TRUST REGION SUBPROBLEM WITH A FIXED NUMBER OF ADDITIONAL LINEAR INEQUALITY CONSTRAINTS HAS POLYNOMIAL COMPLEXITY*

YONG HSIA† AND RUEY-LIN SHEU‡ §

Abstract. The trust region subproblem with a fixed number m additional linear inequality constraints, denoted by $(\mathrm{T_m})$, have drawn much attention recently. The question as to whether Problem $(\mathrm{T_m})$ is in Class P or Class NP remains open. So far, the only affirmative general result is that $(\mathrm{T_1})$ has an exact SOCP/SDP reformulation and thus is polynomially solvable. By adopting an early result of Martínez on local non-global minimum of the trust region subproblem, we can inductively reduce any instance in $(\mathrm{T_m})$ to a sequence of trust region subproblems $(\mathrm{T_0})$. Although the total number of $(\mathrm{T_0})$ to be solved takes an exponential order of m, the reduction scheme still provides an argument that the class $(\mathrm{T_m})$ has polynomial complexity for each fixed m. In contrast, we show by a simple example that, solving the class of extended trust region subproblems which contains more linear inequality constraints than the problem dimension; or the class of instances consisting of an arbitrarily number of linear constraints, namely $\bigcup_{m=1}^{\infty}(\mathrm{T_m})$, is NP-hard. When m is small such as $m = 1, 2$, our inductive algorithm should be more efficient than the SOCP/SDP reformulation since at most 2 or 5 subproblems of $(\mathrm{T_0})$, respectively, are to be handled. In the end of the paper, we improve a very recent dimension condition by Jeyakumar and Li under which $(\mathrm{T_m})$ admits an exact SDP relaxation. Examples show that such an improvement can be strict indeed.

Key words. Trust region subproblem; Computational complexity; Nonconvex quadratic programming; Local non-global minimizer; Semidefinite relaxation; Hidden convexity.

AMS subject classifications. 90C09, 90C10, 90C20

1. Introduction.

The *classical* trust region subproblem, which minimizes a nonconvex quadratic function over the unit ball

$$(\mathrm{T_0}) \quad \min f(x) = \frac{1}{2}x^TQx + c^Tx \quad \text{s.t. } x^Tx \le 1,$$

is an important feature in trust region methods [5, 20]. It is well known that finding an ϵ-optimal solution of $(\mathrm{T_0})$ has polynomial complexity [6, 18] and efficient algorithms

*This research was supported by Taiwan National Science Council under grant 102-2115-M- 006-010, by National Center for Theoretical Sciences (South), by National Natural Science Foundation of China under grant 11001006 and 91130019/A011702, and by the fund of State Key Laboratory of Software Development Environment under grant SKLSDE-2013ZX-13.

†State Key Laboratory of Software Development Environment, LMIB of the Ministry of Education, School of Mathematics and System Sciences, Beihang University, Beijing, 100191, P. R. China (dearyxia@gmail.com).

‡Department of Mathematics, National Cheng Kung University, Taiwan (rsheu@mail.ncku.edu.tw).

§Corresponding author.

for solving ($\mathrm{T_0}$) can be found in [7, 11, 13]. Moreover, problem ($\mathrm{T_0}$) is also a special case of a quadratic problem subject to a quadratic inequality constraint (QP1QC). It was proved that, under Slater's condition, (QP1QC) admits a tight SDP relaxation and its optimal solution can be found through a matrix rank one decomposition procedure [16, 12].

Extensions of ($\mathrm{T_0}$), sometimes termed as the *extended* trust region subproblem, consider problems such as adding to ($\mathrm{T_0}$) several linear inequality constraints [19] or imposing a full-dimensional ellipsoid [4]. In particular, we are interested in the following variant:

$$(\mathrm{T_m}) \quad \min f(x) = \frac{1}{2}x^TQx + c^Tx$$

$$\text{s.t. } x^Tx \le 1, \tag{1.1}$$

$$a_i^Tx \le b_i,\ i = 1,\dots,m, \tag{1.2}$$

which arises from applying trust region methods to solve constrained nonlinear programs [5]. We notice that some NP-hard combinatorial optimization problems also have the similar formulation. A typical example is to rewrite the standard quadratic program

$$\begin{aligned}(\mathrm{QPS}) \quad & \min x^TQx \\ & \text{s.t. } e^Tx = 1, \\ & \quad x \ge 0\end{aligned}$$

as a special case of ($\mathrm{T_m}$), where e is the vector of all ones. To do so, let $y = (x_1,\dots,x_{n-1})^T$. By replacing $e^Tx = 1,\ x \ge 0$ with $x_n = 1 - e^Ty \ge 0,\ y \ge 0$, we can express the standard quadratic program (QPS) in terms of variable y as

$$\min \begin{pmatrix} y \\ 1 - e^Ty \end{pmatrix}^T Q \begin{pmatrix} y \\ 1 - e^Ty \end{pmatrix}$$

$$\text{s.t. } 1 - e^Ty \ge 0, \tag{1.3}$$

$$y \ge 0. \tag{1.4}$$

It is easy to see that $0 \le y \le e$ and

$$y^Ty = \sum_{i=1}^{n-1} y_i^2 \le \sum_{i=1}^{n-1} y_i \le 1.$$

In other words, by imposing a redundant constraint $y^Ty \le 1$ to (1.3)-(1.4), we enforce (QPS) to have an equivalent extended trust region subproblem reformulation as follows:

$$(\mathrm{QPS-TRS}) \quad \min \begin{pmatrix} y \\ 1 - e^Ty \end{pmatrix}^T Q \begin{pmatrix} y \\ 1 - e^Ty \end{pmatrix}$$

$$\begin{aligned} \text{s.t. } & 1 - e^T y \geq 0, \\ & y \geq 0, \\ & y^T y \leq 1. \end{aligned}$$

Since (QPS) is NP-hard (as it captures the NP-hard combinatorial problem to find the cardinality number of the maximum stable set in a graph), so is (QPS-TRS). Let $(\mathrm{T_{n+1}})$ represent the class of extended trust region subproblems which always has the number of linear inequality constraints exceeding the problem dimension by one. We can immediately conclude from the example (QPS-TRS) that $(\mathrm{T_{n+1}})$ must be NP-hard. The implication is that solving the subclass of extended trust region subproblems which contains more linear inequality constraints than the problem dimension; or solving the most general extended trust region subproblems consisting of an arbitrarily number of linear constraints, namely $\bigcup_{m=1}^{\infty}(\mathrm{T_m})$, should be difficult.

