncertainty set (5.6), where $l_i = 1,\ \overline{\Delta} = 0,\ w_i^1 = 0$ and $\beta_i^1 = 1,\ i = 1, \ldots, m$. In this case $\mathcal{U}_i$ reduces to the matrix norm uncertainty set of the form

$$\mathcal{U}_i = \{\tilde{A}_i^{(0)} + \Delta_i : \Delta_i \in \mathbb{R}^{k_i \times (n+1)}, \|\Delta_i\|_F \leq \rho\}. \tag{5.7}$$

An SDP characterization of robust solution of second-order cone programming problem was established in [7].

Corollary 5.3 (Matrix norm uncertainty) *Let $x \in \mathbb{R}^n$. For problem (RSOCP) with $\mathcal{U}$ defined as in (5.7), assume that $\rho > 0$. Then x solves (RSOCP) if and only if $(x, \lambda, \lambda_i^0, \lambda_i^1) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$ solves the following linear semi-definite programming problem:*

$$\min_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R},\lambda_i^0,\lambda_i^1\geq 0} \times \left\{\lambda : \begin{pmatrix} I_k & I_k \otimes \tilde{x} & A^{(0)}x - a^{(0)} \\ (I_k \otimes \tilde{x})^T & \lambda^0 I_{k_i(n+1)} & 0 \\ (A^{(0)}x - a^{(0)})^T & 0 & \lambda - \lambda_i^0\rho^2 - \lambda_i^1 \end{pmatrix} \succeq 0\right\},$$

for some $\lambda \in \mathbb{R}$ and $\lambda_i^0, \lambda_i^l \geq 0$, where $\tilde{x} = (x^T, -1)^T \in \mathbb{R}^{n+1}$.

Proof Let $l_i = 1,\ \overline{\Delta} = 0,\ w_i^1 = 0$ and $\beta_i^1 = 1,\ i = 1, \ldots, m$. Then, $s_i = \dim\text{span}\{w_i^1\} = 0$, and so, $k_i \geq 1 \geq s_i + 1$. Moreover, as $\rho > 0$, the strict feasibility condition is satisfied for $\Delta = 0$. Thus, the conclusion follows from Theorem 5.2. □

Consider another special case of the uncertainty set (5.6), where $k_i = k - 1,\ \overline{\Delta}_i = 0,\ \beta_i^l = -\beta_i^{l+k-1} = 1$ and $w_i^l = -w_i^{l+k-1} = w^l,\ l = 1, \ldots, 2k_i,\ i = 1, \ldots, m$. In this case, the uncertainty set $\mathcal{U}_i$ simplifies to the intersection of k many ellipsoids of the form

$$\begin{aligned}\mathcal{U}_i &= \{\tilde{A}_i^{(0)} + \Delta_i : \Delta_i \in \mathbb{R}^{k\times(n+1)}, \|\Delta_i\|_F \leq \rho_i, -1 \leq (w^l)^T \operatorname{vec}\Delta_i \leq 1, l = 1, \ldots, k-1\} \\ &= \{\tilde{A}_i^{(0)} + \Delta_i : \Delta_i \in \mathbb{R}^{k\times(n+1)}, \operatorname{Tr}(\Delta_i^T\Delta_i) \leq \rho_i^2, \operatorname{Tr}(\Delta_i^T C^l \Delta_i) \leq 1, l = 1, \ldots, k-1\},\end{aligned} \tag{5.8}$$

where $C^l = (w^l)(w^l)^T,\ l = 1, \ldots, k-1$. The following robust solution characterization in terms of SDP has been given in Beck [4].

Corollary 5.4 [4, Section 4.3] (Intersection of many ellipsoids uncertainty) *Let $x \in \mathbb{R}^n$. For problem (RSOCP) with $\mathcal{U}_i$ defined as in (5.8), assume that $\rho_i > 0$. A point $x \in \mathbb{R}^n$ solves (RSOCP) if and only if $(x, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^n \times \mathbb{R}_+^{2k-1} \times \cdots \times \mathbb{R}_+^{2k-1}$ solves the following linear semi-definite programming problem:*

