

---

# Trust-region and $p$ -regularized subproblems: local nonglobal minimum is the second smallest objective function value among all first-order stationary points

Jiulin Wang · Mengmeng Song · Yong Xia

Received: date / Accepted: date

**Abstract** The local nonglobal minimizer of trust-region subproblem, if it exists, is shown to have the second smallest objective function value among all KKT points. This new property is extended to  $p$ -regularized subproblem. As a corollary, we show for the first time that finding the local nonglobal minimizer of Nesterov-Polyak subproblem corresponds to a generalized eigenvalue problem.

**Keywords** Trust-region subproblem ·  $p$ -regularized subproblem · Nesterov-Polyak subproblem · Local nonglobal minimizer · Monotonicity property

**Mathematics Subject Classification (2000)** 90C20, 90C26, 90C30

## 1 Introduction

The trust-region subproblem (TRS) is to minimize a (possibly nonconvex) quadratic function over a Euclidean ball. It is an essential problem in each iteration of the trust region method for solving nonlinear programming problems [1]. (TRS) has other applications in convex quadratic integer programming [2], bounded linear regression, best rank-1 tensor approximation, and image deconvolution [3].

Nonconvex (TRS) has the property of zero Lagrangian-duality gap, see the recent survey on hidden convexity [4]. There is a necessary and sufficient optimality condition for the global minimizer, established in the early 1980s [5,

---

This research was supported by the Beijing Natural Science Foundation, grant Z180005 and by the National Natural Science Foundation of China under grants 11822103, 11571029, and 11771056.

---

Jiulin Wang · Mengmeng Song · Yong Xia  
LMIB of the Ministry of Education, School of Mathematical Sciences, Beihang University, Beijing, 100191, P. R. China  
E-mail: (wangjiulin@buaa.edu.cn (J. L. Wang); songmengmeng@buaa.edu.cn (M. M. Song); yxia@buaa.edu.cn (Y. Xia, corresponding author) )

6, 7]. In 1994, Martínez [8] proved that (TRS) has at most one local nonglobal minimizer. In 2020, Wang and Xia [9] established the necessary and sufficient optimality condition for the local nonglobal minimizer.

Besides the second-order optimality conditions for local and global minimizers, first-order optimality (KKT) conditions are deeply investigated for (TRS), see for example [10]. Gander [11] showed that the objective function values of KKT points of linear least squares problem with a quadratic equality constraint are monotonic in terms of their Lagrangian multipliers. In order to locate local solutions of the Celis-Dennis-Tapia (CDT) subproblem [12] which minimizes a quadratic function over the intersection of two ellipsoids, Chen and Yuan [13] studied the monotonicity property for KKT points of (CDT) in a partial order relation of corresponding Lagrangian multipliers.

The  $p$ -regularized subproblem ( $p$ -RS) is an unconstrained optimization problem of minimizing a (possibly nonconvex) quadratic function with an additional  $p$ -th power regularization term of the norm of the variables [14, 15]. In literature, if the regularization term is cubic, ( $p$ -RS) is known as Nesterov-Polyak subproblem [16]. For more references, we refer to [17, 18, 19, 20] and references therein. If the regularization term is quartic, ( $p$ -RS) reduces to the double-well potential optimization, which has particular applications in solid mechanics and quantum mechanics [21, 22].

In this paper, based on the monotonicity properties for first-order stationary points, we prove that the local nonglobal minimizer of (TRS), if it exists, has the second smallest objective function value among all KKT points. We then extend the new property to ( $p$ -RS) and show that the local nonglobal minimizer of (TRS) has the smallest objective function value among all critical points. As a corollary, we show for the first time that the local nonglobal minimizer of the Nesterov-Polyak subproblem could be founded by solving a generalized eigenvalue problem.

The rest of this paper is organized as follows. In Section 2, we review some properties for local minimizers of (TRS) and ( $p$ -RS). In Section 3 and 4, we first show that monotonicity properties hold for first-order stationary points of (TRS) and ( $p$ -RS), respectively. As the main result of this paper, we prove that the local nonglobal minimums of (TRS) and ( $p$ -RS) are the second smallest objective function value among all first-order stationary points, respectively. Conclusion and discussions are made in Section 5.

## 2 Preliminaries

In this section, we review some characterizations for local minimizers of (TRS) and ( $p$ -RS).

## 2.1 Characterizations for local minimums of (TRS)

In this subsection, we briefly review some properties of local minimums of nonconvex (TRS):

$$\begin{aligned} \text{(TRS)} \quad \min f(x) &= \frac{1}{2}x^T Qx + c^T x \\ \text{s.t. } x^T x - 1 &\leq 0, \end{aligned}$$

where  $Q \in \mathbb{R}^{n \times n}$  is not positive semidefinite and  $c \in \mathbb{R}^n$ . Up to an eigenvalue decomposition, we can always assume that  $Q$  is a diagonal matrix with  $n$  diagonal elements being  $(0 >) \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$ .