A natural question arises from computational complexity: "Fix a positive integer m. What is the complexity of solving $(\mathrm{T_m})$ for all possible dimensions?" The problem turns out to be more difficult than most people thought. The only affirmative result so far in literature is that $(\mathrm{T_m})$ with $\mathrm{m} = 1$ is polynomially solvable [2, 16]. For $\mathrm{m} = 2$, the polynomial solvability of some special cases of $(\mathrm{T_2})$ were established when a_1 and a_2 are parallel [2, 19]; or when $a_1^T x \leq b_1$ and $a_2^T x \leq b_2$ are non-intersecting in the unit-ball [3]. When $m \geq 2$ and any two inequalities are non-intersecting in the interior of the unit ball, this subclass of $(\mathrm{T_m})$ is also polynomial solvable as shown in [3]. Very recently, Jeyakumar and Li [8] showed that, under the following **dimension condition**, $(\mathrm{T_m})$ is also polynomial solvable [8]:

$$\text{[DC]} \quad \dim \operatorname{Ker}(Q - \lambda_{\min}(Q) I_n) \geq \dim \operatorname{span}\{a_1, \ldots, a_m\} + 1, \tag{1.5}$$

where $\operatorname{Ker}(Q)$ denotes the kernel of Q; $\lambda_{\min}(Q)$ the smallest eigenvalue of Q; $\dim L$ the dimension of a subspace L; and I_n the identity matrix of order n.

All the approaches mentioned above elaborate the polynomial complexity of some $(\mathrm{T_m})$ through an exact SOCP/SDP reformulation [2, 3, 16] such as, for $m = 1$,

$$\begin{aligned} \min\ & \frac{1}{2}\operatorname{trace}(QX) + c^T x \\ \text{s.t. } & \operatorname{trace}(X) \leq 1,\ X \succeq x x^T, \\ & \|b_1 x - X a_1\|_2 \leq b_1 - a_1^T x; \end{aligned}$$

or through a tight SDP relaxation [8]:

$$\text{(P)} \quad \begin{aligned} \min\ & \frac{1}{2}\operatorname{trace}(QX) + c^T x \\ \text{s.t. } & \operatorname{trace}(X) \leq 1, \\ & a_i^T x \leq b_i,\ i = 1, \ldots, m, \end{aligned} \tag{1.6}$$

$$\begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0.$$

In other words, the polynomial solvability is built by way of finding the hidden convexity from the non-convex problems $(\mathrm{T_m})$. The scheme is easily seen to be exorbitant as there are examples in the same papers pointing out that neither the SOCP/SDP reformulation nor the SDP relaxation is tight for general cases of $m = 2$ [2] and of $m = 1$ [8], respectively. According to Burer and Anstreicher [2], "the computational complexity of solving an extended trust region problem is highly dependent on the geometry of the feasible set."

Our basic idea to cope with the complication of the geometry is to think the structure of the polytope directly and reduce the problem $(\mathrm{T_m})$ inductively until $(\mathrm{T_0})$ is reached. To avoid triviality, we assume, throughout the paper, that Q has at least one negative eigenvalue, i.e., $\lambda_{\min}(Q) < 0$. Then, the global minimum of $(\mathrm{T_m})$ must lie on the boundary. The boundary could be part of the unit sphere $x^T x = 1$ or part of the boundary of the polytope intersecting with the unit ball $x^T x \le 1$. In the former case when the global minimum of $(\mathrm{T_m})$ happens to be *solely* on the unit sphere (meaning that it does not lie simultaneously on any boundary of the polytope), it must be at least a local minimum of the trust region subproblem $(\mathrm{T_0})$. This case is polynomially checkable due to an early result of Martínez [10]. In the latter case if it lies on the boundary of the polytope intersecting with the unit ball, it can be found by solving one of the following m subproblems:

$$\begin{aligned} v(\mathrm{T_m^j}) := \min f(x) = \frac{1}{2} x^T Q x + c^T x& \\ \text{s.t. } x^T x \le 1,& \\ a_j^T x = b_j,& \\ a_i^T x \le b_i,\ i = 1, \ldots, j-1, j+1, \ldots, m,& \end{aligned} \tag{1.7}$$

where the superscript j varies from 1 to m. By eliminating one variable using (1.7), problem $(\mathrm{T_m^j})$ can be reduced to a type of problem $(\mathrm{T_{m-1}})$ of $n-1$ dimensional. The procedure can be inductively applied to $(\mathrm{T_{m-1}})$, $(\mathrm{T_{m-2}})$, $\ldots$, and so forth until we run down to one of the three possibilities: either an infeasible subproblem, or a convex programming subproblem; or a subproblem of no linear inequality constraint, i.e., $(\mathrm{T_0})$. Since m is fixed, the number of reduction can not grow exponentially and we thus conclude the polynomial complexity of $(\mathrm{T_m})$ for any fixed positive integer m.

When m is a variable, our induction argument eventually leads to solve an exponential number of $(\mathrm{T_0})$ so it still has to face the curse of dimensionality. However, when m is small, this method can be very efficient. For example, when $m = 1$, it requires to only solve two subproblems of $(\mathrm{T_0})$ (see Section 2 below). By the result of

Martínez [10], ($\mathrm{T_0}$) has a spherical structure of global optimal solution set and possesses at most one local non-global minimizer. To solve it amounts to finding the root of a one-variable convex secular function and hence avoid a generally more tedious large scale SOCP/SDP.

Finally, we provide a new result which improves the dimension condition [DC] in [8] to become

$$\text{[NewDC]} \quad \mathrm{rank}\left([Q - \lambda_{\min}(Q)I_n \; a_1 \; \ldots \; a_m]\right) \le n-1, \tag{1.8}$$

under which ($\mathrm{T_m}$) admits an exact SDP relaxation. We use some example to demonstrate that the improvement can be strict.

2. Global Optimization of ($\mathrm{T_m}$) **and Complexity.** To solve ($\mathrm{T_m}$), we begin with ($\mathrm{T_0}$). Let x^* be a global minimizer of ($\mathrm{T_m}$) whereas X_0^* is the set of all the global minimizers of ($\mathrm{T_0}$). Suppose $X_0^* \bigcap \{x \mid a_i^T x \le b_i,\ i = 1,\ldots,m\} \ne \emptyset$. Then $v(\mathrm{T_m}) = v(\mathrm{T_0})$ and any solution in the intersection is a global optimal solution to ($\mathrm{T_m}$). Otherwise, $v(\mathrm{T_m}) > v(\mathrm{T_0})$. In the former case, the task is to find a common point x^* from the intersection in polynomial time. In the latter case, since we assume that the smallest eigenvalue of Q is negative, we need to examine every piece of the boundary of ($\mathrm{T_m}$). In particular, if $a_i^T x^* < b_i,\ i = 1,\ldots,m$, then x^* must be a local solution of ($\mathrm{T_0}$) residing on the sphere. Both cases require a complete understanding about the structure of X_0^* as well as the local non-global minimizer of ($\mathrm{T_0}$), which have been studied in several important literature such as Moré and Sorensen [11]; Stern and Wolkowicz [15]; Martínez [10]; and Lucidi, Palagi, and Roma [9]. Below is a brief review.