$$\min_{\lambda_i^0,\ldots,\lambda_i^{2(k-1)}\geq 0, x\in\mathbb{R}^n} \times \left\{ a^T x : \begin{pmatrix} I_{k_i} & I_{k_i}\otimes\tilde{x} & B_i^{(0)}x - b_i^{(0)} \\ (I_{k_i}\otimes\tilde{x})^T & \lambda_i^0 I_{k_i(n+1)} & \frac{1}{2}\left(\sum_{j=1}^{k-1}\lambda_i^j w^j - \sum_{j=1}^{k-1}\lambda_i^{j+k-1}w^j\right) \\ (B_i^{(0)}x - b_i^{(0)})^T & \frac{1}{2}\left(\sum_{j=1}^{k-1}\lambda_i^j w^j - \sum_{j=1}^{k-1}\lambda_i^{j+k-1}w^j\right)^T & d_i^2 - \lambda_i^0\rho_i^2 - \sum_{j=1}^{k-1}\lambda_i^j + \sum_{j=1}^{k-1}\lambda_i^{j+k-1} \end{pmatrix} \succeq 0 \right\},$$

for some $\lambda_i = (\lambda_i^0, \lambda_i^1, \ldots, \lambda_i^{2(k-1)}) \in \mathbb{R}_+^{2k-1}$, $i = 1, \ldots, m$, *where* $\tilde{x} = (x^T, -1)^T \in \mathbb{R}^{n+1}$.

Proof Let $k_i = k-1$, $\overline{\Delta}_i = 0$, $w_i^l = -w_i^{l+k-1} = w^l$ and $\beta_i^l = -\beta_i^{l+k-1} = 1$, $l = 1, \ldots, 2k_i$, $i = 1, \ldots, m$. Then, $s_i = \dim \operatorname{span}\{w^1, \ldots, w^{2(k-1)}\} \leq k-1$, and so, $k_i = k \geq s_i + 1$. Moreover, as $\rho_i > 0$, the strict feasibility condition is satisfied for $\Delta = 0$. Thus, the conclusion follows by the preceding theorem. □

6 Extensions and further research

In this section, we present how our approach extends to more general trust-region problems that incorporate uniform convex quadratic inequalities. To examine this, consider the system of quadratic functions, $f(x) = x^T Ax + a^T x + \gamma$, $g_0(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = \|Bx\|^2 + b_i^T x - \beta_i$, $i = 1, \ldots, m$, where $A \in S^{n\times n}$, $B \in \mathbb{R}^{l\times n}$ with $l \in \mathbb{N}$, $a, x_0, b_i \in \mathbb{R}^n$ and $\gamma, \alpha, \beta_i \in \mathbb{R}$. In this case, we can consider the following *extended dimension condition*

$$\dim\big(\operatorname{Ker}(A - \lambda_{\min}(A)I_n) \cap \operatorname{Ker}(B)\big) \geq s + 1, \tag{6.1}$$

where s is the dimension of $\operatorname{span}\{b_1, \ldots, b_m\}$.

Clearly, if the matrix B is zero, then the above quadratic systems and the extended dimension condition reduce to the quadratic systems and its associated dimension condition, studied in Sects. 2–4. On the other hand, in the case when B has rank n, our dimension condition (6.1) fails.

As we see in the following Proposition, the hidden convexity property of Sect. 2 follows for the above general quadratic system under the extended dimension condition.

Proposition 6.1 (Hidden convexity of General Quadratic Systems) *Let* $f(x) = x^T Ax + a^T x + \gamma$, $g_0(x) = \|x - x_0\|^2 - \alpha$ *and* $g_i(x) = \|Bx\|^2 + b_i^T x - \beta_i$, $i = 1, \ldots, m$, $A \in S^{n\times n}$, $B \in \mathbb{R}^{l\times n}$ *with* $l \in \mathbb{N}$, $a, x_0, b_i \in \mathbb{R}^n$ *and* $\gamma, \alpha, \beta_i \in \mathbb{R}$. *Suppose that the extended dimension condition* (6.1) *is satisfied.*

Then,

$$U(f, g_0, g_1, \ldots, g_m) := \{(f(x), g_0(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\} + \mathbb{R}_+^{m+2}$$

is a convex set.