$(x; \lambda) \in \mathbb{R}^n \times \mathbb{R}$  is a KKT point of (TRS) if and only if it satisfies the following KKT system:

$$Qx + c = -\lambda x, \quad (2.1)$$

$$\lambda(x^T x - 1) = 0, \quad (2.2)$$

$$x^T x - 1 \leq 0, \quad \lambda \geq 0. \quad (2.3)$$

In case of  $\lambda > 0$ , (2.2)-(2.3) is equivalent to

$$x^T x - 1 = 0. \quad (2.4)$$

Suppose  $-\lambda \notin \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ . Then by substituting the  $x$ -solution of (2.1) into (2.4), we obtain that  $\lambda$  must be a zero point of the following so-called *secular function*:

$$\varphi(\lambda) = \sum_{i=1}^n \frac{c_i^2}{(\alpha_i + \lambda)^2} - 1. \quad (2.5)$$

The first and second derivatives of  $\varphi(\lambda)$  can be written as follows:

$$\varphi'(\lambda) = -2 \sum_{i=1}^n \frac{c_i^2}{(\alpha_i + \lambda)^3},$$

$$\varphi''(\lambda) = 6 \sum_{i=1}^n \frac{c_i^2}{(\alpha_i + \lambda)^4}.$$

If  $\varphi(\lambda)$  has zero points, then  $c \neq 0$  and hence  $\varphi''(\lambda) > 0$  on each nonempty interval  $(-\alpha_{i+1}, -\alpha_i)$  for  $i \in \{1, 2, \dots, n-1\}$ . It follows that, on each interval,  $\varphi(\lambda)$  is strongly convex and thus has at most two zero points.

In the early 1980s, necessary and sufficient conditions for global minimums of (TRS) are established by Gay [5], Sorensen [6], Moré and Sorensen [7].

**Lemma 2.1** ([5], **Theorem 2.1**)  *$x^*$  is a global minimum of (TRS) if and only if  $x^{*T} x^* = 1$  and there is a  $\lambda^* \in \mathbb{R}$  such that (2.1) holds and*

$$\lambda^* \geq -\alpha_1 (> 0).$$

In 1994, Martínez [8] identified two cases where the local nonglobal minimum of (TRS) does not exist.

**Lemma 2.2** ([8], **Lemmas 3.2, 3.3**) *Suppose either  $\alpha_1 = \alpha_2$  or  $c_1 = 0$ , there is no local nonglobal minimizer of (TRS).*

In the same paper, Martínez [8] proved that (TRS) has at most one local nonglobal minimizer based on the following detailed characterization.

**Lemma 2.3** ([8], **Theorem 3.1(i)**) *If  $\underline{x}$  is a local nonglobal minimizer of (TRS), then there is a nonnegative  $\underline{\lambda} \in (-\alpha_2, -\alpha_1)$  such that (2.1) holds, and  $\underline{\lambda}$  is a zero point of  $\varphi(\lambda)$  and*

$$\varphi'(\underline{\lambda}) \geq 0.$$

Recently, Lemma 2.3 has been updated by Wang and Xia [9].

**Lemma 2.4** ([9], **Theorem 3.1**)  *$\underline{x}$  is a local nonglobal minimizer of (TRS), if and only if there is a nonnegative  $\underline{\lambda} \in (-\alpha_2, -\alpha_1)$  such that (2.1) holds, and  $\underline{\lambda}$  is a zero point of  $\varphi(\lambda)$  and*

$$\varphi'(\underline{\lambda}) > 0.$$

In 1998, Lucidi et al. [10] proved that the strict complementarity condition holds at the local nonglobal minimizer of (TRS).

**Lemma 2.5** ([10], **Proposition 3.5**) *If  $\underline{x}$  is a local nonglobal minimizer of (TRS), the corresponding Lagrangian multiplier is  $\underline{\lambda}$ , then (2.2) holds with*

$$\underline{x}^T \underline{x} - 1 = 0, \quad \underline{\lambda} > 0.$$

In 2017, Adachi et al. [23] observed the following result.

**Lemma 2.6** ([23], **Lemma 3.1**) *If  $(x; \lambda)$  is a KKT point of (TRS) satisfying (2.1) and (2.4), then we have*

$$\det M_1(\lambda) = 0,$$

where

$$M_1(\lambda) := \begin{bmatrix} -I & Q \\ Q & -cc^T \end{bmatrix} - \lambda \begin{bmatrix} 0 & -I \\ -I & 0 \end{bmatrix}.$$

That is,  $\lambda$  is a real generalized eigenvalue of  $M_1(\lambda)$ .

Let  $(x; \lambda)$  satisfying (2.1). We can rewrite  $M_1(\lambda)$  as

$$M_1(\lambda) = \begin{bmatrix} I & \\ & Q + \lambda I \end{bmatrix} \begin{bmatrix} -I & I \\ I & -xx^T \end{bmatrix} \begin{bmatrix} I & \\ & Q + \lambda I \end{bmatrix},$$

Then,

$$\det M_1(\lambda) = (-1)^n (\det(Q + \lambda I))^2 (1 - x^T x). \quad (2.6)$$

The following result follows from Lemma 2.6 and (2.6).