2.1. The local and global minimizer of TRS. Denote by

$$(0 >)\sigma_1 = \ldots = \sigma_k < \sigma_{k+1} \le \ldots \le \sigma_n$$

the eigenvalues of Q, $\Sigma = \mathrm{diag}(\sigma_1,\ldots,\sigma_n)$, $u_1,\ldots,u_n$ the corresponding eigenvectors and $U = [u_1,\ldots,u_n]$. By introducing $y = Ux$, $d = Uc$, we can express ($\mathrm{T_0}$) in terms of y as

$$v(\mathrm{T_0}) = \min_{y^T y \le 1} \frac{1}{2} y^T \Sigma y + d^T y.$$

Similarly, applying the same coordinate change to $a_i^T x \le b_i,\ i = 1,2,\ldots,m$ results in $a_i^T U^T y \le b_i$. Denote by $\tilde{a}_i^T = a_i^T U^T$ and express the polytope $\{x \mid a_i^T x \le b_i,\ i = 1,\ldots,m\}$ in terms of y as $\{y \mid \tilde{a}_i^T y \le b_i,\ i = 1,\ldots,m\}$.

Since Slater's condition is satisfied, all local minimizers of ($\mathrm{T_0}$) must satisfy the following KKT conditions associated with a Lagrange multiplier $\mu \ge 0$:

$$(\Sigma + \mu I_n)y + d = 0;$$

$$\mu(y^T y - 1) = 0;$$
$$y^T y \leq 1.$$

By assumption, Σ is not positive semidefinite, so any y in the interior of the unit-ball ($y^T y < 1$) cannot be a local minimizer. Therefore, the necessary condition for local solutions y can be reduced to finding $\mu \geq 0$ such that

$$(\Sigma + \mu I_n)y + d = 0; \tag{2.1}$$
$$y^T y = 1. \tag{2.2}$$

When μ is not equal to any of the eigenvalues of Q, the matrix $\Sigma + \mu I_n$ is invertible and hence $d \neq 0$. Then, (y, μ) with $y_i = \frac{-d_i}{\sigma_i + \mu}$ is a solution to (2.1)-(2.2) if and only if μ is a root of the following *secular function* [15]

$$\varphi(\mu) = \sum_{i=1}^{n} \frac{d_i^2}{(\sigma_i + \mu)^2} - 1. \tag{2.3}$$

From the first and second derivatives of $\varphi(\mu)$, we also have

$$\phi'(\mu) = \sum_{i=1}^{n} \frac{-2d_i^2}{(\sigma_i + \mu)^3}, \tag{2.4}$$

and

$$\phi''(\mu) = \sum_{i=1}^{n} \frac{6d_i^2}{(\sigma_i + \mu)^4} > 0,$$

which shows that the secular function is strictly convex. It turns out that the global minimum and the local non-global minimum of (T_0) can be distinguished by the position where their corresponding Lagrange multiplier μ locates.

THEOREM 2.1. *([11]) Let (y^*, μ^*) satisfy (2.1) − (2.2). Then, y^* is a global minimum solution to (T_0) if and only if the Lagrange multiplier μ^* satisfies*

$$\mu^* \geq -\sigma_1 (> 0).$$

Based on Theorem 2.1, in order to characterize the global optimal solution set X_0^* of (T_0), we only have to investigate the secular function on the interval $\mu \in [-\sigma_1, \infty)$. Notice that secular functions in (2.3) depend on problem data σ_i and d_i. Our discussion below classifies and analyzes different types of secular functions. The result shows that X_0^* is either a singleton or a k-dimensional sphere where k is the multiplicity of the smallest eigenvalue σ_1.

- Suppose $d_1^2 + \ldots + d_k^2 > 0$. Then, $\lim_{\mu \to -\sigma_1^+} \varphi(\mu) = \infty$; $\lim_{\mu\to\infty} \varphi(\mu) = -1$ and $\varphi(\mu)$ is strictly decreasing on $(-\sigma_1, \infty)$. Therefore, the secular function $\varphi(\mu)$ has a unique solution μ^* on $(-\sigma_1, \infty)$. In this case, y^* defined by

$$y_i^* = -\frac{d_i}{\sigma_i + \mu^*},\ i = 1, \ldots, n. \tag{2.5}$$

  is the unique global minimum solution of (T_0).
- Suppose $d_1^2 + \ldots + d_k^2 = 0$. There are two cases.
  (1) $\mu^* > -\sigma_1$. It can happen only when $d \neq 0$ and $\lim_{\mu \to -\sigma_1^+} \varphi(\mu) > 0$. Then, y^* satisfies (2.5) is the unique global minimizer.
  (2) $\mu^* = -\sigma_1$. By Theorem 2.1, any y^* satisfying

$$(y_1^*)^2 + \ldots + (y_k^*)^2 = 1 - \sum_{i=k+1}^{n} \frac{d_i^2}{(\sigma_i - \sigma_1)^2}, \tag{2.6}$$

$$y_i^* = -\frac{d_i}{\sigma_i - \sigma_1},\ i = k+1, \ldots, n \tag{2.7}$$

  is a global minimizer. Namely, the global minimum solution set X_0^* forms a k-dimensional sphere centered at $(0, \cdots, 0, -\frac{d_{k+1}}{\sigma_{k+1}-\sigma_1}, \cdots, -\frac{d_n}{\sigma_n - \sigma_1})$ with the radius $\sqrt{1 - \sum_{i=k+1}^{n} \frac{d_i^2}{(\sigma_i-\sigma_1)^2}}$.

Secular functions also provide useful information on the local non-global minimizer. In the next theorem, Martínez [10] showed that there is at most one local non-global minimizer $\overline{y}$ in (T_0). The associated Lagrange multiplier $\overline{\mu}$ is nonnegative and lies in $(-\sigma_2, -\sigma_1)$. Moreover, Lucidi et al. [9] showed that strict complementarity holds at the local non-global minimizer.

THEOREM 2.2 ([10, 9]). *Suppose $k \geq 2$ or $d_1 = 0$ when $k = 1$, there is no local non-global minimizer. Otherwise, there is at most one local non-global minimizer $\overline{y}$ to (T_0), and the associated Lagrange multiplier $\overline{\mu}$ satisfies $\overline{\mu} \in (\max\{-\sigma_2, 0\}, -\sigma_1)$ and*

$$\varphi(\overline{\mu}) = 0, \tag{2.8}$$

$$\varphi'(\overline{\mu}) \geq 0. \tag{2.9}$$

Moreover, if $\overline{\mu} \in (\max\{-\sigma_2, 0\}, -\sigma_1)$ (2.8) and $\varphi'(\overline{\mu}) > 0$, then $\overline{y}$ defined as

$$\overline{y}_i = -\frac{d_i}{\sigma_i + \overline{\mu}},\ i = 1, \ldots, n \tag{2.10}$$

is the unique local non-global minimizer.

From the formula $\varphi'(\mu)$ in (2.4), there are several types of convex secular functions on $(-\sigma_2, -\sigma_1)$. It can be convex decreasing, for example, when $d_1^2 + \ldots + d_k^2 = 0$ and some $d_i \neq 0,\ i \geq k+1$ in which case (T_0) can not have a local non-global minimizer; or convex increasing, for example, when $d_1^2 + \ldots + d_k^2 > 0$ and $d_i = 0,\ i \geq k+1$;

or have a global minimum on $(-\sigma_2, -\sigma_1)$. In any case, the necessary conditions (2.8)-(2.9), once valid, must possess only a unique solution $\overline{y}$ of the form (2.10) for $\overline{\mu} \in (\max\{-\sigma_2, 0\}, -\sigma_1)$ since $\varphi(\mu)$ is strictly convex on $(-\sigma_2, -\sigma_1)$. That is, $\overline{y}$ is only a candidate for the local non-global minimizer of (T_0). It could otherwise represent a saddle point rather than a local minimum.