Proof As in the proof of Theorem 2.1, we can assume without loss of generality that A is not positive semidfinite. Define h by $h(x) = \min_{x\in\mathbb{R}^n}\{f(x) : \|x - x_0\|^2 \le \alpha + r,\ \|Bx\|^2 + b_i^T x \le \beta + s_i,\ i = 1,\ldots,m\}$ if $x \in D := \{(r, s_1, \ldots, s_m) : \|x - x_0\|^2 \le \alpha + r,\ \|Bx\|^2 + b_i^T x \le \beta_i + s_i \text{ for some } x \in \mathbb{R}^n\}$ and $h(x) = +\infty$ if $x \notin D$. Using the same line of arguments as in Theorem 2.1, we can easily verify that $U(f, g_0, g_1, \ldots, g_m) = \text{epi}h$. Moreover, h is convex if the minimization problem

$$\min_{x\in\mathbb{R}^n}\left\{f(x) - \lambda_{\min}(A)\|x - x_0\|^2 : \|x - x_0\|^2 \le \alpha + r,\ \|Bx\|^2 + b_i^T x \le \beta_i + s_i\right\}$$

attains its minimum at some $\overline{x} \in \mathbb{R}^n$ with $\|\overline{x}-x_0\|^2 = \alpha+r$ and $\|B\overline{x}\|^2+b_i^T\overline{x} \le \beta_i+s_i$.

Indeed, this optimization problem has a minimizer on the sphere. This follows from the fact that there exists $v \in \mathbb{R}^n\backslash\{0\}$ such that

$$v \in \left(\bigcap_{i=1}^m b_i^\perp\right) \cap \text{Ker}(A - \lambda_{\min}(A)I_n) \cap \text{Ker}(B). \tag{6.2}$$

Otherwise, $\left(\bigcap_{i=1}^m b_i^\perp\right) \cap \text{Ker}(A - \lambda_{\min}(A)I_n) \cap \text{Ker}(B) = \{0\}$. Then it follows from our extended dimension condition, $\dim\big(\text{Ker}(A-\lambda_{\min}(A)I_n)\cap\text{Ker}(B)\big) \ge s+1$, where s is the dimension of $\text{span}\{b_1, \ldots, b_m\}$, that

$$\begin{aligned} n + 1 = (s + 1) + (n - s) &\le \dim(\text{Ker}(A - \lambda_{\min}(A)I_n \cap \text{Ker}(B)) + \dim\left(\bigcap_{i=1}^m b_i^\perp\right) \\ &= \dim\left(\text{Ker}(A - \lambda_{\min}(A)I_n) \cap \text{Ker}(B) + \bigcap_{i=1}^m b_i^\perp\right) \\ &\quad + \dim\left(\bigcap_{i=1}^m b_i^\perp \cap \text{Ker}(A - \lambda_{\min}(A)I_n \cap \text{Ker}(B))\right) \\ &\le n, \end{aligned}$$

which is impossible.

So, the same line of arguments as in Theorem 2.1 gives the desired conclusion. □

Recently, in [3], the authors considered trust region problem with one additional linear inequality constraint:

$$(P_2) \qquad \min\{x^T Ax + a^T x : \|x - x_0\|^2 \le \alpha,\ b_1^T x \le \beta_1\}$$

and showed that strong duality holds for (P_1) whenever $\dim\big(\text{Ker}(A-\lambda_{\min}(A)I_n) \ge 2$. Extending this, we consider the following quadratic optimizations with one additional convex quadratic constraint

$$(GP_2) \qquad \min\{x^T Ax + a^T x : \|x - x_0\|^2 \le \alpha,\ \|Bx\|^2 + b_1^T x \le \beta_1\}.$$

Following similar methods of proof of Sect. 3 and 4 and using the preceding proposition, we derive SDP relaxation and strong duality results for (GP_1) under the following dimension condition "$\dim\big(\mathrm{Ker}(A-\lambda_{\min}(A)I_n)\cap\mathrm{Ker}(B)\big)\geq 2$". However, it should be noted that, this dimension condition fails to be satisfied when B has rank n (the dimension of the underlying space). Indeed, in the case when B has rank n, an example was provided in [30, Page 263 EX_1] showing that the model (GP_2) does not enjoy exact SDP relaxation as well as strong duality in general.