**Corollary 2.1** ([24], **Corollary 3**) *The set of real generalized eigenvalues of  $M_1(\lambda)$  is nonempty. Moreover, if  $\det M_1(\lambda) = 0$ , then either  $-\lambda$  is an eigenvalue of  $Q$  or  $\lambda$  is a zero point of  $\varphi(\lambda)$ .*

2.2 Characterizations for local minimizers of ( $p$ -RS)

In this subsection, we review some characterizations for local minimizers of ( $p$ -RS):

$$(p\text{-RS}) \quad \min_{x \in \mathbb{R}^n} \left\{ g(x) = \frac{1}{2}x^T Qx + c^T x + \frac{\sigma}{p}\|x\|^p \right\},$$

where  $\sigma > 0$ ,  $p > 2$ ,  $\|\cdot\|$  is the Euclidean norm, and  $Q, c$  have the same definitions as in (TRS). The global minimum of ( $p$ -RS) is attainable, even if  $Q$  is not positive semidefinite. Actually,  $g(x)$  is coercive, i.e.,

$$\lim_{\|x\| \rightarrow +\infty} g(x) = +\infty,$$

since  $p > 2$  and then

$$\lim_{\|x\| \rightarrow +\infty} \frac{\frac{1}{2}x^T Qx + c^T x}{\|x\|^p} = 0.$$

We call  $x \in \mathbb{R}^n$  a critical point of ( $p$ -RS) if and only if it satisfies the following first order necessary condition for local minimizers of ( $p$ -RS):

$$\nabla g(x) = (Q + \sigma\|x\|^{p-2}I)x + c = 0, \quad (2.7)$$

where  $\nabla g$  denotes the gradient of  $g(\cdot)$ . Let

$$t = \|x\|^{p-2}, \quad (2.8)$$

then (2.7) can be rewritten as

$$(Q + \sigma t I)x + c = 0. \quad (2.9)$$

Suppose  $-\sigma t \notin \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ . Then by substituting the  $x$ -solution of (2.9) into (2.8), we obtain that  $t$  must be a zero point of the following secular function:

$$h(t) = \sum_{i=1}^n \frac{c_i^2}{(\sigma t + \alpha_i)^2} - t^{\frac{2}{p-2}}. \quad (2.10)$$

In the following, we make an assumption that  $x \neq 0$ , then it holds that  $c \neq 0$  from (2.7). Under this assumption,  $h(t)$  has the same zeros as

$$p(t) = \log \left( \sum_{i=1}^n \frac{c_i^2}{(\sigma t + \alpha_i)^2} \right) - \frac{2}{p-2} \log(t).$$

Hsia et al. [15] proved that  $p''(t) > 0$  on each nonempty interval  $(-\frac{\alpha_{i+1}}{\sigma}, -\frac{\alpha_i}{\sigma})$  for  $i \in \{1, 2, \dots, n-1\}$ .

The first derivatives of  $h(t)$  and  $p(t)$  can be written as follows:

$$h'(t) = -\sum_{i=1}^n \frac{2\sigma c_i^2}{(\sigma t + \alpha_i)^3} - \frac{2}{(p-2)t} \cdot t^{\frac{2}{p-2}}, \quad (2.11)$$

$$p'(t) = -\frac{\sum_{i=1}^n \frac{2\sigma c_i^2}{(\sigma t + \alpha_i)^3}}{\sum_{i=1}^n \frac{c_i^2}{(\sigma t + \alpha_i)^2}} - \frac{2}{(p-2)t}. \quad (2.12)$$

If  $\hat{t}$  is a zero point of  $h(t)$ , according to (2.10), (2.11) and (2.12), we obtain

$$p'(\hat{t}) = h'(\hat{t}) \cdot \hat{t}^{\frac{-2}{p-2}}. \quad (2.13)$$

The following detailed characterizations of local minimizers of ( $p$ -RS) are all due to Hsia et al. [15].

**Lemma 2.7** ([15], **Theorem 2.2**)  *$x^*$  is a global minimizer of ( $p$ -RS) for  $p > 2$  if and only if it is a critical point satisfying (2.7) and*

$$\sigma \|x^*\|^{p-2} + \alpha_1 \geq 0.$$

Moreover, the  $\ell_2$  norms of all the global minimizers are equal.

**Lemma 2.8** ([15], **Lemma 3.1**) *Suppose  $\underline{x}$  is a local non-global minimizer of ( $p$ -RS) for  $p > 2$ . It holds that  $\underline{x}_1 \neq 0$ ,  $\alpha_1 < \alpha_2$ , and*

$$\sigma \|\underline{x}\|^{p-2} + \alpha_2 > 0.$$

**Lemma 2.9** ([15], **Theorem 3.2**) *The point  $\underline{x}$  is a local nonglobal minimizer of ( $p$ -RS) for  $p > 2$  if and only if  $\underline{x}$  is a critical point satisfying (2.7) and  $\underline{t}$  ( $=\|\underline{x}\|^{p-2}$ ) is a root of*

$$h(t) = 0, \quad t \in \left( \max \left\{ -\frac{\alpha_2}{\sigma}, 0 \right\}, -\frac{\alpha_1}{\sigma} \right)$$

such that  $h'(\underline{t}) > 0$ .

