

# Hölderian error bounds and Kurdyka-Łojasiewicz inequality for the trust region subproblem

Rujun Jiang\*      Xudong Li†

May 13, 2022

## Abstract

In this paper, we study the local variational geometry of the optimal solution set of the trust region subproblem (TRS), which minimizes a general, possibly nonconvex, quadratic function over the unit ball. Specifically, we demonstrate that a Hölderian error bound holds globally for the TRS with modulus  $1/4$  and the Kurdyka-Łojasiewicz (KL) inequality holds locally for the TRS with a KL exponent  $3/4$  at any optimal solution. We further prove that unless in a special case, the Hölderian error bound modulus, as well as the KL exponent, is  $1/2$ . Finally, as a byproduct, we further apply the obtained KL property to show that projected gradient methods studied elsewhere for solving the TRS achieve a local sublinear or even linear rate of convergence with probability one by choosing a proper initial point.

## 1 Introduction

In this paper, we consider the following trust region subproblem (TRS)

$$\begin{aligned} (\text{P}_0) \quad & \min \quad f(\mathbf{x}) := \mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{b}^T \mathbf{x} \\ & \text{s.t.} \quad g(\mathbf{x}) := \|\mathbf{x}\|^2 - 1 \leq 0, \end{aligned}$$

where  $\mathbf{A}$  is an  $n \times n$  nonzero real symmetric matrix,  $\mathbf{b} \in \mathbb{R}^n$  and  $\|\cdot\|$  denotes the Euclidean  $l_2$  norm. To avoid the trivial case, we make the blanket assumption that  $n \geq 2$ . Let  $S_0$  be the nonempty optimal solution set of  $(\text{P}_0)$  and  $f^*$  be the optimal value.

Problem  $(\text{P}_0)$  first arises as a subproblem in the trust region method for nonlinear optimization [11, 61], and also admits applications in robust optimization [5] and least squares problems [62]. The generalized trust region subproblem, where the constraint is a general quadratic inequality, is also well studied in the literature [49, 54, 33, 32, 31]. When  $\mathbf{A}$  is not positive semidefinite, problem  $(\text{P}_0)$  is a nonconvex problem and may have a local non-global optimal solution [48, 58]. However, problem  $(\text{P}_0)$  enjoys hidden convexity and the strong duality holds due to the celebrated S-lemma [60]. Various methods have been developed in the literature for solving the TRS, e.g., a safeguarded Newton’s method by solving the so-called secular equation [50], generalized Lanczos methods and its recent variants [25, 65, 64] and a parametrized eigenvalue approach based on semidefinite programming and duality theory [55], to name a few. Hazan and Koren [27] proposed the first linear-time algorithm for the TRS to achieve an  $\epsilon$  optimal solution. Wang and Xia [57] and Ho-Nguyen and Kilinc-Karzan [28] presented a linear-time algorithm to solve the TRS by applying Nesterov’s accelerated gradient descent algorithm to a convex reformulation of  $(\text{P}_0)$ , where such a reformulation was first proposed, to the best of our knowledge, by Flippo and Jansen [20]. Very recently, in [4], a family of first-order methods for solving problem  $(\text{P}_0)$ , including the

---

\*School of Data Science, Fudan University, Shanghai, China, rjjiang@fudan.edu.cn.

†School of Data Science, Fudan University, Shanghai, China, lixudong@fudan.edu.cn.

projected and conditional gradient methods, was shown to converge to a global optimal solution under proper initialization. However, the convergence rates for these methods are still largely unknown.

Error bounds [47, 52] and the Kurdyka-Łojasiewicz (KL) inequality [41, 42, 2] are widely used in the literature for analyzing rate of convergence for various optimization algorithms [45, 46, 47, 52, 29, 67, 66, 56, 2, 22, 35, 36, 40, 39], and for understanding the local variational geometry of the solution set of an optimization problem [14, 19, 15, 12]. In this paper, we devote ourself to a thorough understanding of Hölderian error bounds and the KL inequality for the TRS, which, as far as we know, are not available in the literature. Before stating the definitions of the Hölderian error bounds and the KL inequality for the TRS, let us first define  $B$  as the unit norm ball, i.e.,  $B := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq 1\}$  and  $\delta_B$  as the corresponding indicator function  $\delta_B(\mathbf{x}) = \begin{cases} 0, & \text{if } \|\mathbf{x}\| \leq 1, \\ +\infty, & \text{otherwise.} \end{cases}$

**Definition 1.1** (Hölderian error bounds for the TRS). *The TRS is said to satisfy the global error bound condition with modulus  $\rho$  if there exists a constant  $\tau > 0$  such that*

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^\rho, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Notice that Definition 1.1 is sometimes referred to as the Hölderian growth condition. Specifically, if  $\rho = 1/2$ , it is exactly the quadratic growth condition [14]; and if  $\rho = 1$ , the objective  $f(\mathbf{x}) + \delta_B(\mathbf{x})$  is said to have weak sharp minima [9]. We shall emphasize that the residual function used in Definition 1.1, i.e., the right-hand-side of the inequality, corresponds to the difference between the objective function value at any given point and the optimal value of the TRS. This is different from (sub)gradient based error bounds, e.g., Luo-Tseng error bounds, where residual functions associated with certain (sub)gradient information are used in the definitions. The Hölderian error bound is an important concept in optimization. For example, when  $\rho = 1/2$ , the Hölderian error bound (i.e., the quadratic growth condition) is closely related to the metric regularity conditions, calmness of set-valued mappings, the tilt-stability of local minimizers, and positive-definiteness/semidefiniteness properties of the second-order subdifferential (or generalized Hessian) [1, 14, 15, 16, 12]. It has also been used to analyze the convergence rates of various algorithms [17, 13, 18].

**Definition 1.2** (KL inequality for the TRS). *The TRS is said to satisfy the KL inequality at  $\mathbf{x}^* \in S_0$  with an exponent  $\varrho \in [0, 1)$ , if there exist  $\tau > 0$  and  $\epsilon > 0$  such that*

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^\varrho \leq \tau \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \epsilon),$$

where  $N_B(\mathbf{x})$  is the normal cone of  $B$  at  $\mathbf{x}$ ,  $B(\mathbf{x}^*, \epsilon)$  is the  $\epsilon$ -neighborhood of  $\mathbf{x}^*$ , i.e.,  $B(\mathbf{x}^*, \epsilon) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon\}$ , and by convention the distance of any given point to an empty set is defined as  $\infty$ .

We note that Definition 1.2 is inherited from the definition given in [8], except that our definition focuses only on optimal solutions instead of general stationary points. Existing works of error bounds and the KL inequality related to the TRS are Hölderian error bounds for convex quadratic inequality systems [59] (see also Subsection 3.1), convex piecewise quadratic functions [38], a class of convex composite functions [66, 36], a system consisting of a nonconvex quadratic function and a polyhedron [44], and the KL property for spherical constrained quadratic optimization problems [40, 39, 23] and spherical constrained quartic-quadratic optimization problems [63]. However, none of the above results can be directly applied to the TRS due to the existence of the possibly nonconvex and nonhomogeneous quadratic objective function and the unit ball constraint. A key issue in establishing the KL inequality for the TRS is to derive explicitly the exponent  $\varrho$ , since this exponent plays a central role in analyzing the local convergence rate of many methods. Many existing results [43, 2, 7] only showed the existence of the KL inequality for a given function, where the estimated exponent is often too loose and thus does not offer much help in analyzing the rate of convergence of optimization methods for the TRS (see, e.g., the dimension dependent

KL exponent in [34]). One exception is that Li and Pong [36] developed various calculus rules to deduce explicitly the KL exponent for a large class of problems. However, the TRS is not studied in [36].

Recently, [8] studied the relations between Hölderian error bounds and the KL inequality for convex problems. Specifically, they showed that for *convex problems*, error bounds with moderate residual functions are equivalent to the KL inequalities, and that the sum of the Hölderian error bound modulus and the KL exponent equals one. See also [1, 13, 36, 17, 16] for other related discussions and results. We note that certain equivalent relations have also been derived for “not too nonconvex” functions [30, Proposition 6.8] and lower  $\mathcal{C}^1$  functions [37, Theorem 3.2]. While these results focus on more general nonconvex functions, their assumptions are either not satisfied or difficult to verify for nonconvex TRS problems. As far as we know, the relations between Hölderian error bounds and the KL inequality for the special nonconvex TRS still remain largely unknown.

In this paper, by combining the local geometry of the optimal solution set  $S_0$  and the elegant Hölderian error bound results for convex quadratic inequality systems in [59], we are able to obtain a comprehensive characterization of a Hölderian error bound for the TRS. Specifically, it can be shown that the Hölderian error bound holds globally with the modulus  $\rho = 1/4$  for the *TRS-ill case* (to be defined in (30)) and otherwise,  $\rho = 1/2$ . Then, based on the obtained error bound results, we are able to derive the KL inequality for the TRS. We show that for the TRS, the KL inequality always holds locally with the KL exponent  $\varrho = 3/4$  at any optimal solution (if the TRS is convex, the KL inequality in fact holds globally). More precisely, the KL exponent is  $3/4$  if we are dealing with the *TRS-ill case* and  $1/2$  otherwise at any optimal solution, i.e., the sum of the KL exponent  $\varrho$  and the Hölderian error bound modulus  $\rho$  always equals one for the TRS. Hence, we successfully extend the equivalence between error bounds and the KL inequality from the convex problems [8] to the nonconvex TRS. We shall emphasize here that for the TRS, both error bounds and the KL inequality results, as well as their relations, are new in the literature. Equipped with these thorough understandings, we are able to derive convergence rate results for algorithms for solving the TRS. As a byproduct, the convergence rate of projected gradient methods considered in [4] is studied. Specifically, we show that projected gradient methods converge to a global optimal solution locally sublinearly in the *TRS-ill case* and linearly otherwise.