Theorem 6.1 *For problem (GP_2), suppose that $\dim\big(\mathrm{Ker}(A-\lambda_{\min}(A)I_n)\cap\mathrm{Ker}(B)\big)\geq 2$. Then, (GP_2) admits exact SDP relaxation. Moreover, suppose further that there exists $\overline{x}$ such that $\|\overline{x}-x_0\|^2<\alpha$ and $\|B\overline{x}\|^2+b_1^T\overline{x}<\beta_1$. Then, strong duality holds for problem (GP_2), i.e.,*

$$\min_{x\in\mathbb{R}^n}\left\{x^TAx+a^Tx:\|x-x_0\|^2\leq\alpha,\ \|Bx\|^2+b_1^Tx\leq\beta_1\right\}$$
$$=\max_{\lambda_0,\lambda_1\geq 0}\min_{x\in\mathbb{R}^n}\left\{x^TAx+a^Tx+\lambda_0(\|x-x_0\|^2-\alpha)+\lambda_1(\|Bx\|^2+b_1^Tx\leq\beta_1)\right\}.$$

Proof From Proposition 6.1 and the assumption $\dim\big(\mathrm{Ker}(A-\lambda_{\min}(A)I_n)\cap\mathrm{Ker}(B)\big)\geq 2$, we see that $U(f,g_0,g_1)$ is convex where $f(x)=x^TAx+a^Tx$, $g_0(x)=\|x-x_0\|^2-\alpha$ and $g_1(x)=\|Bx\|^2+b_1^Tx-\beta_1$. So, the first conclusion can be proved following similar line of argument as in Theorem 3.1 while the second conclusion can be proved following similar line argument as in Theorem 4.1 and Corollary 4.1. □

Remark 6.1 A careful examination of the proof of above theorem shows that the conclusion of Theorem 6.1 continues to hold for the quadratic problem $\min\{x^TAx+a^Tx:\|x-x_0\|^2\leq\alpha,\ \|Bx\|^2+b_i^Tx\leq\beta_i, i=1,\ldots,l\}$ under the condition "$\dim\big(\mathrm{Ker}(A-\lambda_{\min}(A)I_n)\cap\mathrm{Ker}(B)\big)\geq l+1$". For simplicity, we only considered (GP_2) with two constraints.

In the special case of (GP_2), where B is the zero matrix, the preceding theorem reduces to [3, Theorem 3.6] (see Corollary 4.2).

The following example illustrates that Theorem 6.1 can be applied to some cases where B is not a zero matrix.

Example 6.1 Consider the following quadratic minimization problem

$$\begin{aligned}(P)\ \min\quad & -x_1^2-x_2^2-x_3^2-2x_1\\ \text{s.t.}\quad & x_1^2+x_2^2+x_3^2+x_1\leq 1,\\ & x_1^2+x_1\leq 0.\end{aligned}$$

This quadratic problem can be written as (GP_2) with $f(x)=x^TAx+a^Tx$ with $A=-I_3$ and $a=(-2,0,0)$, $g_0(x)=\|x-x_0\|^2-\alpha$ with $x_0=(-\frac{1}{2},0,0)$, $\alpha=\frac{5}{4}$, and $g_1(x)=\|Bx\|^2+b_1^Tx-\beta_1$ with $B=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}$, $b_1=(1,0,0)$ and $\beta_1=0$.

One could verify that the strict feasibility condition is satisfied at $\overline{x} = (-\frac{1}{2}, 0, 0)^T$ and

$$\dim\big(\mathrm{Ker}(A - \lambda_{\min}(A) I_n) \cap \mathrm{Ker}(B)\big) = 2.$$

Next, we show that strong duality and exact SDP relaxation hold. To see this, we note that, for any feasible point $x = (x_1, x_2, x_3)$, we have $-x_2^2 - x_3^2 \geq x_1^2 + x_1 - 1$ and $-1 \leq x_1 \leq 0$, and hence,

$$-x_1^2 - x_2^2 - x_3^2 - 2x_1 \geq -x_1 - 1 \geq -1.$$

So, it can be easily seen that the optimal value of (P) is -1 and $(0, 1, 0)$ is a global minimizer. Let $\lambda_0 = 1$ and $\lambda_1 = 1$. Then,

$$\min\{f(x) + \lambda_0 g_0(x) + \lambda_1 g_1(x)\} = \min\{x_1^2 - 1\} = -1 = \min(P).$$

So, the inequalities $\max(D) \leq \min(SDRP) \leq \min(P)$ imply that the strong duality and exact SDP relaxation hold.