Motivated by the observation that the Lagrange multiplier of each KKT point of (TRS) satisfying (2.1) and (2.4) is a generalized eigenvalue of  $M_1(\lambda)$  [23], Lieder [20] pointed out the following result holds for the Nesterov-Polyak subproblem (i.e., ( $p$ -RS) for  $p = 3$ ).

**Lemma 2.10** ([20], **Lemma 3.1**) *If  $x$  is a critical point of the cubic regularization, then  $t$  ( $=\|x\|$ ) is a real generalized eigenvalue of*

$$M_2(t) = \begin{bmatrix} 0 & 0 & 0 & c^T \\ 0 & -\sigma I & 0 & Q \\ 0 & 0 & -\sigma & 0 \\ c & Q & 0 & 0 \end{bmatrix} - t \begin{bmatrix} 0 & 0 & -\sigma & 0 \\ 0 & 0 & 0 & -\sigma I \\ -\sigma & 0 & 0 & 0 \\ 0 & -\sigma I & 0 & 0 \end{bmatrix}.$$

That is,  $t$  is a real root of  $\det M_2(t) = 0$ .

Let  $(x; t)$  satisfying (2.9). We can reformulate  $M_2(t)$  as

$$M_2(t) = \begin{bmatrix} I & & & & \\ & I & & & \\ & & 1 & & \\ & & & Q + \sigma t I & \\ & & & & \end{bmatrix} \begin{bmatrix} 0 & 0 & \sigma t & -x^T \\ 0 & -\sigma I & 0 & I \\ \sigma t & 0 & -\sigma & 0 \\ -x & I & 0 & 0 \end{bmatrix} \begin{bmatrix} I & & & & \\ & I & & & \\ & & 1 & & \\ & & & Q + \sigma t I & \\ & & & & \end{bmatrix},$$

Then, by Schur complement,

$$\det M_2(t) = (-1)^{n+1} \sigma^2 \cdot (\det(Q + \sigma t I))^2 (t^2 - x^T x). \quad (2.14)$$

The following result follows from Lemma 2.10 and (2.14).

**Corollary 2.2** *The set of real generalized eigenvalues of  $M_2(t)$  is nonempty. Moreover, if  $\det M_2(t) = 0$ , then one of the following cases happens: (1)  $t = 0$ ; (2)  $-\sigma t$  is an eigenvalue of  $Q$ ; (3)  $t$  is a zero point of  $p(t)$ .*

### 3 On local nonglobal minimum of (TRS)

We first prove a monotonicity property for KKT points of (TRS).

**Lemma 3.1** *Let  $(x_1; \lambda_1)$  and  $(x_2; \lambda_2)$  be two KKT points for (TRS). Then*

$$\lambda_2 \geq (>) \lambda_1 \implies f(x_2) \leq (<) f(x_1).$$

*Proof* According to (2.1) and the definitions of  $(x_1; \lambda_1)$  and  $(x_2; \lambda_2)$ , we have

$$Qx_1 + c = -\lambda_1 x_1, \quad Qx_2 + c = -\lambda_2 x_2. \quad (3.1)$$

By multiplying  $x_1^T$  and  $x_2^T$  to both sides of the two equalities in (3.1), respectively, we have

$$\begin{aligned} x_1^T Qx_1 + x_1^T c &= -\lambda_1 x_1^T x_1, & x_1^T Qx_2 + x_1^T c &= -\lambda_2 x_1^T x_2, \\ x_2^T Qx_2 + x_2^T c &= -\lambda_2 x_2^T x_2, & x_2^T Qx_1 + x_2^T c &= -\lambda_1 x_2^T x_1. \end{aligned}$$

Then, it holds that

$$\begin{aligned} 2(f(x_2) - f(x_1)) &= (x_2^T Qx_2 + x_2^T c) - (x_1^T Qx_1 + x_1^T c) + x_2^T c - x_1^T c \\ &= (-\lambda_2 x_2^T x_2) - (-\lambda_1 x_1^T x_1) + (-\lambda_1 x_2^T x_1) - (-\lambda_2 x_1^T x_2) \\ &= (\lambda_2 - \lambda_1) x_1^T x_2 + \lambda_1 x_1^T x_1 - \lambda_2 x_2^T x_2 \end{aligned} \quad (3.2)$$

$$\leq (\lambda_2 - \lambda_1) \cdot 1 + \lambda_1 x_1^T x_1 - \lambda_2 x_2^T x_2 \quad (3.3)$$

$$= \lambda_1 (x_1^T x_1 - 1) - \lambda_2 (x_2^T x_2 - 1) = 0, \quad (3.4)$$

where (3.3) follows from Cauchy-Schwarz inequality and (2.3), more precisely,

$$x_1^T x_2 \leq \|x_1\| \cdot \|x_2\| \leq 1, \quad (3.5)$$

the last equality in (3.4) holds due to the complementarity condition (2.2).