The remaining of this paper is organized as follows. In Section 2, we first review some existing results in the literature that will be used in our proof. Then, we conduct a thorough analysis about the Hölderian error bound and the KL inequality for the TRS in Sections 3 and 4, respectively. In Section 5, we study the convergence rate of projected gradient methods for solving the TRS with the help of the KL inequality. We conclude our paper in Section 6.

**Notation.** For any vector  $\mathbf{x} \in \mathbb{R}^n$ , we use  $[\mathbf{x}]_+$  to denote its positive part. We use  $(\cdot)^\dagger$  to denote the *Moore-Penrose* pseudoinverse of a matrix. For any given nonempty closed set  $C$ , the distance between any given vector  $\mathbf{x} \in \mathbb{R}^n$  to  $C$  is denoted by  $\text{dist}(\mathbf{x}, C) := \min_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$ . Meanwhile, define the possibly set-valued projection mapping  $\Pi_C : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  as  $\Pi_C(\mathbf{x}) = \{\mathbf{v} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{v}\| = \text{dist}(\mathbf{x}, C)\}$ . We define  $\mathbf{0}$  as the vector (or matrix) of all zeros and its dimension will be clear from the context.

## 2 Preliminaries

We recall some basic properties associated with the TRS. Let  $\mathbf{A} = \mathbf{P}\mathbf{A}\mathbf{P}^T$  be the spectral decomposition of  $\mathbf{A}$  with  $\mathbf{P}$  being the orthogonal matrix and  $\mathbf{\Lambda} = \text{Diag}(\lambda)$  being the diagonal matrix with  $\lambda_1 \leq \dots \leq \lambda_n$ . When  $\lambda_1 < 0$  (i.e., nonconvex trust region subproblems), we consider the following convex relaxation:

$$\begin{aligned} (\text{P}_1) \quad & \min \tilde{f}(\mathbf{x}) := \mathbf{x}^T(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} - 2\mathbf{b}^T\mathbf{x} + \lambda_1 \\ & \text{s.t. } \|\mathbf{x}\|^2 \leq 1. \end{aligned}$$

Problem (P<sub>1</sub>) is regarded as a relaxation of (P<sub>0</sub>) since

$$[g(\mathbf{x})]_+ = [\|\mathbf{x}\|^2 - 1]_+ = 0 \quad \text{and} \quad \tilde{f}(\mathbf{x}) = f(\mathbf{x}) - \lambda_1(\mathbf{x}^T\mathbf{x} - 1) \leq f(\mathbf{x}) \quad (1)$$

whenever  $\|\mathbf{x}\| \leq 1$  and  $\lambda_1 < 0$ . We shall emphasize here that the relaxation (P<sub>1</sub>) plays a central role in our subsequent analysis. Throughout this paper, we define the solution set of the problem (P<sub>1</sub>) as  $S_1$ . We summarize in the following lemma the corresponding results obtained in [20, Lemmas 1 and 2] to reveal the relations between  $S_0$  and  $S_1$ .

**Lemma 2.1.** *Suppose  $\lambda_1 < 0$ . Then we have the following:*

1. for all  $x^* \in S_0$ ,  $\|\mathbf{x}^*\| = 1$ ,
2. for any  $\mathbf{x}^* \in S_0$ ,  $\tilde{f}(\mathbf{x}^*) = f(\mathbf{x}^*)$ , and
3. hence  $S_0 = S_1 \cap \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$ .

*Proof.* The claim 1 follows from [20, Lemma 1]. The claim 2 can be verified by noting (1) and  $\|\mathbf{x}^*\| = 1$  for all  $\mathbf{x}^* \in S_0$ . From [20, Lemmas 1 and 2], we know that  $\emptyset \neq S_1 \cap \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\} \subset S_0$ . Now let  $\mathbf{x}_1^* \in S_1 \cap \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$ . Then for all  $\mathbf{x}_0^* \in S_0$ , we have

$$\tilde{f}(\mathbf{x}_1^*) = f(\mathbf{x}_1^*) = f(\mathbf{x}_0^*) = \tilde{f}(\mathbf{x}_0^*),$$

i.e.,  $\mathbf{x}_0^* \in S_1 \cap \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$ . Therefore,  $S_0 \subset S_1 \cap \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$ , which is just claim 3.  $\square$

Lemma 2.1 also implies the well-known optimality conditions [20, 11] for the TRS, i.e.,  $\mathbf{x}^* \in \mathbb{R}^n$  is an optimal solution to problem (P<sub>0</sub>) if and only if for some  $\lambda^* \in \mathbb{R}$ ,  $(\mathbf{x}^*, \lambda^*)$  satisfies the following KKT conditions:

$$\begin{aligned} \|\mathbf{x}^*\|^2 &\leq 1, \\ (\mathbf{A} + \lambda^* \mathbf{I})\mathbf{x}^* &= \mathbf{b}, \\ \mathbf{A} + \lambda^* \mathbf{I} &\succeq \mathbf{0}, \\ \lambda^* &\geq 0, \\ \lambda^*(1 - \|\mathbf{x}^*\|^2) &= 0 \quad (\text{complementary slackness}). \end{aligned}$$

For the nonconvex TRS, i.e.,  $\lambda_1 < 0$ , it is well-known that the problem can be categorized into easy and hard cases<sup>1</sup> ([21]). A brief review about the two cases is given as follows:

1. In the easy case,  $\mathbf{b} \not\perp \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ , which implicitly implies that  $\lambda^* > -\lambda_1$ . In this case, the optimal solution is unique, and is given by  $\mathbf{x}^* = (\mathbf{A} + \lambda^* \mathbf{I})^{-1} \mathbf{b}$  with  $\|\mathbf{x}^*\| = 1$ .
2. In the hard case,  $\mathbf{b} \perp \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ . In this case, the optimal solution may not be unique. In fact, the optimal solution is given by either  $\mathbf{x}^* = (\mathbf{A} + \lambda^* \mathbf{I})^{-1} \mathbf{b}$  for some optimal Lagrange multiplier  $\lambda^* > -\lambda_1$  (or called hard case 1 in [21]) or  $\mathbf{x}^* = (\mathbf{A} - \lambda_1 \mathbf{I})^\dagger \mathbf{b} + \mathbf{v}$ , where  $\mathbf{v} \in \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$  such that  $\|\mathbf{x}^*\| = 1$  (or called hard case 2 in [21]). Particularly, the case with  $\mathbf{v} = \mathbf{0}$  is called hard case 2 (i), and the case with  $\mathbf{v} \neq \mathbf{0}$  is called hard case 2 (ii).

We summarize the above characterizations in the following table.

Easy case	Hard case 1	Hard case 2 (i)	Hard case 2 (ii)
$\mathbf{b} \not\perp \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ (implies $\lambda^* > -\lambda_1$ )	$\mathbf{b} \perp \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ and $\lambda^* > -\lambda_1$	$\mathbf{b} \perp \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ , $\lambda^* = -\lambda_1$ and $\ (\mathbf{A} - \lambda_1 \mathbf{I})^\dagger \mathbf{b}\  = 1$	$\mathbf{b} \perp \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ , $\lambda^* = -\lambda_1$ and $\ (\mathbf{A} - \lambda_1 \mathbf{I})^\dagger \mathbf{b}\  < 1$

Table 1: Different cases for the nonconvex TRS, i.e.,  $\lambda_1 < 0$ .

<sup>1</sup>We should point out that in this paper we use the categories of the easy and hard cases only for the nonconvex case, which is slightly different from the categories in [21].

### 3 Hölderian error bounds for the TRS

In this section, we present our main results on Hölderian error bounds for the TRS. Mainly, our analysis will be divided into two cases. First, we consider the convex case, i.e., the case with  $\lambda_1 \geq 0$ . The more challenging nonconvex case with  $\lambda_1 < 0$  will be discussed later.

#### 3.1 Case with $\lambda_1 \geq 0$

In this case, problem (P<sub>0</sub>) is convex and Hölderian error bounds for the TRS can be obtained by applying the elegant error bound results derived in [59] for convex quadratic inequalities. Let  $[m] := \{1, 2, \dots, m\}$ . We recall the main definitions and results in [59].

**Definition 3.1.** *Consider the inequality system*

$$q_i(\mathbf{x}) \leq 0, \quad \forall i \in [m].$$

*An inequality  $q_i(\mathbf{x}) \leq 0$  in the system is said to be singular if  $q_i(\mathbf{x}) = 0$  for any solution  $\mathbf{x}$  to the system. Thus an inequality  $q_i(\mathbf{x}) \leq 0$  is nonsingular if there is a solution  $\mathbf{x}_i$  to the system such that  $q_i(\mathbf{x}_i) < 0$ . If every inequality in the system is singular, we say that the inequality system is singular.*

**Definition 3.2.** *Let  $S$  be a singular system of inequalities. We say that  $S$  is critical, if either one of the following two conditions holds:*

1. *at most one of the inequalities is nonlinear; or,*
2. *after any one of the nonlinear inequalities is deleted, all the remaining nonlinear inequalities become nonsingular.*

**Definition 3.3.** *An inequality in a system is called irregular if it is nonlinear, singular, and contained in no critical subsystem.*

The following definition defines a concept of the degree of singularity, which will be used to determine the modulus of Hölderian error bounds for a convex quadratic inequality system.

**Definition 3.4.** *Let*

$$q_i(\mathbf{x}) \leq 0, \quad \forall i \in [m].$$

*be a system of inequalities. If there is no nonlinear, singular inequality, we define the degree of singularity of this system to be zero. If there is at least one such inequality, we define the degree of singularity of the system to be one plus the number of irregular inequalities.*

The main technical result we will use is the following global error bound for convex quadratic inequality systems.