2.2. The intersection of X_0^* and a polytope. In this subsection, we are concerned with the following decision problem:

$$X_0^* \bigcap \{y \mid \tilde{a}_i^T y \le b_i,\ i = 1, \ldots, m\} \neq \emptyset. \tag{2.11}$$

Since X_0^* is a k-dimensional sphere as expressed in (2.6)-(2.7), we first reduce the n-dimensional polytope $\{y \mid \tilde{a}_i^T y \le b_i,\ i = 1, \ldots, m\}$ to the same k dimension by fixing y_i at $-\frac{d_i}{\sigma_i - \sigma_1}$ for $i = k+1, \ldots, n$ and assume that

$$\{u \in R^k \mid \tilde{a}_i^T \left(u^T, -\frac{d_{k+1}}{\sigma_{k+1} - \sigma_1}, \cdots, -\frac{d_n}{\sigma_n - \sigma_1} \right)^T \le b_i,\ i = 1, \ldots, m\}$$

is non-empty. Otherwise, there would be no intersection between X_0^* and the polytope.

If X_0^* is a singleton, it is easy to check because both sets are convex. However, when X_0^* is a nonconvex sphere, it is in general difficult to determine whether (2.11) is true. Our procedure to answer yes/no for (2.11) might depend exponentially on the number m of linear constraints, but only polynomially on the problem dimension n. Therefore, when m is a fixed constant, our method has polynomial complexity to answer (2.11).

To begin, let $L = \{u \in R^p \mid Hu \le g\} = \{u \in R^p \mid h_i^T u \le g_i,\ i = 1, \ldots, m\}$; $B = \{u \in R^p \mid u^T u \le r\}$ and $\partial B = \{u \in R^p \mid u^T u = r\}$. That is, we conduct the analysis for any p-dimensional space.

Lemma 2.1. *Let $H \in R^{m \times p}$ be column dependent and $g \in R^m$. Then, the polytope L is either infeasible or unbounded.*

Proof. Let $H_1, \ldots, H_p$ be the columns of H. Since H is column dependent, there is a nonzero $z \in R^p$ such that $Hz = 0$. If u_0 is feasible with $Hu_0 \le g$, then $H(u_0 + \beta z) \le g$ for any scalar β. The polytope L is hence unbounded. □

Lemma 2.2. *Let $H \in R^{m \times p}$ be column independent and $g \in R^m$. Assume that there is a u_0 satisfying $Hu_0 \le g$. Then, the polytope L is bounded if and only if the optimal value f^* of the following linear programming is nonnegative*

$$\begin{aligned} f^* = \quad \min \quad & e^T H u \\ \text{s.t.} \quad & Hu \le 0, \\ & \|u\|_\infty \le 1, \end{aligned}$$

where $\|u\|_\infty := \max_i |u_i|$. Moreover, if $f^* < 0$, the optimal solution d to the linear programming is an extreme direction of the unbounded polytope L.

Proof. Suppose L is unbounded. It contains at least one extreme ray, denoted by $\{u_0 + \beta z \mid \beta \geq 0\}$ where $z \neq 0$ and $\|z\|_\infty \leq 1$ such that $H(u_0 + \beta z) \leq g$. This can happen only when $Hz \leq 0$. Since $z \neq 0$ and H is column independent, we have $Hz \neq 0$, i.e., $e^T Hz < 0$ and hence $f^* < 0$.

On the other hand, suppose $f^* < 0$ and d is optimal to the linear programming. It implies that $Hd \leq 0$ and $d \neq 0$. Consequently, $\{u_0 + \beta d \mid \beta \geq 0\}$ is contained in the polytope L, which is therefore unbounded. □

Lemma 2.3. *Let $H \in R^{m\times p}$ and $g \in R^m$, where m is fixed and p is arbitrary. For any given $r > 0$, it is polynomially checkable whether $\{u \in R^p \mid Hu \leq g, u^T u = r\}$ is empty. Moreover, if the set is nonempty, a feasible point can be found in polynomial time.*

Proof. Since both L and B are convex, we can either find a $\hat{u} \in L \bigcap B$ or conclude that $L \bigcap B = \emptyset$ in polynomial time. For example, consider the convex program

$$\hat{\delta} = \min_{\{(u,v) \mid u \in L, v \in B\}} \|u - v\|^2. \tag{2.12}$$

If $\hat{\delta} > 0$, then $L \bigcap B = \emptyset$. Otherwise, when $\hat{\delta} = 0$, any optimal solution $(\hat{u}, \hat{v})$ to (2.12) would imply that $\hat{u} = \hat{v}$ is in the intersection. If, furthermore, it happens that $\hat{u} \in \partial B$, then $L \bigcap \partial B \neq \emptyset$. Otherwise, we have $\hat{u}^T \hat{u} < r$.

Since B is a *full* dimensional ball in R^p, the only possibility that $\hat{\delta} = 0$ but $L \bigcap \partial B = \emptyset$ is when L is bounded and contained entirely in the interior of B. By Lemmas 2.1 and 2.2, the polytope L is bounded if and only if the columns of H are linearly independent and the linear programming in Lemma 2.2 has a nonnegative optimal value. If L is indeed bounded, $m \geq p$ and we enumerate all the vertices of L.

Suppose $\widetilde{u}$ is a vertex point and assume, without loss of generality, that $h_i^T \widetilde{u} = g_i$ for $i = 1, ..., r$ and $h_i^T \widetilde{u} < g_i$ for $i = r + 1, ..., m$. Then we conclude that $\text{rank}\{h_1, \ldots, h_r\} = p$. If this is not true, there is a vector $\eta \neq 0$ such that $h_i^T \eta = 0$ for $i = 1, ..., r$. Then both $\widetilde{u} + \epsilon\eta$ and $\widetilde{u} - \epsilon\eta$ are feasible solutions of L for sufficiently small $\epsilon > 0$, contradicting to $\widetilde{u}$ being a vertex of L. Therefore, to enumerate all extreme points of L, it is sufficient to pick all p linearly independent vectors out of $\{h_1, \ldots, h_m\}$ and then check the feasibility. It follows that L has at most $C(p, m) = O(m^{\min\{p,m-p\}})$ vertices, denoted by $z_1, \ldots, z_t$. Since m is assumed to be fixed and $p \leq m$, the number t cannot exceed a constant factor depending on m. It is also obvious that if $z_i^T z_i < r$ for $i = 1, \ldots, t$, the polytope L is in the strict interior of B and thus $L \bigcap \partial B = \emptyset$. If there exists an index j_0 such that $z_{j_0}^T z_{j_0} = r$, then $z_{j_0} \in L \bigcap \partial B$.