The equality  $f(x_2) = f(x_1)$  holds true if and only if the inequality in (3.3) holds as an equality, that is

$$(\lambda_2 - \lambda_1)(x_1^T x_2 - 1) = 0. \quad (3.6)$$

In case of  $\lambda_2 > \lambda_1$ , (3.6) is equivalent to

$$x_1^T x_2 = 1. \quad (3.7)$$

It implies from (3.5) and (3.7) that

$$x_1 = x_2 \neq 0. \quad (3.8)$$

Combining the assumption  $\lambda_2 > \lambda_1$  with (2.1) and (3.8) yields a contradiction. Therefore,  $f(x_2) = f(x_1)$  if and only if  $\lambda_2 = \lambda_1$ . The proof is complete.  $\square$

**Remark 3.1** *Let  $(x_1; \lambda_1)$  and  $(x_2; \lambda_2)$  be two KKT points for (TRS) satisfying  $\|x_1\| = \|x_2\|$ , since*

$$x_1^T x_1 - x_1^T x_2 = x_2^T x_2 - x_1^T x_2 = \frac{1}{2} \|x_1 - x_2\|^2,$$

*it follows from (3.2) that*

$$f(x_2) - f(x_1) = \frac{\lambda_1 - \lambda_2}{4} \|x_1 - x_2\|^2,$$

*which is presented in [11] for the relaxed least square problem.*

As a corollary of Lemma 3.1, the following result holds.

**Corollary 3.1** *If  $(x^*; \lambda^*)$  is KKT point of (TRS),  $x^*$  is a global minimizer, then  $\lambda^*$  is the largest Lagrangian multiplier among all KKT points. Moreover, all the global minimizers of (TRS) share the same Lagrangian multiplier  $\lambda^*$ .*

Besides, according to Lemma 3.1, Lemma 2.1, Lemma 2.6 and Corollary 2.1, we can recover the following result for the global minimizer of (TRS).

**Corollary 3.2 ([23], Theorem 3.2)** *If  $x^*$  is a global minimizer of (TRS),  $\lambda^*$  is the corresponding Lagrangian multiplier, then  $\lambda^*$  is equal to the largest real generalized eigenvalue of  $M_1(\lambda)$ .*

We now present the main result of this section.

**Theorem 3.1** *Let  $(x_1; \lambda_1)$  and  $(x_2; \lambda_2)$  be any two different KKT points for (TRS). If  $x_1$  is a local nonglobal minimizer of (TRS) and  $x_2$  is not a global minimizer of (TRS), then  $f(x_1) < f(x_2)$ .*



*Proof* According to Lemma 3.1, it is sufficient to prove  $\lambda_2 < \lambda_1$ .

If  $x_2^T x_2 < 1$ , then (2.2) implies that  $\lambda_2 = 0$ . According to Lemma 2.5, it holds that  $\lambda_1 > 0$ . Then, we obtain

$$\lambda_2 = 0 < \lambda_1.$$

If  $x_2^T x_2 = 1$ , then either  $-\lambda_2 \in \{\alpha_1, \alpha_2, \dots, \alpha_n\}$  or  $\lambda_2$  is a zero point of  $\varphi(\lambda)$  (2.5). Since  $x_2$  is not a global minimizer of (TRS), it follows from Lemma 2.1 that

$$\lambda_2 < -\alpha_1.$$

According to Lemma 2.3,  $\lambda_1$  is the largest zero point of  $\varphi(\lambda)$  in  $(-\alpha_2, -\alpha_1)$ . Then we have

$$\lambda_2 < \lambda_1 < -\alpha_1.$$

The proof is complete.  $\square$

According to Lemma 2.2, if there is a local nonglobal minimizer of (TRS), then we have  $\alpha_1 \neq \alpha_2$  and  $c_1 \neq 0$ . Then,  $\det M_1(-\alpha_1) \neq 0$ . Otherwise, it is not difficult to verify that  $c_1 = 0$ , which leads to a contradiction. Combining  $\det M_1(-\alpha_1) \neq 0$ , the proof of Theorem 3.1, Lemma 2.1, Lemma 2.6 and Corollary 2.1 together, we can recover the following result.

**Corollary 3.3 ([24], Theorem 5)** *If  $\underline{x}$  is a local nonglobal minimizer of (TRS),  $\underline{\lambda}$  is the corresponding Lagrangian multiplier, then  $\underline{\lambda}$  is equal to the second largest real generalized eigenvalue of  $M_1(\lambda)$ .*

#### 4 On local nonglobal minimum of (p-RS)

We first prove a monotonicity property for critical points of (p-RS).