Finally, we note that our approach and results in the present work suggest that the exact SDP-relaxation and strong duality may extend to multi-variate polynomial problems with a norm constraint and linear inequalities under an appropriate dimension condition. Moreover, it would be interesting to examine further potential applications of strong duality to robust optimization problems. These will be our future research direction and will be examined in a forthcoming study.

7 Appendix: Technical results

For the sake of self-containment, in this Section, we provide known technical results on hidden convexity of quadratic systems, S-lemma and tractable classes of robust optimization.

7.1 Hidden convexity of quadratic systems

The basic and probably the most useful result on the joint-range convexity of homogeneous quadratic functions, known as Dine's Theorem [11], states as follows:

Lemma 7.1 (Dine's Theorem) [11] *Let $A_1, A_2 \in S^n$. Then, the set $\{(x^T A_1 x, x^T A_2 x) : x \in \mathbb{R}^n\}$ is convex.*

Dine's theorem is known to fail for three homogeneous in general. Polyak [25] established the following joint-range convexity result for three homogeneous quadratic functions under a positive definite condition on the matrices involved.

Lemma 7.2 (Polyak's Lemma [25, Theorem 2.1]) *Let $n \geq 2$ and let $A_1, A_2, A_3 \in S^n$. Suppose that there exist $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ such that $\gamma_1 A_1 + \gamma_2 A_2 + \gamma_3 A_3 \succ 0$. Then the set $\{(x^T A_1 x, x^T A_2 x, x^T A_3 x) : x \in \mathbb{R}^n\}$ is convex.*

7.2 S-lemma and approximate S-lemma

Using Dine's Theorem, Yakubovich (cf [24]) obtained the following fundamental *S*-lemma which has played a key role in many areas of control and optimization.

Lemma 7.3 (*S*-lemma [24]) *Let* $A_1, A_2 \in S^n$, $a_1, a_2 \in \mathbb{R}^n$ *and* $\alpha_1, \alpha_2 \in \mathbb{R}$. *Suppose that there exists* $x_0 \in \mathbb{R}^n$ *such that* $x_0^T A_2 x_0 + a_2^T x_0 + \alpha_2 < 0$. *Then the following statements are equivalent:*

(i) $x^T A_2 x + a_2^T x + \alpha_2 \leq 0 \Rightarrow x^T A_1 x + a_1^T x + \alpha_1 \geq 0$;
(ii) $(\exists \lambda \geq 0)(\forall x \in \mathbb{R}^n)(x^T A_1 x + a_1^T x + \alpha_1) + \lambda(x^T A_2 x + a_2^T x + \alpha_2) \geq 0$.

For a homogeneous quadratic system with multiple convex quadratic constraints, Ben-Tal et al. [7] derived the following approximate S-lemma which provides an estimate between an associated quadratic optimization problem and its SDP relaxation.

Lemma 7.4 (Approximate S-lemma [7, Lemma A.6]) *Let* $R, H_0, H_1, \ldots, H_K$ *be symmetric* $(p \times p)$ *matrices such that* $H_i \succeq 0$, $i = 1, \ldots, K$ *and* $\sum_{k=0}^K \lambda_i H_i \succ 0$, *for some* $\lambda_i \geq 0, i = 0, \ldots, K$. *Consider the following quadratically constrained quadratic problem*

$$(QCQ) \quad \max_{y \in \mathbb{R}^p} \{y^T R y : y^T H_0 y \leq 1, y^T H_i y \leq 1, \quad i = 1, \ldots, K\}$$

and the semidefinite optimization problem

$$(SDP) \quad \min_{\mu_0, \ldots, \mu_K \geq 0} \left\{ \sum_{i=0}^K \mu_i : \sum_{i=0}^K \mu_k H_k \succeq R \right\}.$$

Then, $\max(QCQ) \leq \min(SDP) \leq \rho^2 \max(QCQ)$ *where* $\rho = \sqrt{2 \log(6 \sum_{i=1}^K \operatorname{rank} H_k)}$.