**Lemma 3.5** (Theorem 3.1 in [59]). *Suppose*

$$q_i(\mathbf{x}) \leq 0, \quad \forall i \in [m]$$

*is a convex quadratic system with a nonempty solution set  $S$  and let  $[m] = K \oplus J$ , where  $K$  is the index set of all the nonsingular constraints and  $J$  is the index set of all the singular constraints. Then there exists a constant  $\tau > 0$  such that,*

$$\text{dist}(\mathbf{x}, S) \leq \tau \left( \sum_{i=1}^m [q_i(\mathbf{x})]_+ + \sum_{j \in J} [q_j(\mathbf{x})]_+^{1/2^d} \right), \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (2)$$

*where  $d$  is the degree of singularity of the system<sup>2</sup>.*

<sup>2</sup>The original inequality in [59, Theorem 3.1] is  $\text{dist}(\mathbf{x}, S) \leq \tau (\|[q_K(\mathbf{x})]_+\| + \|[q_J(\mathbf{x})]_+\| + \|[q_J(\mathbf{x})]_+\|^{1/2^d})$ , from which inequality (2) follows directly with a possibly rescaling of the constant  $\tau$ .

We note that one important feature of Lemma 3.5 is that the exponent of the term  $[q_J(\mathbf{x})]_+$  in the above inequality is related to  $d$ , the degree of singularity of the system. As one can observe later, it is this special and computable quantity that makes our analysis possible. Indeed, when Lemma 3.5 is applied to system

$$f(\mathbf{x}) - f^* \leq 0, \quad g(\mathbf{x}) \leq 0, \quad (3)$$

the main task will be computing its degree of singularity. We first consider a case that the system is minimal, i.e., deleting either inequality yields the system nonsingular.

**Lemma 3.6.** *Assume that  $\lambda_1 \geq 0$  and  $\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) < f^*$ . Then there exists a constant  $\tau > 0$  such that*

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (4)$$

*Proof.* From the definition of  $f^*$ , we know that the solution set of system (3) is  $S_0$ . Since  $\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) < f^*$ , it further holds that system (3) is singular. Moreover, we see that when either inequality in (3) is deleted, the remaining inequality is nonsingular. Hence, from Definition 3.2, we know that system (3) is critical. Therefore, there is no irregular inequality in (3) and the degree of singularity  $d$  of system (3) equals to 1. Hence, by using Lemma 3.5, we have that for all  $\mathbf{x} \in \mathbb{R}^n$ ,

$$\text{dist}(\mathbf{x}, S_0) \leq \tau_1 \left( [f(\mathbf{x}) - f^*]_+ + [g(\mathbf{x})]_+ + [f(\mathbf{x}) - f^*]_+^{1/2} + [g(\mathbf{x})]_+^{1/2} \right). \quad (5)$$

By Weierstrass theorem, we know that  $[f(\mathbf{x}) - f^*]_+$  is upper bounded over the unit ball, i.e., there exists a constant  $M > 0$  such that

$$M = \max \{ [f(\mathbf{x}) - f^*]_+ \mid \|\mathbf{x}\| \leq 1 \}.$$

If  $M \leq 1$ , inequality (5) implies that (4) holds with  $\tau = 2\tau_1$ . If  $M > 1$ , we have from (5) that

$$\text{dist}(\mathbf{x}, S_0) \leq 2\tau_1 (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \text{ with } [f(\mathbf{x}) - f^*]_+ \leq 1,$$

and

$$\begin{aligned} \text{dist}(\mathbf{x}, S_0) &\leq \tau_1 (M + (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}), \\ &\leq \tau_1 (M + 1) (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \text{ with } [f(\mathbf{x}) - f^*]_+ > 1. \end{aligned}$$

Combining all the above discussions, we see that (4) holds with  $\tau = \max\{2, M + 1\}\tau_1$ .  $\square$

Now consider the case that  $\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) = f^*$ . In this case, it is easy to see that  $\{f(\mathbf{x}) - f^* \leq 0\}$  is a singular system. Hence, the degree of singularity of system (3) and the corresponding error bound modulus depend on the singularity of the second inequality. Indeed, since  $\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) = f^* > -\infty$ , we know that  $\mathbf{b} \in \text{Range}(\mathbf{A})$  and for all  $\mathbf{x} \in \mathbb{R}^n$ ,

$$f(\mathbf{x}) - f^* = (\mathbf{x} - \tilde{\mathbf{x}})^T \mathbf{A} (\mathbf{x} - \tilde{\mathbf{x}}),$$

where  $\tilde{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{b}$  and  $\|\tilde{\mathbf{x}}\| \leq 1$ . Therefore, the optimal solution set of problem (P) can be written as

$$S_0 = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \tilde{\mathbf{x}} + \mathbf{d}, \|\mathbf{x}\| \leq 1, \mathbf{d} \in \text{Null}(\mathbf{A}) \}. \quad (6)$$

With these discussions in hand, we are ready to derive a Hölderian error bound in the following lemma.

**Lemma 3.7.** *Assume that  $\lambda_1 \geq 0$  and  $\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) = f^*$ . Then, it holds that*

1. if  $\lambda_1 > 0$ , then

$$\text{dist}(\mathbf{x}, S_0) \leq \sqrt{\frac{1}{\lambda_1}} (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \in \mathbb{R}^n;$$

2. if  $\lambda_1 = 0$  and  $\|\tilde{\mathbf{x}}\| = 1$ , then  $S_0 = \{\tilde{\mathbf{x}}\}$  and there exists a constant  $\tau > 0$  such that

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/4}, \quad \forall \mathbf{x} \in \mathbb{R}^n;$$

3. if  $\lambda_1 = 0$  and  $\|\tilde{\mathbf{x}}\| < 1$ , then there exists a constant  $\tau > 0$  such that

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Proof. Case 1 follows directly from the fact that  $f(\mathbf{x}) - f^* = (\mathbf{x} - \tilde{\mathbf{x}})^T \mathbf{A}(\mathbf{x} - \tilde{\mathbf{x}}) \geq \lambda_1 \|\mathbf{x} - \tilde{\mathbf{x}}\|^2$  and  $S_0 = \{\tilde{\mathbf{x}}\}$ .

For case 2, since  $\lambda_1 = 0$  and  $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| < 1\} \cap S_0 = \emptyset$ , from (6), we know that  $S_0 = \{\tilde{\mathbf{x}}\}$ , i.e.,  $S_0$  is a singleton. Thus, the second inequality  $g(\mathbf{x}) \leq 0$  in system (3) is singular. Meanwhile, the assumption that  $\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) = f^*$  implies that the only critical subsystem of (3) is the first inequality  $f(\mathbf{x}) - f^* \leq 0$ . Therefore,  $g(\mathbf{x}) \leq 0$  is irregular and the degree of singularity of (3) equals to 2. Then, Lemma 3.5 asserts that there exists a constant  $\tau_1 > 0$  such that

$$\text{dist}(\mathbf{x}, S_0) \leq \tau_1 \left( [f(\mathbf{x}) - f^*]_+ + [g(\mathbf{x})]_+ + [f(\mathbf{x}) - f^*]_+^{1/4} + [g(\mathbf{x})]_+^{1/4} \right)$$

for all  $\mathbf{x} \in \mathbb{R}^n$ . Following the same arguments in the proof of Lemma 3.6, we know that there exists a constant  $\tau > 0$  such that

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/4}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

In case 3, we have  $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| < 1\} \cap S_0 \neq \emptyset$  and thus the second inequality  $\|\mathbf{x}\|^2 - 1 \leq 0$  is nonsingular. Hence the degree of singularity of system (3) equals to 1 as there are no irregular inequalities. Then, Lemma 3.5 asserts that there exists a constant  $\tau_2 > 0$  such that

$$\text{dist}(\mathbf{x}, S_0) \leq \tau_2 \left( [f(\mathbf{x}) - f^*]_+ + [g(\mathbf{x})]_+ + [f(\mathbf{x}) - f^*]_+^{1/2} + [g(\mathbf{x})]_+^{1/2} \right)$$

for all  $\mathbf{x} \in \mathbb{R}^n$ . Similar as the proof of Lemma 3.6, we have

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

This completes the proof for the lemma. □

### 3.2 Case with $\lambda_1 < 0$

In this section, we turn our interests to the nonconvex TRS. As is stated in the preliminary section, the nonconvex TRS includes the easy and hard cases. We will derive error bounds for them separately.

Before diving into the proofs, we shall discuss the main ideas here. Since  $\lambda_1 < 0$ , the first inequality in the quadratic inequality system (3) is nonconvex. Fortunately, the following quadratic inequality system

$$\tilde{f}(\mathbf{x}) - f^* \leq 0, \quad g(\mathbf{x}) \leq 0, \tag{7}$$

derived from the reformulation (P<sub>1</sub>) (see Lemma 2.1), is always a convex one. Hence, we can apply Lemma 3.5 to system (7) and then use the relations between the solution sets of systems (3) and (7) to establish meaningful error bounds for nonconvex quadratic inequality system (3).

We first study the easy case and hard case 1. In these two cases, the solution for problem (P) is unique and has unit norm.