**Lemma 4.1** *Let  $x_1$  and  $x_2$  be two critical points of (p-RS) for  $p > 2$ . Then*

$$\|x_2\| \geq (>)\|x_1\| \implies g(x_2) \leq (<)g(x_1).$$

*Proof* According to (2.7) and the definitions of  $x_1$  and  $x_2$ , we have

$$(Q + \sigma\|x_1\|^{p-2}I)x_1 + c = 0, \quad (Q + \sigma\|x_2\|^{p-2}I)x_2 + c = 0. \quad (4.1)$$

By multiplying  $x_1^T$  and  $x_2^T$  to both sides of the two equalities in (4.1), respectively, we have

$$\begin{aligned} x_1^T Q x_1 + x_1^T c + \sigma\|x_1\|^p &= 0, & x_1^T Q x_2 + x_1^T c + \sigma\|x_2\|^{p-2}x_1^T x_2 &= 0, \\ x_2^T Q x_2 + x_2^T c + \sigma\|x_2\|^p &= 0, & x_2^T Q x_1 + x_2^T c + \sigma\|x_1\|^{p-2}x_2^T x_1 &= 0. \end{aligned}$$

Then,  $g(x_2) - g(x_1)$  is equal to

$$\begin{aligned} & \frac{1}{2}((x_2^T Q x_2 + x_2^T c) - (x_1^T Q x_1 + x_1^T c)) + \frac{1}{2}(x_2^T c - x_1^T c) + \frac{\sigma}{p}(\|x_2\|^p - \|x_1\|^p) \\ &= \left(\frac{\sigma}{p} - \frac{\sigma}{2}\right)(\|x_2\|^p - \|x_1\|^p) + \frac{\sigma}{2}(\|x_2\|^{p-2} - \|x_1\|^{p-2})x_1^T x_2. \end{aligned} \quad (4.2)$$

Then by (4.2), we have

$$\|x_2\| = \|x_1\| \implies g(x_2) = g(x_1). \quad (4.3)$$

Next, we show that

$$\|x_2\| > \|x_1\| \implies g(x_2) < g(x_1). \quad (4.4)$$

If  $x_1 = 0$  and  $x_2 \neq 0$ , it follows from (4.2) and  $p > 2$  that

$$g(x_2) - g(x_1) = \left(\frac{\sigma}{p} - \frac{\sigma}{2}\right)\|x_2\|^p < 0. \quad (4.5)$$

If  $x_1 \neq 0$ , let  $\|x_2\| = \alpha\|x_1\|$ , where  $\alpha \geq 1$ . Substituting  $\|x_2\| = \alpha\|x_1\|$  to (4.2) yields that

$$\begin{aligned} g(x_2) - g(x_1) &= \left(\frac{\sigma}{p} - \frac{\sigma}{2}\right)(\alpha^p - 1)\|x_1\|^p + \frac{\sigma}{2}(\alpha^{p-2} - 1)\|x_1\|^{p-2} \cdot x_1^T x_2 \\ &\leq \frac{\sigma}{2}\|x_1\|^p \cdot \left(\frac{2-p}{p}(\alpha^p - 1) + \alpha^{p-1} - \alpha\right). \end{aligned} \quad (4.6)$$

where (4.6) follows from Cauchy-Schwarz inequality, i.e.,

$$x_1^T x_2 \leq \|x_1\| \cdot \|x_2\| = \alpha\|x_1\|^2.$$

Define

$$\phi(\alpha) = \frac{2-p}{p}(\alpha^p - 1) + \alpha^{p-1} - \alpha.$$

The first and second derivatives of  $\phi(\alpha)$  can be written as follows:

$$\begin{aligned} \phi'(\alpha) &= (p-1)\alpha^{p-2} - (p-2)\alpha^{p-1} - 1, \\ \phi''(\alpha) &= (p-1)(p-2)(1-\alpha)\alpha^{p-3}. \end{aligned} \quad (4.7)$$

If  $\alpha = 1$ , we can verify that

$$\phi(1) = \phi'(1) = \phi''(1) = 0. \quad (4.8)$$

If  $\alpha > 1$ , according to (4.7) and (4.8), we have

$$\phi''(\alpha) < \phi''(1) = 0 \implies \phi'(\alpha) < \phi'(1) = 0 \implies \phi(\alpha) < \phi(1) = 0. \quad (4.9)$$

Then (4.4) holds due to (4.5) and (4.9). Combining (4.3) with (4.4), the proof is complete.  $\square$

If  $p = 3$ , according to Lemma 4.1, Lemma 2.7, Lemma 2.10 and Corollary 2.2, we recover the following result.

**Corollary 4.1 ([20], Corollary 3.2)** *If  $x^*$  is a global minimizer of (p-RS) for  $p = 3$ , then  $t^* = \|x^*\|$  is equal to the largest real generalized eigenvalue of  $M_2(t)$ .*

Now we present the main result of this section.

**Theorem 4.1** *Let  $x_1$  and  $x_2$  be any two different critical points of (p-RS) for  $p > 2$ . If  $x_1$  is a local nonglobal minimizer of (p-RS) and  $x_2$  is not a global minimizer, then  $g(x_1) < g(x_2)$ .*

*Proof* Let  $t_i = \|x_i\|^{p-2}$  for  $i = 1, 2$ . According to Lemma 4.1, it is sufficient to prove  $\|x_2\| < \|x_1\|$  or equivalently,  $t_2 < t_1$ .