7.3 Tractable classes of robust optimization problems

The following tractable classes of robust optimization problems are known.

7.3.1 Robust least squares problems [3,13]

Consider the following robust least squares programming problem:

$$(RLSP) \quad \min_{x \in \mathbb{R}^n} \max_{(A,a) \in \mathcal{U}} \|Ax - a\|^2,$$

where $\mathcal{U} \subseteq \mathbb{R}^{k \times n} \times \mathbb{R}^k = \mathbb{R}^{k \times (n+1)}$, is an uncertainty set. Then, (RLSP) can be equivalently rewritten as a semidefinite programming problem under the following two cases:

(i) $\mathcal{U}$ is an ellipsoid (see [13]), i.e., $\mathcal{U} = \{(A^{(0)}, a^{(0)}) + \Delta : \Delta \in \mathbb{R}^{k \times (n+1)}, \|\Delta - \overline{\Delta}\|_F \leq \rho\}$;

(ii) $k \geq 2$ and $\mathcal{U}$ is the intersection of two ellipsoids (see [3]), i.e, $\mathcal{U} = \{(A^{(0)}, a^{(0)}) + \Delta : \Delta \in \mathbb{R}^{k\times(n+1)}, \ \mathrm{Tr}(\Delta B_j \Delta) \leq \rho_j^2, \ j = 1, 2\}$ where $B_j \in S^{n\times n}$ satisfying $\gamma_1 B_1 + \gamma_2 B_2 \succ 0$ for some $\gamma_1, \gamma_2 \geq 0$.

7.3.2 Robust second-order cone programming problems [4,7]

Consider the following robust second order cone programming problem:

$$(RSOCP) \ \min_{x\in\mathbb{R}^n} \ a^T x \\ \text{s.t.} \ \ \|B_i x - b_i\| \leq d_i, \ \forall (B_i, b_i) \in \mathcal{U}_i, \quad i = 1, \ldots, m,$$

where $\mathcal{U}_i \subseteq \mathbb{R}^{k_i\times n} \times \mathbb{R}^{k_i} = \mathbb{R}^{k_i\times(n+1)}$, $i = 1, \ldots, m$, is an uncertainty set. Then, (RSOCP) can be equivalently rewritten as a semidefinite programming problem under the following two cases:

(i) $\mathcal{U}_i$ is an ellipsoid (see [7]), i.e., $\mathcal{U}_i = \{(B_i^{(0)}, b_i^{(0)}) + \Delta_i : \Delta_i \in \mathbb{R}^{k_i\times(n+1)}, \|\Delta_i - \overline{\Delta}_i\|_F \leq \rho_i\}$;

(ii) $\mathcal{U}_i$ is the intersection of at most k many ellipsoids (see [4]), i.e, $k_i = k$ with $k \in \mathbb{N}$ and $\mathcal{U}_i = \{(B_i^{(0)}, b_i^{(0)}) + \Delta : \Delta \in \mathbb{R}^{k\times(n+1)}, \ \|C_j \Delta^T\|_F^2 \leq \rho_j^2, \ j = 1, \ldots, k\}$, where $C_j \in \mathbb{R}^{(n+1)\times(n+1)}$ such that there exist $\mu_j \in \mathbb{R}$ such that $\sum_{j=1}^k \mu_j C_j^T C_j \succ 0$.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. Ser. B **95**, 351–370 (2003)
2. Burer, S., Anstreicher, K.M.: Second-order cone constraints for extended trust-region problems. Preprint, Optimization online, March (2011)
3. Beck, A., Eldar, Y.C.: Strong duality in nonconvex quadratic optimization with two quadratic constraints. SIAM J. Optim. **17**, 844–860 (2006)
4. Beck, A.: Convexity properties associated with nonconvex quadratic matrix functions and applications to quadratic programming. J. Optim. Theory Appl. **142**, 1–29 (2009)
5. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization: Analysis, Algorithms and Engineering Applications. SIAM-MPS, Philadelphia (2000)
6. Ben-Tal, A., Ghaoui, L.E., Nemirovski, A.: *Robust Optimization*. Princeton Series in Applied Mathematics (2009)
7. Ben-Tal, A., Nemirovski, A., Roos, C.: Robust solutions of uncertain quadratic and conic quadratic problems. SIAM J. Optim. **13**(2), 535–560 (2002)
8. Bertsimas, D., Brown, D., Caramanis, C.: Theory and applications of robust optimization. SIAM Rev. **53**, 464–501 (2011)
9. Bertsimas, D., Pachamanova, D., Sim, M.: Robust linear optimization under general norms. Oper. Res. Lett. **32**, 510–516 (2004)
10. Conn, A.R., Gould, N.I.M., Toint, P.L.: Trust-Region Methods. MPS-SIAM Series in Optimization, SIAM, Philadelphia, PA (2000)
11. Dines, L.L.: On the mapping of quadratic forms. Bull. Am. Math. Soc. **47**, 494–498 (1941)
12. Fradkov, A.L., Yakubovich, V.A.: The S-procedure and duality relations in nonconvex problems of quadratic programming. Leningrad, Russia. Vestn. Leningr. Univ. **6**, 101–109 (1979)
13. El Ghaoui, L., Lebret, H.: Robust solution to least-squares problems with uncertain data. SIAM J. Matrix Anal. Appl. **18**(4), 1035–1064 (1997)