**Lemma 3.8.** *In the easy case or hard case 1, there exists some constant  $\tau > 0$  such that*

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Proof. In both the easy case and hard case 1, we see that  $\min_{\mathbf{x} \in \mathbb{R}^n} \tilde{f}(\mathbf{x}) < f^*$ . Indeed, for the easy case, since  $\mathbf{b} \notin \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ , we have  $\mathbf{b} \notin \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$ . Then, it holds that  $\min \tilde{f}(\mathbf{x}) = -\infty < f^*$ . For hard case 1, the optimal solution for  $\min_{\mathbf{x} \in \mathbb{R}^n} \tilde{f}(\mathbf{x})$  is achieved by  $(\mathbf{A} - \lambda_1 \mathbf{I})^\dagger \mathbf{b}$ , whose norm is larger than 1. This gives  $\min_{\mathbf{x} \in \mathbb{R}^n} \tilde{f}(\mathbf{x}) < f^*$ . Similar to the case studied in Lemma 3.6, the degree of singularity of system (7) equals to 1. Hence, there exists some constant  $\tau > 0$  such that for all  $\mathbf{x} \in \mathbb{R}^n$ ,

$$\text{dist}(\mathbf{x}, S_1) \leq \tau(\tilde{f}(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}.$$

From (1), we see that

$$\text{dist}(\mathbf{x}, S_1) \leq \tau(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Note that in both the easy case and hard case 1, since  $\min_{\mathbf{x} \in \mathbb{R}^n} \tilde{f}(\mathbf{x}) < f^*$ , we have that  $\|\mathbf{x}\| = 1$  for all  $\mathbf{x} \in S_1$ . This further implies that  $S_1 = S_0$  (in fact,  $S_1 = S_0 = \{(\mathbf{A} - \lambda^* \mathbf{I})^\dagger \mathbf{b}\}$  for a unique  $\lambda^* > -\lambda_1$  such that  $\|(\mathbf{A} - \lambda^* \mathbf{I})^\dagger \mathbf{b}\| = 1$ ) and thus completes the proof.  $\square$

In hard case 2,  $\mathbf{b}$  is orthogonal to the eigenspace of matrix  $\mathbf{A}$  corresponding to the smallest eigenvalue, i.e.,  $\mathbf{b} \perp \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$  and  $f^* = \min_{\mathbf{x} \in \mathbb{R}^n} \tilde{f}(\mathbf{x})$ . Denote  $\bar{\mathbf{x}} = (\mathbf{A} - \lambda_1 \mathbf{I})^\dagger \mathbf{b}$ . Then, we have

$$\begin{aligned} S_0 &= \{\bar{\mathbf{x}} + \mathbf{v} \in \mathbb{R}^n \mid \|\bar{\mathbf{x}} + \mathbf{v}\| = 1, \mathbf{v} \in \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})\} \quad \text{and} \\ S_1 &= \{\bar{\mathbf{x}} + \mathbf{v} \in \mathbb{R}^n \mid \|\bar{\mathbf{x}} + \mathbf{v}\| \leq 1, \mathbf{v} \in \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})\}. \end{aligned} \quad (8)$$

Let us first consider the hard case 2 (i).

**Lemma 3.9.** *In the hard case 2 (i), there exists  $\tau > 0$  such that*

$$\text{dist}(\mathbf{x}, S_0) \leq \tau(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/4}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Proof. From the assumption, we know that  $S_0 = S_1 = \{\bar{\mathbf{x}}\}$ . By applying Lemma 3.7 case 2 to problem (P<sub>1</sub>), we obtain that

$$\text{dist}(\mathbf{x}, S_0) = \text{dist}(\mathbf{x}, S_1) \leq \tau(\tilde{f}(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/4}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

The conclusion then follows directly from (1).  $\square$

Next we focus on the hard case 2 (ii), in which  $S_1$  is not a singleton since  $\|\bar{\mathbf{x}}\| < 1$ . To establish the desired error bound inequality in this case, we need to characterize the connection between  $\text{dist}(\mathbf{x}, S_0)$  and  $\text{dist}(\mathbf{x}, S_1)$ . For this purpose, we establish the following technical lemma.

**Lemma 3.10.** *Given  $\mathbf{x} \in \mathbb{R}^n$ , let  $\mathbf{x}_1 := \Pi_{S_1}(\mathbf{x})$  be the projection of  $\mathbf{x}$  onto the convex set  $S_1$ . In the hard case 2 (ii), it holds that*

1. if  $\mathbf{x}_1 = \bar{\mathbf{x}}$ , then  $\Pi_{S_0}(\mathbf{x}) = S_0$  and  $\text{dist}(\mathbf{x}, S_0) = \sqrt{\|\mathbf{x} - \bar{\mathbf{x}}\|^2 + 1 - \|\bar{\mathbf{x}}\|^2}$ ;
2. else if  $\mathbf{x}_1 \neq \bar{\mathbf{x}}$ , then  $\Pi_{S_0}(\mathbf{x}) = \bar{\mathbf{x}} + t_0 \mathbf{v}_1$  with  $\mathbf{v}_1 = \frac{\mathbf{x}_1 - \bar{\mathbf{x}}}{\|\mathbf{x}_1 - \bar{\mathbf{x}}\|} \in \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$  and  $t_0 \geq \|\mathbf{x}_1 - \bar{\mathbf{x}}\|$  such that  $\|\bar{\mathbf{x}} + t_0 \mathbf{v}_1\| = 1$ .

Proof. In the first case, since  $\mathbf{x}_1 = \Pi_{S_1}(\mathbf{x}) = \bar{\mathbf{x}}$ , we know that

$$\langle \mathbf{x} - \bar{\mathbf{x}}, \mathbf{z} - \bar{\mathbf{x}} \rangle \leq 0, \quad \forall \mathbf{z} \in S_1,$$

which, together with the structure of  $S_1$ , implies that  $\mathbf{x} - \bar{\mathbf{x}} \in \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$ . Thus, for any  $\mathbf{y} \in S_0$ , we have that

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x} - \bar{\mathbf{x}} + \bar{\mathbf{x}} - \mathbf{y}\|^2 = \|\mathbf{x} - \bar{\mathbf{x}}\|^2 + \|\bar{\mathbf{x}} - \mathbf{y}\|^2,$$



where the second equality follows from the facts that  $\mathbf{y} - \bar{\mathbf{x}} \in \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$  and thus  $\langle \mathbf{x} - \bar{\mathbf{x}}, \bar{\mathbf{x}} - \mathbf{y} \rangle = 0$ . This, together with  $1 = \|\mathbf{y}\| = \|\bar{\mathbf{x}} - \mathbf{y}\|^2 + \|\bar{\mathbf{x}}\|^2$  (due to  $\mathbf{y} - \bar{\mathbf{x}} \in \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$  and  $\bar{\mathbf{x}} \in \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$ ), implies

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{\|\mathbf{x} - \bar{\mathbf{x}}\|^2 + 1 - \|\bar{\mathbf{x}}\|^2}, \quad \forall \mathbf{y} \in S_0.$$

This completes the proof for the first case.

We next consider the second case. If  $\|\mathbf{x}_1\| = 1$ , we know that  $\mathbf{x}_1 \in \Pi_{S_0}(\mathbf{x})$ . For any  $\tilde{\mathbf{x}} \in \Pi_{S_0}(\mathbf{x})$ , it holds that

$$\tilde{\mathbf{x}} \in S_0 \subseteq S_1 \quad \text{and} \quad \text{dist}(\mathbf{x}, S_1) = \|\mathbf{x} - \mathbf{x}_1\| = \text{dist}(\mathbf{x}, S_0) = \|\mathbf{x} - \tilde{\mathbf{x}}\|,$$

and consequently,  $\tilde{\mathbf{x}} = \Pi_{S_1}(\mathbf{x}) = \mathbf{x}_1$  due to the uniqueness of the projection onto the convex set  $S_1$ . Hence,  $\Pi_{S_0}(\mathbf{x})$  is a singleton, i.e.,

$$\Pi_{S_0}(\mathbf{x}) = \mathbf{x}_1 = \bar{\mathbf{x}} + \|\mathbf{x}_1 - \bar{\mathbf{x}}\| \frac{\mathbf{x}_1 - \bar{\mathbf{x}}}{\|\mathbf{x}_1 - \bar{\mathbf{x}}\|}.$$

Now we consider the case with  $\|\mathbf{x}_1\| < 1$ . Without loss of generality, assume that the null space of  $\mathbf{A} - \lambda_1 \mathbf{I}$  is spanned by an orthogonal basis  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  with some  $k \geq 1$ ,  $\mathbf{v}_1$  being the non-zero vector  $\frac{\mathbf{x}_1 - \bar{\mathbf{x}}}{\|\mathbf{x}_1 - \bar{\mathbf{x}}\|}$  and all other  $\mathbf{v}_i$  being unit-norm vectors. Then, we can rewrite the solution set of (P<sub>0</sub>) as  $S_0 = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \bar{\mathbf{x}} + \sum_{i=1}^k \alpha_i \mathbf{v}_i, \|\mathbf{x}\| = 1\}$ . Since  $\mathbf{x}_1 = \Pi_{S_1}(\mathbf{x})$ , we have

$$\langle \mathbf{x} - \mathbf{x}_1, \mathbf{z} - \mathbf{x}_1 \rangle \leq 0, \quad \forall \mathbf{z} \in S_1.$$

This, together with (8) and  $\|\mathbf{x}_1\| < 1$ , implies that  $\langle \mathbf{x} - \mathbf{x}_1, \mathbf{d} \rangle = 0$  for all  $\mathbf{d} \in \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ . Therefore, there exists some vector  $\mathbf{s} \in \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$  such that  $\mathbf{x} = \mathbf{x}_1 + \mathbf{s}$ . Consider the projection of  $\mathbf{x}$  onto  $S_0$ :

$$\min \|\mathbf{x} - \mathbf{z}\|^2 \quad \text{s.t.} \quad \mathbf{z} \in S_0, \quad (9)$$

which, due to  $\mathbf{x} = \bar{\mathbf{x}} + \|\mathbf{x}_1 - \bar{\mathbf{x}}\| \mathbf{v}_1 + \mathbf{s}$  and  $\mathbf{z} = \bar{\mathbf{x}} + \sum_{i=1}^k \mu_i \mathbf{v}_i$ , is equivalent to

$$\min_{\mu_1, \dots, \mu_k} \left\| \|\mathbf{x}_1 - \bar{\mathbf{x}}\| \mathbf{v}_1 + \mathbf{s} - \sum_{i=1}^k \mu_i \mathbf{v}_i \right\|^2 \quad \text{s.t.} \quad \left\| \bar{\mathbf{x}} + \sum_{i=1}^k \mu_i \mathbf{v}_i \right\| = 1. \quad (10)$$