Finally, if there is some index j_0 such that $z_{j_0}^T z_{j_0} > r$, since $\hat{u}^T \hat{u} < s$, the line segment $[\hat{u}, z_{j_0}]$ must intersect ∂B at one point. Similarly, when L is unbounded, solving the linear programming in Lemma 2.2 yields an extreme direction d at $\hat{u}$, along which an intersection point at $L \bigcap \partial B$ can be easily found. The proof is complete. □

2.3. Iterative Reduction Procedure for Global Optimization. Assume that $X_0^* \bigcap \{y \in R^k \mid \tilde{a}_i^T \left(y^T, -\frac{d_{k+1}}{\sigma_{k+1}-\sigma_1}, \cdots, -\frac{d_n}{\sigma_n-\sigma_1}\right)^T \le b_i,\ i=1,\ldots,m\} = \emptyset$. That is, the global minimum of $(\mathrm{T_0})$ does not help solve $(\mathrm{T_m})$ so that we have to analyze directly the boundary of $\{x \in R^n | x^Tx \le 1, a_i^Tx \le b_i,\ i=1,2,\ldots,m\}$ and the local non-global minimizer of $(\mathrm{T_0})$.

The geometry of the boundary could be expressly complicate, specified by one or several inequalities (linear or quadratic inequalities) becoming active. However, if we consider the boundaries "one piece at a time", the global minimizer x^* of $(\mathrm{T_m})$ must belong to and thus globally minimize at least one of the following candidate subproblems:

$$
\begin{aligned}
v(\mathrm{T_m^0}) := \min f(x) = \frac{1}{2}x^TQx + c^Tx& \\
\text{s.t. } x^Tx \le 1,& \\
a_i^Tx < b_i,\ i=1,2,\ldots,m;&
\end{aligned}
\tag{2.13}
$$

and for $j=1,2,\ldots,m$

$$
\begin{aligned}
v(\mathrm{T_m^j}) := \min f(x) = \frac{1}{2}x^TQx + c^Tx& \\
\text{s.t. } x^Tx \le 1,& \\
a_j^Tx = b_j,& \\
a_i^Tx \le b_i,\ i=1,\ldots,j-1,j+1,\ldots,m.&
\end{aligned}
\tag{2.14}
$$

It is clear that $v(\mathrm{T_m}) = \min\{v(\mathrm{T_m^0}), v(\mathrm{T_m^1}), \ldots, v(\mathrm{T_m^m})\}$.

According to the previous analysis, there is at most one local non-global minimizer of the trust region subproblem $(\mathrm{T_0})$, and the only legitimate candidate is $\overline{x}_0 = U^T\overline{y}$ where $\overline{y}$ is the unique solution to (2.8)-(2.10) for some $\overline{\mu} \in (\max\{-\sigma_2, 0\}, -\sigma_1)$. Therefore, if $f(\overline{x}_0) < \min\{v(\mathrm{T_m^1}), \ldots, v(\mathrm{T_m^m})\}$ and $\overline{x}_0$ satisfies (2.13), then

$$
v(\mathrm{T_m^0}) \le f(\overline{x}_0) < \min\{v(\mathrm{T_m^1}), \ldots, v(\mathrm{T_m^m})\}
$$

and consequently x^* must solve $(\mathrm{T_m^0})$. Since x^* is in the interior of the polytope, and since the polytope has no intersection with X_0^*, x^* must be a local non-global minimizer of $(\mathrm{T_0})$. Since $\overline{x}_0$ is the unique candidate, it follows that x^* coincides with $\overline{x}_0$ and $\overline{x}_0$ is indeed a local minimizer.

The above argument also implies that, if the unique candidate $\overline{x}_0$ does not satisfy (2.13), since there is no alternative candidate, the optimal solution x^* can not be found from solving $(\mathrm{T_m^0})$. In addition, when $f(\overline{x}_0) \ge \min\{v(\mathrm{T_m^1}), \ldots, v(\mathrm{T_m^m})\}$, x^* should be retrieved from solving one of $(\mathrm{T_m^j})$, $j=1,\ldots,m$ so that $(\mathrm{T_m^0})$ need not be considered either.

As a summary, we have

$$v(\mathrm{T_m}) = \begin{cases} v(\mathrm{T_0}), & \text{if } X_0^* \bigcap \{x \mid a_i^T x \le b_i, \forall i\} \neq \emptyset; \\ f(\overline{x}_0), & \text{if } a_i^T \overline{x}_0 < b_i, \forall i = 1, \ldots, m \text{ and} \\ & f(\overline{x}_0) < \min\{v(\mathrm{T_m^1}), \ldots, v(\mathrm{T_m^m})\}; \\ \min\{v(\mathrm{T_m^1}), \ldots, v(\mathrm{T_m^m})\}, & \text{o.w.} \end{cases} \tag{2.15}$$

It remains to show how to solve $(\mathrm{T_m^j}),\ j = 1, \ldots, m$. Our idea is to eliminate one variable using the equation (2.14) and maintains the same structure as minimizing a quadratic function over the intersection of a ball centered at 0 with some polytope.

Let $P_j \in R^{n\times(n-1)}$ be a column-orthogonal matrix such that $a_j^T P_j = 0$. Let z_0 be a feasible solution to (2.14). Then $z_0 - P_j P_j^T z_0$ is also feasible to (2.14). Using the null-space representation, we have

$$\{x \in R^n \mid a_j^T x = b_j\} = \{z_0 - P_j P_j^T z_0 + P_j z \mid z \in R^{n-1}\} \tag{2.16}$$

and

$$\begin{aligned} x^T x &= (z_0 - P_j P_j^T z_0 + P_j z)^T (z_0 - P_j P_j^T z_0 + P_j z) \\ &= z_0^T (I - P_j P_j^T)(I - P_j P_j^T) z_0 + 2 z_0^T (I - P_j P_j^T) P_j z + z^T P_j^T P_j z \\ &= z_0^T (I - P_j P_j^T) z_0 + z^T z. \end{aligned}$$

Suppose $z_0^T (I - P_j P_j^T) z_0 > 1$. Then $x^T x > 1$ for all z in the null space of the column space $\{\alpha a_j | \alpha \in R\}$. It indicates that $(\mathrm{T_m^j})$ is infeasible since $\{x \mid\ x^T x \le 1\} \bigcap \{x \mid\ a_j^T x = b_j\} = \emptyset$. Otherwise, we can equivalently express $(\mathrm{T_m^j})$ as:

$$\begin{aligned} v(\mathrm{T_m^j}) = &\min f(z_0 - P_j P_j^T z_0 + P_j z) \\ &\text{s.t. } z^T z \le 1 - z_0^T (I - P_j P_j^T) z_0, \\ &\quad a_i^T (z_0 - P_j P_j^T z_0 + P_j z) \le b_i,\ i = 1, \ldots, m,\ i \neq j \end{aligned} \tag{2.17}$$

which is again an extended trust region subproblem of $n-1$ variables equipped with $m-1$ linear inequality constraints. If the subproblem $(\mathrm{T_m^j})$ is a convex program, it can be globally solved. Otherwise, it is reduced to an instance in $(\mathrm{T_{m-1}})$ with at least one negative eigenvalue. Sometimes, more redundant constraints can be also removed after the reduction. For example, we can delete the i-th constraint in (2.17) provided $a_i^T P_j = 0$ and $a_i^T (z_0 - P_j P_j^T z_0) \le b_i$.