14. Jeyakumar, V., Lee, G.M., Li, G.Y.: Alternative theorems for quadratic inequality systems and global quadratic optimization. SIAM J. Optim. **20**(2), 983–1001 (2009)
15. Jeyakumar, V., Li, G.: Strong duality in robust convex programming: complete characterizations. SIAM J. Optim. **20**, 3384–3407 (2010)
16. Jeyakumar, V., Li, G.: Exact SDP Relaxations for classes of nonlinear semidefinite programming problems. Oper. Res. Lett. (2012). doi:10.1016/j.orl.2012.09.006
17. Jeyakumar, V., Huy, N.Q., Li, G.: Necessary and sufficient conditions for S-lemma and nonconvex quadratic optimization. Optim. Eng. **10**, 491–503 (2009)
18. Jeyakumar, V., Li, G.: A robust von-Neumann minimax theorem for zero-sum games under bounded payoff uncertainty. Oper. Res. Lett. **39**(2), 109–114 (2011)
19. Jeyakumar, V., Rubinov, A.M., Wu, Z.Y.: Non-convex quadratic minimization problems with quadratic constraints: global optimality conditions. Math. Program. Ser. A **110**, 521–541 (2007)
20. Jeyakumar, V., Wolkowicz, H.: Zero duality gaps in infinite-dimensional programming. J. Optim. Theory Appl. **67**, 87–108 (1990)
21. More, J.J.: Generalizations of the trust region subproblem. Optim. Methods Softw. **2**, 189–209 (1993)
22. Pardalos, P., Romeijn, H.: Handbook in Global Optimization, vol. 2. Kluwer Academic Publishers, Dordrecht (2002)
23. Peng, J.M., Yuan, Y.X.: Optimality conditions for the minimization of a quadratic with two quadratic constraints. SIAM J. Optim. **7**, 579–594 (1997)
24. Pólik, I., Terlaky, T.: A survey of the S-Lemma. SIAM Rev. **49**, 371–418 (2007)
25. Polyak, B.T.: Convexity of quadratic transformation and its use in control and optimization. J. Optim. Theory Appl. **99**, 563–583 (1998)
26. Powell, M.J.D., Yuan, Y.: A trust region algorithm for equality constrained optimization. Math. Program. **49**(91), 189–211 (1990)
27. Stern, R.J., Wolkowicz, H.: Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. SIAM J. Optim. **5**, 286–313 (1995)
28. Sturm, J.F., Zhang, S.Z.: On cones of nonnegative quadratic functions. Math. Oper. Res. **28**, 246–267 (2003)
29. Yakubovich, Y.A.: The S-procedure in nonlinear control theory. Vestn. Leningr. Univ. **4**(1), 62–77 (1971)
30. Ye, Y.Y., Zhang, S.Z.: New results of quadratic minimization. SIAM J. Optim. **14**, 245–267 (2003)
31. Yuan, Y.X.: On a subproblem of trust region algorithms for constrained optimization. Math. Program. **47**, 53–63 (1990)