Since  $\mathbf{s}, \mathbf{v}_1, \dots, \mathbf{v}_k$  are orthogonal to each other and  $\|\mathbf{v}_i\| = 1$  for  $i = 1, \dots, k$ , the objective in (10) can be further written as:

$$\left\| \|\mathbf{x}_1 - \bar{\mathbf{x}}\| \mathbf{v}_1 + \mathbf{s} - \sum_{i=1}^k \mu_i \mathbf{v}_i \right\|^2 = \|\mathbf{x}_1 - \bar{\mathbf{x}}\|^2 - 2\mu_1 \|\mathbf{x}_1 - \bar{\mathbf{x}}\| + \sum_{i=1}^k \mu_i^2 + \|\mathbf{s}\|^2.$$

We further note that for any feasible solution  $(\mu_1, \dots, \mu_k)$  to (10), it holds that  $\|\bar{\mathbf{x}}\|^2 + \sum_{i=1}^k \mu_i^2 = 1$ . Hence, (10), and consequently (9), can be equivalently rewritten as

$$\min_{\mu_1, \dots, \mu_k} \|\mathbf{x}_1 - \bar{\mathbf{x}}\|^2 - 2\mu_1 \|\mathbf{x}_1 - \bar{\mathbf{x}}\| + 1 - \|\bar{\mathbf{x}}\|^2 + \|\mathbf{s}\|^2 \quad \text{s.t.} \quad \|\bar{\mathbf{x}}\|^2 + \sum_{i=1}^k \mu_i^2 = 1,$$

whose unique optimal solution is clearly  $\mu_1^* = \sqrt{1 - \|\bar{\mathbf{x}}\|^2}$  and  $\mu_i^* = 0$  for  $i = 2, \dots, k$ . Therefore, the projection problem (9) has a unique optimal solution that takes the form  $\mathbf{z}^* = \bar{\mathbf{x}} + t_0 \mathbf{v}_1$  with  $t_0 > \|\mathbf{x}_1 - \bar{\mathbf{x}}\|$  such that  $\|\mathbf{z}^*\| = 1$ .  $\square$

Now we are ready to present the connection between  $\text{dist}(\mathbf{x}, S_0)$  and  $\text{dist}(\mathbf{x}, S_1)$  in the hard case 2 (ii).

**Lemma 3.11.** *In the hard case 2 (ii), there exists some constant  $\gamma \in (0, \pi/2)$  with*

$$\sin \gamma = \frac{1 - \|\bar{\mathbf{x}}\|}{\sqrt{(1 - \|\bar{\mathbf{x}}\|)^2 + 1 - \|\bar{\mathbf{x}}\|^2}} = \sqrt{\frac{1 - \|\bar{\mathbf{x}}\|}{2}}$$

such that

$$\text{dist}(\mathbf{x}, S_1) + \sqrt{1 - \|\mathbf{x}\|^2} \geq \text{dist}(\mathbf{x}, S_0) \sin \gamma, \quad \forall \mathbf{x} \in \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq 1\}. \quad (11)$$

*Proof.* For any given  $\mathbf{x} \in \mathbb{R}^n$  with  $\|\mathbf{x}\| \leq 1$  and any given  $\mathbf{x}_0 \in \Pi_{S_0}(\mathbf{x})$ , define  $\mathbf{v} = \mathbf{x}_0 - \bar{\mathbf{x}}$ . Since  $\|\bar{\mathbf{x}}\| < 1$  in the hard case 2 (ii) and  $\|\mathbf{x}_0\| = 1$ , we have that  $\|\mathbf{v}\| \neq 0$ . Define line segment  $L := \{\bar{\mathbf{x}} + \alpha \mathbf{v} \in \mathbb{R}^n \mid \alpha \in [0, 1]\}$ . We note from (8) and  $\bar{\mathbf{x}} = (\mathbf{A} - \lambda_1 \mathbf{I})^\dagger \mathbf{b}$  that  $\langle \mathbf{v}, \bar{\mathbf{x}} \rangle = 0$ . If  $\bar{\mathbf{x}} = \Pi_{S_1}(\mathbf{x})$ , then  $\Pi_{S_1}(\mathbf{x}) \in L$ ; otherwise, by case 2 in Lemma 3.10, we know that  $\mathbf{x}_0 = \Pi_{S_0}(\mathbf{x}) = \bar{\mathbf{x}} + t(\Pi_{S_1}(\mathbf{x}) - \bar{\mathbf{x}})$  for some constant  $t \geq 1$ . Hence,  $\Pi_{S_1}(\mathbf{x}) - \bar{\mathbf{x}} = \mathbf{v}/t$ , i.e.,  $\Pi_{S_1}(\mathbf{x}) \in L$ . Since  $L \subseteq S_1$ , it further holds that

$$\Pi_L(\mathbf{x}) = \Pi_{S_1}(\mathbf{x}). \quad (12)$$

Let  $\Pi_L(\mathbf{x}) = \bar{\mathbf{x}} + \alpha^* \mathbf{v}$ . By the definition of  $L$  and the properties of the projection operator  $\Pi_L$ , it is not difficult to see that

$$\langle \mathbf{x} - \Pi_L(\mathbf{x}), \mathbf{v} \rangle \begin{cases} = 0, & \text{if } \alpha^* \in (0, 1), \\ \leq 0, & \text{if } \alpha^* = 0, \\ \geq 0, & \text{if } \alpha^* = 1. \end{cases}$$

We first consider the case where  $\langle \mathbf{x} - \Pi_L(\mathbf{x}), \mathbf{v} \rangle > 0$ . In this case, we have  $\alpha^* = 1$  and thus

$$\text{dist}(\mathbf{x}, S_1) = \|\mathbf{x} - \Pi_{S_1}(\mathbf{x})\| = \|\mathbf{x} - \Pi_L(\mathbf{x})\| = \|\mathbf{x} - (\bar{\mathbf{x}} + \mathbf{v})\| = \|\mathbf{x} - \mathbf{x}_0\| = \text{dist}(\mathbf{x}, S_0),$$

i.e., (11) holds trivially. Then, we argue that  $\langle \mathbf{x} - \Pi_L(\mathbf{x}), \mathbf{v} \rangle < 0$  cannot occur. Indeed, if this is not the case, we must have  $\alpha^* = 0$  and  $\langle \mathbf{x} - \Pi_L(\mathbf{x}), \mathbf{v} \rangle < 0$  and thus  $\langle \mathbf{x} - \bar{\mathbf{x}}, \mathbf{v} \rangle < 0$ . Hence we further have that

$$\begin{aligned} \|\mathbf{x} - (\bar{\mathbf{x}} - \mathbf{v})\|^2 &= \|\mathbf{x} - \bar{\mathbf{x}}\|^2 + 2\langle \mathbf{x} - \bar{\mathbf{x}}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \\ &< \|\mathbf{x} - \bar{\mathbf{x}}\|^2 - 2\langle \mathbf{x} - \bar{\mathbf{x}}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \\ &= \|\mathbf{x} - (\bar{\mathbf{x}} + \mathbf{v})\|^2 = \|\mathbf{x} - \mathbf{x}_0\|^2. \end{aligned} \quad (13)$$

Since  $\langle \bar{\mathbf{x}}, \mathbf{v} \rangle = 0$ , it holds that  $\|\bar{\mathbf{x}} - \mathbf{v}\|^2 = \|\bar{\mathbf{x}}\|^2 + \|\mathbf{v}\|^2 = \|\bar{\mathbf{x}} + \mathbf{v}\|^2 = 1$ . This, together with the definition of  $S_0$  in (8), implies that  $\bar{\mathbf{x}} - \mathbf{v} \in S_0$ . Since  $\mathbf{x}_0 \in \Pi_{S_0}(\mathbf{x})$ , it holds that  $\|\mathbf{x} - \mathbf{x}_0\| \leq \|\mathbf{x} - (\bar{\mathbf{x}} - \mathbf{v})\|$ , which contradicts (13).

Thus, in the subsequent analysis, we only need to focus on the case where  $\langle \mathbf{x} - \Pi_L(\mathbf{x}), \mathbf{v} \rangle = 0$ .

We first consider the case where  $\bar{\mathbf{x}} = \mathbf{0}$ . In this case,  $\mathbf{v} = \mathbf{x}_0$ . Let  $\mathbf{u} = \mathbf{x} - \Pi_L(\mathbf{x}) = \mathbf{x} - \alpha^* \mathbf{v}$ . Hence we have  $\mathbf{u} \perp \mathbf{v}$ . It holds that  $\text{dist}(\mathbf{x}, S_1) = \|\mathbf{u}\|$ , and

$$\text{dist}(\mathbf{x}, S_0) \leq \|\mathbf{x} - \mathbf{x}_0\| = \|\mathbf{u} - (1 - \alpha^*)\mathbf{v}\|.$$

We then claim that (11) holds. To see this, since

$$\text{dist}(\mathbf{x}, S_1)^2 + 1 - \|\mathbf{x}\|^2 \leq \left( \text{dist}(\mathbf{x}, S_1) + \sqrt{1 - \|\mathbf{x}\|^2} \right)^2,$$

it suffices to show

$$\text{dist}(\mathbf{x}, S_0)^2 \sin^2 \gamma \leq \text{dist}(\mathbf{x}, S_1)^2 + 1 - \|\mathbf{x}\|^2,$$

which is equivalent to

$$\|\mathbf{u} + (1 - \alpha^*)\mathbf{v}\|^2 \leq 2(\|\mathbf{u}\|^2 + 1 - \|\mathbf{x}\|^2) \quad (14)$$

due to  $\sin \gamma = 1/\sqrt{2}$  (since  $\bar{\mathbf{x}} = \mathbf{0}$ ),  $\mathbf{x} = \mathbf{u} + \alpha^* \mathbf{v}$  and  $\mathbf{u} \perp \mathbf{v}$ . Meanwhile, (14) is further equivalent to

$$\|\mathbf{u}\|^2 + (1 - \alpha^*)^2 \|\mathbf{v}\|^2 + 2\|\alpha^* \mathbf{v}\|^2 \leq 2,$$

which is trivial since  $\|\mathbf{x}\|^2 = \|\mathbf{u}\|^2 + \|\alpha^*\mathbf{v}\|^2 \leq 1$ ,  $0 \leq \alpha^* \leq 1$  and  $\|\mathbf{v}\| = 1$ .