Iteratively applying (2.15), we will eventually terminate when further reducing the subproblem causes (i) infeasibility; (ii) a convex programming subproblem; or (iii) a classical trust region subproblem (with no linear constraint left). Let s be the smallest number such that any $s+1$ inequalities are either row-dependent; or row-independent but non-intersecting within the ball. By this inductive way, there are at most $m \times (m-1) \times \cdots \times (m-s+1)$ trust region subproblems to be solved. The special case $s = 2$ has been polynomially solved in [3] recently. Since m is assumed to be fixed, the total number of reduction iterations is bounded by a constant factor of m. We thus have proved that

THEOREM 2.3. *For each fixed m, $(\mathrm{T_m})$ is polynomially solvable.*

As examples, when $m = 1$, the inductive procedure (2.15) requires to solve two trust region subproblems: one is $(\mathrm{T_0})$ and the other one is reduced from (T_1^1). For $m = 2$, the three subproblems $(\mathrm{T_0})$; (T_2^1) and (T_2^2) need be solved. The latter two can be further reduced to two trust region subproblems each. In total, at most five trust region subproblems are necessary for solving $(\mathrm{T_2})$. Moreover, when $m = 1,\ n \geq 2$, the polytope is unbounded and there is no need to enumerate the vertices in checking the decision problem (2.11) for a possible intersection. Same as $m = 2,\ n \geq 3$.

3. Improved Dimension Condition for Exact SDP Relaxation. In this section, we improve the very recent dimension condition by Jeyakumar and Li [8] under which $(\mathrm{T_m})$ admits an exact SDP relaxation.

3.1. Hidden Convexity of some special $(\mathrm{T_m})$. Without loss of generality, we may assume that

$$\{x \mid\ a_i^T x \leq b_i,\ i = 1, \ldots, m\} \text{ has a strictly interior solution.} \tag{3.1}$$

Otherwise, there is a $j \in \{1, \ldots, m\}$ such that $a_j^T x = b_j$ for all feasible x. According to (2.16), $(\mathrm{T_m})$ can be reduced to a similar problem with $m-1$ linear constraints. Moreover, we can further assume Slater condition holds for $(\mathrm{T_m})$, i.e., $(\mathrm{T_m})$ has a strictly interior solution. The following proposition shows that the failure of Slater condition for $(\mathrm{T_m})$ implies triviality.

PROPOSITION 3.1. *Under Assumption (3.1), $(\mathrm{T_m})$ has a unique feasible solution if and only if it has no interior solution.*

Proof. Obviously, when there is only a unique solution for $(\mathrm{T_m})$, it cannot be an interior point. Now suppose $(\mathrm{T_m})$ has two feasible solutions: $y \neq z$. Since $\alpha y + (1-\alpha) z$ is also feasible for any $\alpha \in [0, 1]$, and both $\|y\|_2 \leq 1,\ \|z\|_2 \leq 1$, we can always obtain some feasible solution w such that $w^T w < 1$ and $a_i^T w \leq b_i,\ i = 1, 2, \ldots, m$.

According to Assumption (3.1), there is an x^c such that $a_i^T x^c < b_i,\ i = 1, \ldots, m$. Then for sufficient small $\epsilon > 0$, we have $y(\epsilon) := \epsilon x^c + (1-\epsilon) w$ satisfies $y(\epsilon)^T y(\epsilon) < 1$ and $a_i^T y(\epsilon) < b_i$ for $i = 1, \ldots, m$, which contradicts the no-interior assumption. □

Now we present the main result in this section.

THEOREM 3.1. *Under the assumption*

$$[\text{NewDC}]\quad \operatorname{rank}\left([Q - \lambda_{\min}(Q) I_n\ a_1\ \ldots\ a_m]\right) \leq n - 1, \tag{3.2}$$

we have $v(\mathrm{T_m}) = v(\mathrm{P})$ *where* (P) *is the standard SDP relaxation of* $(\mathrm{T_m})$ *as defined in (1.6). Besides, the condition [NewDC] (3.2) is more general than [DC] (1.5).*

Proof. Consider the following convex quadratically constrained quadratic program:

$$(\mathrm{T_m^c}) \quad \min \frac{1}{2}x^T(Q-\lambda_{\min}(Q)I_n)x + c^Tx + \frac{1}{2}\lambda_{\min}(Q)$$

$$\text{s.t. } x^Tx \le 1, \tag{3.3}$$

$$a_i^Tx \le b_i,\ i=1,\dots,m. \tag{3.4}$$

Since the feasible region of $(\mathrm{T_m^c})$ is nonempty and compact, $-\infty < v(\mathrm{T_m^c}) < +\infty$. According to Proposition 6.5.6 ([1], page 380), there is no duality gap between $(\mathrm{T_m^c})$ and its Lagrangian dual

$$\begin{aligned}(\mathrm{D^c}) \quad & \sup \frac{1}{2}\lambda_{\min}(Q) - \tau - \lambda - \sum_{i=1}^m \mu_i b_i \\ & \text{s.t. } \begin{pmatrix} Q-\lambda_{\min}(Q)I_n + 2\lambda I_n & c+\sum_{i=1}^m \mu_i a_i \\ c^T + \sum_{i=1}^m \mu_i a_i^T & 2\tau \end{pmatrix} \succeq 0. \\ & \lambda \ge 0,\ \mu \ge 0.\end{aligned}$$

The conic dual of $(\mathrm{D^c})$ is

$$\begin{aligned}(\mathrm{P^c}) \quad & \min \frac{1}{2}\mathrm{trace}((Q-\lambda_{\min}(Q)I_n)X) + c^Tx + \frac{1}{2}\lambda_{\min}(Q) \\ & \text{s.t. } \mathrm{trace}(X) \le 1, \\ & \quad a_i^Tx \le b_i,\ i=1,\dots,m, \\ & \quad \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0.\end{aligned}$$

Notice that Slater condition for $(\mathrm{T_m})$ implies that $(\mathrm{P^c})$ has a strictly feasible solution. It is trivial to see that $(\mathrm{D^c})$ also has an interior feasible solution. According to the conic duality theorem [17], $v(\mathrm{P^c}) = v(\mathrm{D^c})$ and both optimal values are attained.

Let x^* and (λ^*, μ^*) be the optimal solution of $(\mathrm{T_m^c})$ and $(\mathrm{D^c})$, respectively. Then λ^* and μ^* are also the corresponding Lagrangian multipliers of (3.3)-(3.4). Let I^* be the index set of active linear constraints at x^*, i.e., $I^* = \{i \mid a_i^Tx^* = b_i\}$.