We first prove that  $c \neq 0$ . As  $x_1$  is a local nonglobal minimizer of (p-RS), according to Lemma 2.7, we have

$$\sigma\|x_1\|^{p-2} + \alpha_1 < 0. \quad (4.10)$$

Then, it follows from (4.10), (2.7) and Lemma 2.8 that  $c_1 \neq 0$ , and hence  $c \neq 0$ .

Next, as  $c \neq 0$ , (2.7) implies that  $x_1, x_2 \neq 0$ . Then,  $p(t)$  is well defined. We can conclude that either  $-\sigma t_2 \in \{\alpha_1, \alpha_2, \dots, \alpha_n\}$  or  $t_2$  is a zero point of  $p(t)$ . Since  $x_2$  is not a global minimizer of (p-RS), it follows from Lemma 2.7 that

$$t_2 < -\frac{\alpha_1}{\sigma}.$$

Since  $p(t)$  is strongly convex for  $t \in (\max\{-\frac{\alpha_2}{\sigma}, 0\}, -\frac{\alpha_1}{\sigma})$ , it has at most two zero points in this interval. According to Lemma 2.9 and (2.13), we have

$$p'(t_1) > 0.$$

Thus  $t_1$  is the largest zero point of  $p(t)$  in the interval  $(\max\{-\frac{\alpha_2}{\sigma}, 0\}, -\frac{\alpha_1}{\sigma})$ . Then we have

$$t_2 < t_1 < -\frac{\alpha_1}{\sigma}.$$

The proof is complete.  $\square$

In Section 3, we have shown that if there is a local nonglobal minimizer of (TRS), then  $\det M_1(-\alpha_1) \neq 0$ . Similarly, we can show that if there is a local nonglobal minimizer of (p-RS) for  $p = 3$ , then  $\det M_2(-\frac{\alpha_1}{\sigma}) \neq 0$ . Combing this result with the proof of Theorem 4.1, Lemmas 2.7, 2.10 and Corollary 2.2, we give the following result.

**Corollary 4.2** *If  $\underline{x}$  is the local nonglobal minimizer of the Nesterov-Polyak subproblem (i.e., (p-RS) for  $p = 3$ ), then  $\underline{t} = \|\underline{x}\|$  is equal to the second largest real generalized eigenvalue of  $M_2(t)$ .*

## 5 Conclusion and discussions

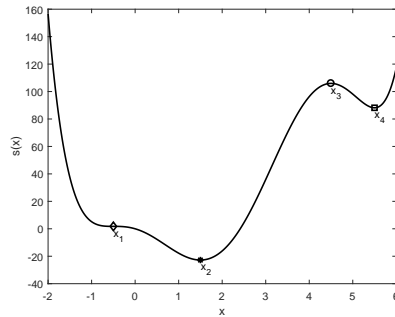
This paper shows that there are monotonicity properties for first-order stationary points of trust-region subproblem (TRS) and  $p$ -regularized subproblem (p-RS). Based on these properties, we point out that all the global minimizers of (TRS) share the same Lagrangian multiplier and recover the result that finding a global minimizer of (TRS) (or (p-RS)) corresponds to a generalized eigenvalue problem. As the main contribution, we prove that the local

nonglobal minimizer of (TRS) (or ( $p$ -RS)), if it exists, has the second smallest objective function value among all KKT points (critical points). For the Nesterov-Polyak subproblem (i.e., ( $p$ -RS) with  $p = 3$ ), we show for the first time that the local nonglobal minimizer, if it exists, could also be founded by solving a generalized eigenvalue problem.

Our main results for ( $p$ -RS) or (TRS) may fail to hold for more general optimization problems. Consider the problem of minimizing a univariate sextic function:

$$s(x) = \frac{1}{6}x^6 - \frac{21}{10}x^5 + \frac{57}{8}x^4 - \frac{1}{12}x^3 - \frac{435}{32}x^2 - \frac{297}{32}x.$$

There are four critical points, denoted by  $x_i$ ,  $i = 1, 2, 3, 4$ . As demonstrated in Fig. 1,  $x_1 = -0.5$  is neither a local minimizer nor a local maximizer,  $x_2 = 1.5$  is a global minimizer,  $x_3 = 4.5$  is a local maximizer, and  $x_4 = 5.5$  is a local nonglobal minimizer. It is observed from Fig. 1 that, the unique local nonglobal minimum is not the second smallest objective function value among all critical points. Moreover, the above unconstrained minimization problem



**Fig. 1** Variation of  $s(x)$  with four critical points satisfying  $s(x_2) < s(x_1) < s(x_4) < s(x_3)$ .

can be reformulated as the following nonconvex quadratic optimization with a quintic constraint:

$$\begin{aligned} \min_{x, y \in \mathbb{R}} \quad & xy \\ \text{s.t.} \quad & y = \frac{1}{6}x^5 - \frac{21}{10}x^4 + \frac{57}{8}x^3 - \frac{1}{12}x^2 - \frac{435}{32}x - \frac{297}{32}. \end{aligned}$$

One can verify that the constrained problem has four KKT points, corresponding the four critical points of  $s(x)$ .