In the following of this proof, we consider the remaining case where  $\bar{\mathbf{x}} \neq \mathbf{0}$ . In this case, we derive a lower bound of  $\|\mathbf{x} - \Pi_{S_1}(\mathbf{x})\|$ , which is equivalent to  $\|\mathbf{x} - \Pi_L(\mathbf{x})\|$  due to (12), by considering the following optimization problem:

$$\min_{\mathbf{z} \in \mathbb{R}^n} \{ \|\mathbf{z} - \Pi_L(\mathbf{z})\| \mid \langle \mathbf{z} - \Pi_L(\mathbf{z}), \mathbf{v} \rangle = 0, \|\mathbf{z}\| = \|\mathbf{x}\|, \|\mathbf{z} - \mathbf{x}_0\| = \text{dist}(\mathbf{x}, S_0) \}. \quad (15)$$

From (12) and  $\langle \mathbf{x} - \Pi_L(\mathbf{x}), \mathbf{v} \rangle = 0$ , we see that problem (15) has a non-empty closed and bounded feasible set as  $\mathbf{x}$  is always a feasible solution. Since the objective is continuous in  $\mathbf{z}$ , by Weierstrass theorem we know that problem (15) has a non-empty and compact solution set. Let  $\mathbf{z}^*$  be any optimal solution to problem (15). We know that  $\text{dist}(\mathbf{x}, S_1) = \|\mathbf{x} - \Pi_L(\mathbf{x})\| \geq \|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\|$  as  $\mathbf{x}$  is a feasible solution to (15). Since  $\|\mathbf{x}\| = \|\mathbf{z}^*\|$ , it further holds that

$$\text{dist}(\mathbf{x}, S_1) + \sqrt{1 - \|\mathbf{x}\|^2} \geq \|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| + \sqrt{1 - \|\mathbf{z}^*\|^2}. \quad (16)$$

Recall that  $\bar{\mathbf{x}} \neq \mathbf{0}$  and  $\mathbf{v} \neq \mathbf{0}$ . For any  $\mathbf{z} \in \mathbb{R}^n$ , since  $\mathbf{v} \perp \bar{\mathbf{x}}$ , we note that there exist  $\lambda, \mu \in \mathbb{R}$  and  $\mathbf{u} \in \mathbb{R}^n$  satisfying  $\langle \mathbf{u}, \bar{\mathbf{x}} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle = 0$  (note that  $\mathbf{u} = \mathbf{0}$  if  $n = 2$ ) such that

$$\mathbf{z} = \lambda \bar{\mathbf{x}} + \mu \mathbf{v} + \mathbf{u}.$$

Given the structure of  $L$ , we know that  $\Pi_L(\mathbf{z}) = \bar{\mathbf{x}} + \alpha \mathbf{v}$  for some  $\alpha \in [0, 1]$ . Now, if, in addition,  $\langle \mathbf{z} - \Pi_L(\mathbf{z}), \mathbf{v} \rangle = 0$ , it then holds that  $\mu = \alpha \in [0, 1]$ , and since  $\mathbf{x}_0 - \Pi_L(\mathbf{z}) = (\bar{\mathbf{x}} + \mathbf{v}) - (\bar{\mathbf{x}} + \mu \mathbf{v}) = (1 - \mu)\mathbf{v}$ , we further know that  $\langle \mathbf{z} - \Pi_L(\mathbf{z}), \mathbf{x}_0 - \Pi_L(\mathbf{z}) \rangle = 0$ . Hence, for any feasible solution  $\mathbf{z}$  to problem (15), it holds that

$$\|\mathbf{z} - \Pi_L(\mathbf{z})\|^2 = \|\mathbf{z} - \mathbf{x}_0\|^2 - \|\mathbf{x}_0 - \Pi_L(\mathbf{z})\|^2 = \text{dist}^2(\mathbf{x}, S_0) - (1 - \mu)^2 \|\mathbf{v}\|^2. \quad (17)$$

Therefore, problem (15) can be equivalently reformulated as

$$\begin{aligned} \min_{\mathbf{u} \in \mathbb{R}^n, \mu, \lambda \in \mathbb{R}} \quad & \mu \\ \text{s.t.} \quad & \lambda^2 \|\bar{\mathbf{x}}\|^2 + \mu^2 \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 = \|\mathbf{x}\|^2, \\ & (\lambda - 1)^2 \|\bar{\mathbf{x}}\|^2 + (\mu - 1)^2 \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 = \text{dist}^2(\mathbf{x}, S_0), \\ & \langle \mathbf{u}, \bar{\mathbf{x}} \rangle = 0, \langle \mathbf{u}, \mathbf{v} \rangle = 0, \mu \in [0, 1]. \end{aligned} \quad (18)$$

We proceed the proof by considering two cases where the dimension of problem (15) is  $n = 2$  or  $n \geq 3$ .

**Case I:** If  $n = 2$ , then for any feasible solution  $(\mu, \lambda, \mathbf{u})$  to (18), it holds that  $\mathbf{u} \equiv \mathbf{0}$ . Let  $(\lambda^*, \mu^*, \mathbf{0})$  be an optimal solution to (18). Then,  $\mathbf{z}^* = \lambda^* \bar{\mathbf{x}} + \mu^* \mathbf{v}$  is an optimal solution to (15). Define  $\tilde{\mathbf{z}} = \lambda^* \beta \bar{\mathbf{x}} + \mu^* \mathbf{v}$ , where  $\beta \geq 1$  is some constant such that  $\|\tilde{\mathbf{z}}\| = 1$ . Note that the above construction of  $\tilde{\mathbf{z}}$  implies that  $\Pi_L(\tilde{\mathbf{z}}) = \Pi_L(\mathbf{z}^*)$ . Let  $\tilde{\theta}$  be the angle between  $\tilde{\mathbf{z}} - \mathbf{x}_0$  and  $\bar{\mathbf{x}} - \mathbf{x}_0$ . Recalling  $\mathbf{x}_0 = \bar{\mathbf{x}} + \mathbf{v}$ , it is easy to verify that

$$\langle \tilde{\mathbf{z}} - \mathbf{x}_0, \bar{\mathbf{x}} - \mathbf{x}_0 \rangle = (1 - \mu^*) \|\mathbf{v}\|^2 \geq 0.$$

If  $\mu^* = 1$ , we have from (17) that  $\|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| = \text{dist}(\mathbf{x}, S_0)$ . Hence, we have

$$\sqrt{1 - \|\mathbf{x}\|^2} + \text{dist}(\mathbf{x}, S_1) \geq \sqrt{1 - \|\mathbf{z}^*\|^2} + \|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| \geq \text{dist}(\mathbf{x}, S_0) \geq \text{dist}(\mathbf{x}, S_0) \sin \gamma,$$

where the first inequality follows from (16). If  $\mu^* \in [0, 1)$ , since  $0 < \|\bar{\mathbf{x}}\| < 1$ , we know that  $\|\mathbf{v}\| > 0$ , and thus,  $\langle \tilde{\mathbf{z}} - \mathbf{x}_0, \bar{\mathbf{x}} - \mathbf{x}_0 \rangle > 0$ , i.e.,  $\tilde{\theta} \in [0, \pi/2)$ . We further argue that  $\tilde{\theta} \neq 0$ . Suppose for contradiction that  $\tilde{\theta} = 0$ , i.e.,  $\tilde{\mathbf{z}} - \mathbf{x}_0$  and  $\bar{\mathbf{x}} - \mathbf{x}_0$  are parallel. Since  $\tilde{\mathbf{z}} - \mathbf{x}_0 = (\lambda^* \beta - 1)\bar{\mathbf{x}} + (\mu^* - 1)\mathbf{v}$  and  $\bar{\mathbf{x}} - \mathbf{x}_0 = \mathbf{v}$ , it holds that  $\lambda^* \beta = 1$ , i.e.,  $\tilde{\mathbf{z}} = \bar{\mathbf{x}} + \mu^* \mathbf{v}$ . It then follows that

$$\|\tilde{\mathbf{z}}\|^2 = \|\bar{\mathbf{x}}\|^2 + (\mu^*)^2 \|\mathbf{v}\|^2 < \|\bar{\mathbf{x}}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{x}_0\| = 1,$$

contradicting  $\|\tilde{\mathbf{z}}\| = 1$ . Thus,  $\tilde{\theta} \in (0, \pi/2)$ . Note that  $\gamma \in (0, \pi/2)$  can be regarded as the angle between  $\frac{\tilde{\mathbf{x}}}{\|\tilde{\mathbf{x}}\|} - \mathbf{x}_0$  and  $\tilde{\mathbf{x}} - \mathbf{x}_0$ . Then, geometric arguments assert that  $\tan \tilde{\theta} \geq \tan \gamma$ , i.e.,  $\tilde{\theta} \geq \gamma$  (see Figure 1 for the illustration). Hence,

$$\|\tilde{\mathbf{z}} - \Pi_L(\tilde{\mathbf{z}})\| = \|\tilde{\mathbf{z}} - \Pi_L(\mathbf{z}^*)\| = \|\tilde{\mathbf{z}} - \mathbf{x}_0\| \sin \tilde{\theta} \geq \|\tilde{\mathbf{z}} - \mathbf{x}_0\| \sin \gamma.$$