Suppose $\lambda^* > 0$. By the complementarity, $x^{*T}x^* = 1$ and hence $v(\mathrm{T_m}) = v(\mathrm{T_m^c})$. Now assume $\lambda^* = 0$. Corresponding to the Lagrangian multiplier $(0, \mu^*)$, all the feasible solutions x of $(\mathrm{T_m^c})$ which satisfy the following KKT conditions are optimal to $(\mathrm{T_m^c})$:

$$\begin{cases} (Q-\lambda_{\min}(Q)I_n)x + c + \sum_{i=1}^m \mu_i^* a_i = 0, \\ a_i^Tx = b_i,\ \forall i \in I^*. \end{cases} \tag{3.5}$$

Since Assumption [NewDC] (3.2) implies that

$$\dim \operatorname{Ker} [Q - \lambda_{\min}(Q)I_n \ a_1 \ \dots \ a_m]^T = n - \operatorname{rank}([Q - \lambda_{\min}(Q)I_n \ a_1 \ \dots \ a_m]) \geq 1,$$

there exists $z \neq 0$ such that $[Q - \lambda_{\min}(Q)I_n \ a_1 \ \dots \ a_m]^T z = 0$. Then, $x^* + \beta z, \ \forall \beta \in R$ satisfies the KKT system (3.5) among which

$$\|x^* + \beta^* z\|_2^2 - 1 = (z^T z)\beta^{*2} + 2(x^{*T} z)\beta^* + x^{*T} x^* - 1 = 0$$

where

$$\beta^* := \frac{-x^{*T} z + \sqrt{(x^{*T} z)^2 + z^T z(1 - x^{*T} x^*)}}{z^T z}.$$

It follows that $x^* + \beta^* z$ solves $(\mathrm{T}_\mathrm{m}^\mathrm{c})$ and consequently $v(\mathrm{T}_\mathrm{m}) = v(\mathrm{T}_\mathrm{m}^\mathrm{c})$. In other words, the problem (T_m) can be equivalently solved by the convex program $(\mathrm{T}_\mathrm{m}^\mathrm{c})$ under Assumption [NewDC].

To see the SDP relaxation (P) is indeed tight, we introducing $\widetilde{\lambda} = \lambda - \frac{1}{2}\lambda_{\min}(Q)$ and reformulate (D^c) as

$$\text{(3.6)} \qquad (\mathrm{D}^\mathrm{c}) \quad \begin{aligned} &\sup -\tau - \widetilde{\lambda} - \sum_{i=1}^m \mu_i b_i \\ &\text{s.t.} \begin{pmatrix} Q + 2\widetilde{\lambda} I_n & c + \sum_{i=1}^m \mu_i a_i \\ c^T + \sum_{i=1}^m \mu_i a_i^T & 2\tau \end{pmatrix} \succeq 0 \\ &\widetilde{\lambda} \geq -\frac{1}{2}\lambda_{\min}(Q), \ \mu \geq 0. \end{aligned}$$

Comparing (3.6) with the Lagrangian dual problem of (T_m):

$$(\mathrm{D}) \quad \begin{aligned} &\sup -\tau - \lambda - \sum_{i=1}^m \mu_i b_i \\ &\text{s.t.} \begin{pmatrix} Q + 2\lambda I_n & c + \sum_{i=1}^m \mu_i a_i \\ c^T + \sum_{i=1}^m \mu_i a_i^T & 2\tau \end{pmatrix} \succeq 0 \\ &\lambda \geq 0, \ \mu \geq 0, \end{aligned}$$

we find that $v(\mathrm{D}) \geq v(\mathrm{D}^\mathrm{c})$ since it is assumed that $\lambda_{\min}(Q) < 0$. Moreover, it can be easily verified that (D) is the conic dual of the standard SDP relaxation (P). Since Slater condition holds for both (P) and (D), it implies that $v(\mathrm{P}) = v(\mathrm{D})$.

As a summary, we have the following chain of inequalities

$$v(\mathrm{T}_\mathrm{m}) \geq v(\mathrm{P}) = v(\mathrm{D}) \geq v(\mathrm{D}^\mathrm{c}) = v(\mathrm{T}_\mathrm{m}^\mathrm{c}) = v(\mathrm{T}_\mathrm{m}),$$

which proves $v(\mathrm{T}_\mathrm{m}) = v(\mathrm{P})$.

Finally, we can verify that [NewDC] (3.2) actually improves [DC] (1.5) by the following derivation:

$$\begin{aligned} & \dim \mathrm{Ker}(Q - \lambda_{\min}(Q)I_n) \geq \dim \mathrm{span}\{a_1, \ldots, a_m\} + 1 \\ \Longleftrightarrow\ & n - \mathrm{rank}(Q - \lambda_{\min}(Q)I_n) \geq \dim \mathrm{span}\{a_1, \ldots, a_m\} + 1 \\ \Longleftrightarrow\ & \mathrm{rank}(Q - \lambda_{\min}(Q)I_n) + \mathrm{rank}([a_1, \ldots, a_m]) \leq n - 1 \\ \Longrightarrow\ & \mathrm{rank}\left([Q - \lambda_{\min}(Q)I_n\ a_1\ \ldots\ a_m]\right) \leq n - 1. \end{aligned}$$

The proof of the theorem is thus completed. □

For the special case $m = 1$, the condition [NewDC] (3.2) can be easily satisfied except that when the smallest eigenvalue of Q does not repeat (i.e., $k = 1$, so that $\mathrm{rank}\,(Q - \lambda_{\min}(Q)I_n) = n-1$) and a_1 happens to be in $\mathrm{Ker}(Q-\lambda_{\min}(Q)I_n)$. Namely, we have

Corollary 3.1. *When $m = 1$, $v(\mathrm{T}_1) > v(\mathrm{P})$ happens only when the smallest eigenvalue of Q does not repeat and $a_1 \in \mathrm{Ker}(Q - \lambda_{\min}(Q)I_n)$.*

In general, in order to make the condition [NewDC] (3.2) hold for $m \geq 2$, the smallest eigenvalue of Q must repeat at least once or $a_1, \ldots, a_m$ are in the range space of $Q - \lambda_{min}(Q)I_n$. .

3.2. Examples. The following examples illustrate the applicability of the condition [NewDC]. Example 3.1 shows that [NewDC] strictly improves [DC]. Example 3.2 shows that SDP is not tight when [NewDC] fails. Finally, Example 3.3 gives a special type of (T_1) which has a tight SDP *without* any condition.