It seems that there is still a little room for possible extension. We conclude this paper with the following conjecture:

**Conjecture 5.1** *For (CDT) subproblem or nonconvex quadratic optimization with a quartic constraint, the smallest local nonglobal minimum has the second smallest objective function value among all KKT points.*

## References

1. Y. YUAN, *Recent advances in trust region algorithms*, Math. Program., 151(1) (2015), pp. 249–281.
2. C. BUCHHEIM, R. HÜBNER, A. SCHÖBEL, *Ellipsoid bounds for convex quadratic integer programming*, SIAM J. Optim., 25(2) (2015), pp. 741–769.
3. A. H. PHAN, M. YAMAGISHI, D. MANDIC, A. CICHOCKI, *Quadratic programming over ellipsoids with applications to constrained linear regression and tensor decomposition*, Neural Comput. Appl., 32(2020), pp. 7097–7120.
4. Y. XIA, *A survey of hidden convex optimization*, J. Oper. Res. Soc. China., 8(1) (2020), pp. 1–28.
5. D. M. GAY, *Computing optimal locally constrained steps*, SIAM J. Sci. Stat. Comput., 2(2) (1981), pp. 186–197.
6. D. C. SORENSEN, *Newton's method with a model trust region modification*, SIAM J. Numer. Anal., 19(2) (1982), pp. 409–426.
7. J. J. MORÉ, D. C. SORENSEN, *Computing a trust region step*, SIAM J. Sci. Statist. Comput., 4(3) (1983), pp. 553–572.
8. J. M. MARTÍNEZ, *Local minimizers of quadratic functions on Euclidean balls and spheres*, SIAM J. Optim., 4(1) (1994), pp. 159–176.
9. J. WANG, Y. XIA, *Closing the gap between necessary and sufficient conditions for local nonglobal minimizer of trust region subproblem*, SIAM J. Optim., 30(3) (2020), pp. 1980–1995.
10. S. LUCIDI, L. PALAGI, M. ROMA, *On some properties of quadratic programs with a convex quadratic constraint*, SIAM J. Optim., 8(1) (1998), pp. 105–122.
11. W. GANDER, *Least squares with a quadratic constraint*, Numer. Math., 36(3) (1980), pp. 291–307.
12. M. R. CELIS, J. E. DENNIS, R. A. TAPIA, *A trust region strategy for nonlinear inequality constrained optimization*, in Numerical Optimization, R.T. Boggs, R.H. Byrd, and R.B. Schnabel, eds. (1984), SIAM, Philadelphia, PA, (1985), pp. 71–82.
13. X. CHEN, Y. YUAN, *On local solutions of the Celis-Dennis-Tapia subproblem*, SIAM J. Optim., 10(2) (2000), pp. 359–383.
14. N. I. M. GOULD, D. P. ROBINSON, H. SUE THORNE, *On solving trust-region and other regularised subproblems in optimization*, Math. Program. Comput., 2(1) (2010), pp. 21–57.
15. Y. HSIA, R. L. SHEU, Y. YUAN, *Theory and application of  $p$ -regularized subproblems for  $p > 2$* , Optim. Method Softw., 32(5) (2017), pp. 1059–1077.
16. Y. NESTEROV, B. T. POLYAK, *Cubic regularization of Newton method and its global performance*, Math. Program., 108(1) (2006), pp. 177–205.
17. M. WEISER, P. DEUFLHARD, B. ERDMANN, *Affine conjugate adaptive Newton methods for nonlinear elastomechanics*, Optim. Methods Softw., 22(3) (2007), pp. 413–431.
18. C. CARTIS, N. I. M. GOULD, PH. L. TOINT, *Adaptive cubic regularisation methods for unconstrained optimization. Part I: motivation, convergence and numerical results*, Math. Program., 127(2) (2011), pp. 245–295.
19. C. CARTIS, N. I. M. GOULD, PH. L. TOINT, *Adaptive cubic regularisation methods for unconstrained optimization. Part II: worst-case function-and derivative-evaluation complexity*, Math. Program., 130(2) (2011), pp. 295–319.
20. F. LIEDER, *Solving large-Scale cubic regularization by a generalized eigenvalue problem*, SIAM J. Optim., 30(4) 2020, pp. 3345–3358.
21. S. C. FANG, D. GAO, G. X. LIN, R. L. SHEU, W. XING, , *Double well potential function and its optimization in the  $n$ -dimensional real space: part I*, J. Ind. Manag. Optim., 13(3) (2017), pp. 1291–1305.
22. Y. XIA, R. L. SHEU, S. C. FANG, W. XING, *Double well potential function and its optimization in the  $n$ -dimensional real space: part II*, J. Ind. Manag. Optim., 13(3) (2017), pp. 1307–1328.
23. S. ADACHI, S. IWATA, Y. NAKATSUKASA, A. TAKEDA, *Solving the trust-region subproblem by a generalized eigenvalue problem*, SIAM J. Optim., 27(1) (2017), pp. 269–291.
24. M. SALAH, A. TAATI, H. WOLKOWICZ, *Local nonglobal minima for solving large-scale extended trust-region subproblems*, Comput. Optim. Appl., 66(2) (2017), pp. 223–244.