First consider the case where  $\|\tilde{\mathbf{z}} - \mathbf{x}_0\| \geq \|\mathbf{z}^* - \mathbf{x}_0\|$ . In this case, we know that

$$\begin{aligned} \sqrt{1 - \|\mathbf{z}^*\|^2} + \|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| &\geq \|\tilde{\mathbf{z}} - \mathbf{z}^*\| + \|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| \\ &\geq \|\tilde{\mathbf{z}} - \Pi_L(\mathbf{z}^*)\| \geq \|\tilde{\mathbf{z}} - \mathbf{x}_0\| \sin \gamma \geq \|\mathbf{z}^* - \mathbf{x}_0\| \sin \gamma = \text{dist}(\mathbf{x}, S_0) \sin \gamma, \end{aligned} \quad (19)$$

where the first inequality follows from the facts that  $\langle \tilde{\mathbf{z}} - \mathbf{z}^*, \mathbf{z}^* \rangle = (\beta - 1)(\lambda^*)^2 \|\tilde{\mathbf{x}}\|^2 \geq 0$  and  $1 = \|\tilde{\mathbf{z}}\|^2 = \|\tilde{\mathbf{z}} - \mathbf{z}^*\|^2 + \|\mathbf{z}^*\|^2 + 2\langle \tilde{\mathbf{z}} - \mathbf{z}^*, \mathbf{z}^* \rangle \geq \|\tilde{\mathbf{z}} - \mathbf{z}^*\|^2 + \|\mathbf{z}^*\|^2$ . Next, we consider the case where  $\|\tilde{\mathbf{z}} - \mathbf{x}_0\| < \|\mathbf{z}^* - \mathbf{x}_0\|$ . To proceed, define by  $\theta \in (0, \pi/2)$  the angle between  $\mathbf{z}^* - \mathbf{x}_0$  and  $\tilde{\mathbf{x}} - \mathbf{x}_0$ . Since  $\tilde{\theta} \in (0, \pi/2)$ , and  $\cos \theta = \frac{\|\mathbf{x}_0 - \Pi_L(\mathbf{z}^*)\|}{\|\mathbf{z}^* - \mathbf{x}_0\|} < \frac{\|\mathbf{x}_0 - \Pi_L(\mathbf{z}^*)\|}{\|\tilde{\mathbf{z}} - \mathbf{x}_0\|} = \cos \tilde{\theta}$ , we see that  $\sin \theta > \sin \tilde{\theta}$ . Then, it holds that

$$\|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| = \|\mathbf{z}^* - \mathbf{x}_0\| \sin \theta > \|\tilde{\mathbf{z}} - \mathbf{x}_0\| \sin \tilde{\theta} \geq \text{dist}(\mathbf{x}, S_0) \sin \tilde{\theta} \geq \text{dist}(\mathbf{x}, S_0) \sin \gamma. \quad (20)$$

From (16), (19) and (20), we see that (11) holds.

**Case II:** If  $n \geq 3$ , we observe that  $\mathbf{u}$  can be eliminated from problem (18) by using the fact that  $\|\tilde{\mathbf{x}}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{x}_0\|^2 = 1$ . In particular, problem (18) can be rewritten as follows:

$$\begin{aligned} \min_{\mu, \lambda \in \mathbb{R}} \quad & \mu \\ \text{s.t.} \quad & \lambda^2 \|\tilde{\mathbf{x}}\|^2 + \mu^2 \|\mathbf{v}\|^2 \leq \|\mathbf{x}\|^2, \\ & (\lambda - 1)^2 \|\tilde{\mathbf{x}}\|^2 + (\mu - 1)^2 \|\mathbf{v}\|^2 \leq \text{dist}^2(\mathbf{x}, S_0), \\ & \|\mathbf{x}\|^2 + 1 - 2\lambda \|\tilde{\mathbf{x}}\|^2 - 2\mu \|\mathbf{v}\|^2 = \text{dist}^2(\mathbf{x}, S_0), \\ & \mu \in [0, 1], \end{aligned} \quad (21)$$

where the equality constraint comes from eliminating  $\|\mathbf{u}\|^2$  in the first two constraints in (18). Indeed, for any feasible solution  $(\lambda, \mu, \mathbf{u})$  to (18),  $(\lambda, \mu)$  is feasible to (21). Meanwhile, if  $(\lambda, \mu)$  is feasible to (21), since  $n \geq 3$ , one can always find  $\mathbf{u} \in \mathbb{R}^n$  satisfying  $\|\mathbf{u}\|^2 = \|\mathbf{x}\|^2 - (\lambda^2 \|\tilde{\mathbf{x}}\|^2 + \mu^2 \|\mathbf{v}\|^2) \geq 0$  and  $\langle \mathbf{u}, \tilde{\mathbf{x}} \rangle = 0$ ,  $\langle \mathbf{u}, \mathbf{v} \rangle = 0$  such that  $(\lambda, \mu, \mathbf{u})$  is a feasible solution to (18). Consequently,  $(\lambda^*, \mu^*, \mathbf{u}^*)$  is an optimal solution to (18) if and only if  $(\lambda^*, \mu^*)$  is an optimal solution to (21).

Let  $\mu^*$  be the optimal value of problem (21). Then,  $(\lambda^*, \mu^*)$  with  $\lambda^* = (1 + \|\mathbf{x}\|^2 - 2\mu^* \|\mathbf{v}\|^2 - \text{dist}^2(\mathbf{x}, S_0))/2\|\tilde{\mathbf{x}}\|^2$  is the unique optimal solution to problem (21). We consider three cases here:

- (i) Case  $\mu^* \in (0, 1)$ . Note that the first three constraints in (21) result a line segment, denoted by  $F$ , where its endpoints are the two intersection points (note that since  $\mathbf{x}$  is a feasible solution of (15) with  $\mathbf{u} = \mathbf{0}$ , the two ellipses must intersect) of the two ellipses:

$$\left\{ (\lambda, \mu) \mid \lambda^2 \|\tilde{\mathbf{x}}\|^2 + \mu^2 \|\mathbf{v}\|^2 = \|\mathbf{x}\|^2 \right\}$$

and

$$\left\{ (\lambda, \mu) \mid (\lambda - 1)^2 \|\tilde{\mathbf{x}}\|^2 + (\mu - 1)^2 \|\mathbf{v}\|^2 = \text{dist}^2(\mathbf{x}, S_0) \right\}.$$

Hence the feasible set to (21) can be written as  $F \cap \{(\lambda, \mu) \mid 0 \leq \mu \leq 1\}$ . Since  $\|\tilde{\mathbf{x}}\| \neq 0$ , the equality constraint in (21) implies that  $F$  cannot be parallel to the  $\lambda$ -axis. Now,  $\mu^* \in (0, 1)$  implies that the optimal solution to (21) must be an endpoint of  $F$ . This further implies that the first two inequality constraints in (21) are active at the optimal solution  $(\lambda^*, \mu^*)$ . Then,  $(\lambda^*, \mu^*, \mathbf{0})$  is an optimal solution to (18). Therefore,  $\mathbf{z}^* = \lambda^* \tilde{\mathbf{x}} + \mu^* \mathbf{v}$  is an optimal solution to problem (15). The desired result then follows from the same arguments in Case I.

(ii) Case  $\mu^* = 1$ . In this case, we have from (17) that  $\|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| = \text{dist}(\mathbf{x}, S_0)$ . Hence, we have

$$\sqrt{1 - \|\mathbf{x}\|^2} + \text{dist}(\mathbf{x}, S_1) \geq \sqrt{1 - \|\mathbf{x}\|^2} + \|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| \geq \text{dist}(\mathbf{x}, S_0) \geq \text{dist}(\mathbf{x}, S_0) \sin \gamma,$$

where the first inequality follows from (16).

(iii) Case  $\mu^* = 0$ . In this case,  $(\lambda^*, 0, \mathbf{u}^*)$  with some  $\mathbf{u}^*$  satisfying  $\|\mathbf{u}^*\|^2 = \|\mathbf{x}\|^2 - (\lambda^*)^2 \|\bar{\mathbf{x}}\|^2$  and  $\langle \mathbf{u}^*, \bar{\mathbf{x}} \rangle = 0$ ,  $\langle \mathbf{u}^*, \mathbf{v} \rangle = 0$  is an optimal solution to (18). Then,  $\mathbf{z}^* = \lambda^* \bar{\mathbf{x}} + \mathbf{u}^*$  is an optimal solution to (15) and  $\Pi_L(\mathbf{z}^*) = \bar{\mathbf{x}}$ . Let  $\tilde{\mathbf{z}} := \mathbf{z}^* + \beta \mathbf{v} = \lambda^* \bar{\mathbf{x}} + \mathbf{u}^* + \beta \mathbf{v}$  with  $\beta \geq 0$  such that  $\|\tilde{\mathbf{z}}\| = 1$ . Then, we see that

$$\|\tilde{\mathbf{z}} - \bar{\mathbf{x}}\|^2 = (\lambda^* - 1)^2 \|\bar{\mathbf{x}}\|^2 + \|\mathbf{u}^*\|^2 + \beta^2 \|\mathbf{v}\|^2 = 1 - 2\lambda^* \|\bar{\mathbf{x}}\|^2 + \|\bar{\mathbf{x}}\|^2, \quad (22)$$

where the second equality holds since  $\|\tilde{\mathbf{z}}\| = 1$ . Since  $\mathbf{x}_0 = \bar{\mathbf{x}} + \mathbf{v}$ , it holds from the last equality constraint in (15) that  $\text{dist}^2(\mathbf{x}, S_0) = \|\mathbf{z}^* - \mathbf{x}_0\|^2 = (\lambda^* - 1)^2 \|\bar{\mathbf{x}}\|^2 + \|\mathbf{u}^*\|^2 + \|\mathbf{v}\|^2$ . Then, we see that

$$\frac{\text{dist}^2(\mathbf{x}, S_0)}{\|\tilde{\mathbf{z}} - \bar{\mathbf{x}}\|^2} = \frac{(\lambda^* - 1)^2 \|\bar{\mathbf{x}}\|^2 + \|\mathbf{u}^*\|^2 + \|\mathbf{v}\|^2}{1 - 2\lambda^* \|\bar{\mathbf{x}}\|^2 + \|\bar{\mathbf{x}}\|^2} \quad (23)$$