Example 3.1. *Consider an instance of* (T_1) *where $n = 2$, $b_1 = 0$ and*

$$Q = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix},\ c = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ a_1^T = \begin{bmatrix} 0 & -1 \end{bmatrix}.$$

Since $\lambda_{\min}(Q) = -2$, $Q - \lambda_{\min}(Q)I_n = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$ and

$$\dim \mathrm{Ker}(Q - \lambda_{\min}(Q)I_n) = 1,\ \dim \mathrm{span}\{a_1, \ldots, a_m\} = 1,$$

the dimension condition [DC] (1.5) fails. On the other hand, the new dimension condition [NewDC] (3.2) holds since

$$\mathrm{rank}\left([Q - \lambda_{\min}(Q)I_n\ a_1\ \ldots\ a_m]\right) = \mathrm{rank}\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & -1 \end{bmatrix}\right) = 1 \leq n - 1 = 1.$$

According to Theorem 3.1, there is no relaxation gap between (T_1) and (P) which is verified by

$$v(\mathrm{T}_1) = -1 = v(\mathrm{P}).$$

Example 3.2. *Consider an instance of* (T_1) *where* $n = 2$, $b_1 = 0$ *and*

$$Q = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}, \; c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; a_1 = \begin{bmatrix} -1 & 0 \end{bmatrix}.$$

One can check that $\lambda_{\min}(Q) = -2$, $Q - \lambda_{\min}(Q)I_n = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$ and the new dimension condition [NewDC] (3.2) fails since

$$\mathrm{rank}\left([Q - \lambda_{\min}(Q)I_n \; a_1 \; \ldots \; a_m]\right) = \mathrm{rank}\left(\begin{bmatrix} 0 & 0 & -1 \\ 0 & 4 & 0 \end{bmatrix}\right) = 2 \nleq n - 1 = 1.$$

In this example, the SDP relaxation is not tight because it has

$$v(\mathrm{T}_1) = 0 > -1 = v(\mathrm{P}).$$

Example 3.3. *Let* $c = 0$, $m = 1$ *and* $b_1 = 0$. *That is, we consider*

$$\begin{aligned} \min f(x) &= \frac{1}{2}x^T Q x \\ \text{s.t. } x^T x &\le 1, \\ a_1^T x &\le 0. \end{aligned}$$

It is not difficult to see that, let $v_{\min}$ be the unit eigenvector of Q corresponding to $\lambda_{\min}(Q)$, then $x^* = -\mathrm{sign}(a_1^T v_{\min})v_{\min}$ where

$$\mathrm{sign}(t) = \begin{cases} 1, & \text{if t} > 0 \\ -1, & \text{otherwise.} \end{cases}$$

solves (T_1) and the optimal value is $\frac{1}{2}\lambda_{\min}(Q)$. With simple verification, we can also see that $(X^*, x^*) = (v_{\min}v_{\min}^T, -\mathrm{sign}(a_1^T v_{\min})v_{\min})$ solves the related SDP relaxation with the same optimal value $\frac{1}{2}\lambda_{\min}(Q)$. In other words,

$$v(\mathrm{T}_1) = v(\mathrm{P}) = \frac{1}{2}\lambda_{\min}(Q)$$

with no preliminary condition.

4. Conclusion. From the study, we observe that solving the extended trust region subproblem $(\mathrm{T_m})$ by finding the hidden convexity is a very restrictive idea. Even for $m = 1$, there are simple examples with a positive relaxation gap (c.f. Example 3.2), which in turns require a much more complicate SOCP/SDP formulation to catch the hidden convexity. Contrarily, we think the problem more directly from the structure of the polytope. If m is fixed, since in the last resort we can enumerate all the vertices which depends only on m, the complexity for $(\mathrm{T_m})$ should be only a function of m and thus be fixed too. This is indeed the case as we have exhibited in this paper. Our scheme also has to base on the understanding of results in 1990's for characterizing the global minimum and the local non-global minimum of the classical trust region subproblem. Although the results have been cited for quite a number of times in literature, we still feel that they did not receive noticeable attention. The induction technique reduces $(\mathrm{T_m})$ to a couple of small-sized trust region subproblems $(\mathrm{T_0})$. When m is not too large, it can be very efficient as only solving the root of a one-dimensional convex secular function is needed. Our new dimension condition [NewDC] has its own interest too. When the problem has high multiplicity of the smallest eigenvalue or the coefficient vector of the linear inequality constraints are in the space spanned by the eigenvectors corresponding to non-minimal eigenvalues of the Hessian matrix of the objective function, $(\mathrm{T_m})$ can be solved directly by its SDP reformulation.

REFERENCES

[1] D. P. Bertsekas, A. Nedić, A. E. Ozdaglar, Convex Analysis and Optimization, Athena Scientific, Belmont, MA, 2003

[2] S. Burer and K. M. Anstreicher, Second-Order-Cone Constraints for Extended Trust-Region Subproblems, SIAM J. Optim., 23(1): 432 - 451, 2013

[3] S. Burer and B. Yang, The Trust-Region Subproblem with Non-Intersecting Linear Constraints, working paper

[4] M. R. Celis, J. E. Dennis, and R. A. Tapia. A trust region strategy for nonlinear equality constrained optimization. In Numerical Optimization, 1984 (Boulder, Colo., 1984), pages 71–82. SIAM, Philadelphia, PA, 1985.

[5] A. R. Conn, N. I. M. Gould, and P. L. Toint. Trust-Region Methods. MPS/SIAM Series on Optimization. SIAM, Philadelphia, PA, 2000

[6] M. Fu, Z.-Q. Luo, and Y. Ye. Approximation algorithms for quadratic programming. J. Combinatorial Optimization, 2:29–50, 1998.

[7] N. I. M. Gould, S. Lucidi, M. Roma, and P. L. Toint. Solving the trust-region subproblem using the Lanczos method. SIAM J. Optim., 9(2):504–525, 1999

[8] V. Jeyakumar, G. Y. Li, Trust-Region Problems with Linear Inequality Constraints: Exact SDP Relaxation, Global Optimality and Robust Optimization, Math. Programming, to appear

[9] S. Lucidi, L. Palagi and M. Roma, On some properties of quadratic programs with a convex quadratic constraint, SIAM J. Optim. 8, 105–122, 1998

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

[11] J. J. Moré and D. C. Sorensen. Computing a trust region step. SIAM J. Sci. Statist. Comput., 4(3):553–572, 1983.

[12] I. Pólik and T. Terlaky, A Servey of S-lemma. SIAM review. 49(3), 371-418 (2007)

[13] F. Rendl and H. Wolkowicz. A semidefinite framework for trust region subproblems with applications to large scale minimization. Math. Programming, 77(2, Ser. B):273–299, 1997

[14] N. Z. Shor, Quadratic optimization functions problems. *Soviet Journal of Computer and System Sciences*, 6, 137–161, 1987.

[15] R. J. Stern and H. Wolkowicz, Trust region problems and nonsymmetric eigenvalue perturba-

tions, SIAM J. Matrix Anal. Appl. 15, 755–778, 1994

[16] J. F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. Math. Oper. Res., 28(2):246–267, 2003

[17] L. Vandenberghe and S. Boyd, Semidefinite programming. SIAM Review 38, 49-95, 1996

[18] Y. Ye. A new complexity result on minimization of a quadratic function with a sphere constraint. In C. Floudas and P. Pardalos, editors, Recent Advances in Global Optimization. Princeton University Press, Princeton, NJ, 1992.

[19] Y. Ye and S. Zhang. New results on quadratic minimization. SIAM J. Optim., 14(1):245–267, 2003

[20] Y. Yuan. On a subproblem of trust region algorithms for constrained optimization. Math. Programming, 47(1), (Ser. A):53–63, 1990