$$= 1 + \frac{((\lambda^*)^2 - 1) \|\bar{\mathbf{x}}\|^2 + \|\mathbf{u}^*\|^2}{1 - 2\lambda^* \|\bar{\mathbf{x}}\|^2 + \|\bar{\mathbf{x}}\|^2} \quad (24)$$

$$\leq 1 + \frac{1 - \|\bar{\mathbf{x}}\|^2}{1 - 2\lambda^* \|\bar{\mathbf{x}}\|^2 + \|\bar{\mathbf{x}}\|^2} \quad (25)$$

$$\leq 1 + \frac{1 - \|\bar{\mathbf{x}}\|^2}{1 - 2\|\bar{\mathbf{x}}\| + \|\bar{\mathbf{x}}\|^2} \quad (26)$$

$$= \frac{2}{1 - \|\bar{\mathbf{x}}\|} = \frac{1}{(\sin \gamma)^2}, \quad (27)$$

where (23) follows from (22), (24) follows from  $\|\mathbf{x}_0\|^2 = \|\bar{\mathbf{x}}\|^2 + \|\mathbf{v}\|^2 = 1$ , (25) follows from  $\|\mathbf{z}^*\|^2 = \|\lambda^* \bar{\mathbf{x}}\|^2 + \|\mathbf{u}^*\|^2 \leq 1$  and (26) follows from  $\|\lambda^* \bar{\mathbf{x}}\| \leq 1$ . Meanwhile, it holds that

$$1 - \|\mathbf{z}^*\|^2 = 1 - (\lambda^*)^2 \|\bar{\mathbf{x}}\|^2 - \|\mathbf{u}^*\|^2 = \beta^2 \|\mathbf{v}\|^2 = \|\tilde{\mathbf{z}} - \mathbf{z}^*\|^2. \quad (28)$$

Hence, we have

$$\begin{aligned} \sqrt{1 - \|\mathbf{x}\|^2} + \text{dist}(\mathbf{x}, S_1) &\geq \sqrt{1 - \|\mathbf{z}^*\|^2} + \|\mathbf{z}^* - \Pi_L(\mathbf{z}^*)\| \\ &= \|\tilde{\mathbf{z}} - \mathbf{z}^*\| + \|\mathbf{z}^* - \bar{\mathbf{x}}\| \\ &\geq \|\tilde{\mathbf{z}} - \bar{\mathbf{x}}\| \\ &\geq \text{dist}(\mathbf{x}, S_0) \sin \gamma, \end{aligned}$$

where the first inequality follows from (16), the first equality follows from (28) and  $\Pi_L(\mathbf{z}^*) = \bar{\mathbf{x}}$ , and the last inequality follows from (27).

We have shown that (11) holds and thus completed the proof of Lemma 3.11.  $\square$

Note that Lemma 3.7 provides certain error bound inequality involving  $\text{dist}(\mathbf{x}, S_1)$  for the convex problem (P<sub>1</sub>) and Lemma 3.11 connects  $\text{dist}(\mathbf{x}, S_1)$  and  $\text{dist}(\mathbf{x}, S_0)$ . Using these results, we obtain in the following lemma the desired error bound result in the hard case 2 (ii).

**Lemma 3.12.** *In the hard case 2 (ii), there exists a constant  $\tau > 0$  such that for all  $\mathbf{x} \in \mathbb{R}^n$ ,*

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}.$$

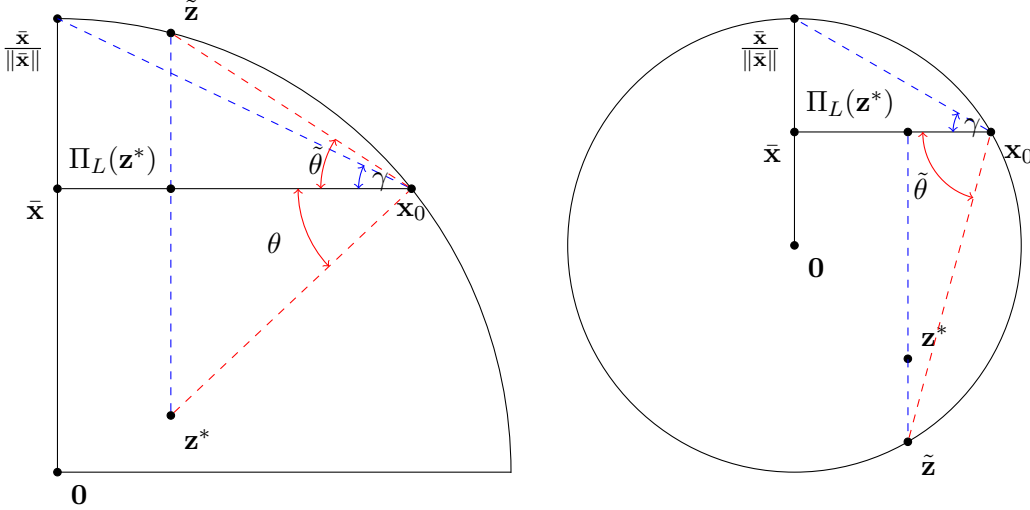


Figure 1: Illustration of two possible scenarios of the positions of  $\tilde{\mathbf{z}}$ . In either case, it holds that  $\tilde{\theta} \geq \gamma$ . In the first case, it also holds that  $\cos \theta \leq \cos \tilde{\theta}$  (or equivalently,  $\sin \theta \geq \sin \tilde{\theta}$ ).

Proof. Recall that  $\tilde{f}(\mathbf{x}) := f(\mathbf{x}) - \lambda_1(\mathbf{x}^T \mathbf{x} - 1)$  is convex. Note that in hard case 2 (ii), it holds that  $\min \tilde{f}(\mathbf{x}) = f^*$  and  $\|\bar{\mathbf{x}}\| < 1$ . Then, by Lemma 3.7 item 3, we know that there exist a constant  $\tau_1 > 0$  such that

$$\text{dist}(\mathbf{x}, S_1) \leq \tau_1 (\tilde{f}(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (29)$$

Without loss of generality, one can assume that  $\tau_1 \geq 1/\sqrt{-\lambda_1}$ , i.e.,  $2\tau_1^2\lambda_1 + 2 \leq 0$ . By Lemma 3.11, we know that

$$\text{dist}(\mathbf{x}, S_0) \sin \gamma \leq \text{dist}(\mathbf{x}, S_1) + \sqrt{1 - \|\mathbf{x}\|^2}, \quad \forall \|\mathbf{x}\| \leq 1,$$

where  $\sin \gamma = \frac{1 - \|\bar{\mathbf{x}}\|}{\sqrt{(1 - \|\bar{\mathbf{x}}\|)^2 + 1 - \|\bar{\mathbf{x}}\|^2}}$ . Together with (29), this implies that for all  $\|\mathbf{x}\| \leq 1$ ,

$$\begin{aligned} \text{dist}^2(\mathbf{x}, S_0) \sin^2 \gamma &\leq 2\text{dist}^2(\mathbf{x}, S_1) + 2(1 - \|\mathbf{x}\|^2) \\ &\leq 2\tau_1^2 (\tilde{f}(\mathbf{x}) - f^* + \delta_B(\mathbf{x})) + 2(1 - \|\mathbf{x}\|^2) \\ &= 2\tau_1^2 (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x})) + (2 + 2\tau_1^2\lambda_1)(1 - \|\mathbf{x}\|^2) \\ &\leq 2\tau_1^2 (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x})), \end{aligned}$$

where the last inequality holds since  $2 + 2\tau_1^2\lambda_1 \leq 0$ . Thus, we know that

$$\text{dist}(\mathbf{x}, S_0) \leq \frac{\sqrt{2}\tau_1}{\sin \gamma} (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

We complete the proof of this lemma.  $\square$

Now, with Lemmas 3.7, 3.8, 3.9 and 3.12, we are able to summarize the situations in the following *TRS-ill case* in which the Hölderian error bound modulus is 1/4:

$$\lambda_1 \leq 0, \quad \mathbf{b} \in \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I}) \quad \text{and} \quad \left\| (\mathbf{A} - \lambda_1 \mathbf{I})^\dagger \mathbf{b} \right\| = 1. \quad (30)$$

After all these preparations, we arrive at the following theorem.

**Theorem 3.13.** *For the trust region subproblem problem  $(P_0)$ , there exists a constant  $\tau_{EB} > 0$  such that*

$$\text{dist}(\mathbf{x}, S_0) \leq \tau_{EB} (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^\rho, \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

where  $\rho = \begin{cases} 1/4, & \text{for the TRS-ill case (30),} \\ 1/2, & \text{otherwise.} \end{cases}$

The above theory in fact shows that the Hölderian error bound always holds with modulus  $1/4$  for all cases due to the facts that the function value  $f(\mathbf{x})$  is bounded in the unit ball and the inequality  $t^{1/2} \leq t^{1/4}$  holds for all  $t \in (0, 1)$ .

## 4 KL inequality for the TRS

In this section, based on the previously developed error bound results, we derive the KL inequality associated with the TRS. To prove the KL inequality, let us first recall part of the results from [8] that essentially state that in the *convex* setting, Hölderian error bounds imply the KL inequality. In fact, the equivalence between these two concepts for *convex* problems is also obtained in [8].

**Lemma 4.1** (Corollary 6 (ii) in [8]). *Let  $H$  be a real Hilbert space. Set*

$$\mathcal{K}(0, +\infty) = \{\varphi \in C^0[0, +\infty) \cap C^1(0, +\infty), \varphi(0) = 0, \varphi \text{ is concave and } \varphi' > 0\}.$$

*Let  $h : H \rightarrow (-\infty, +\infty]$  be a proper, convex and lower-semicontinuous function, with  $\min h = 0$ . Let  $\varphi \in \mathcal{K}(0, +\infty), c > 0$ . Let  $S = \{x \in H \mid h(x) = \min h(x)\}$ .*

*Then if  $s\varphi'(s) \geq c\varphi(s)$  for all  $s > 0$ , and  $\varphi(h(x)) \geq \text{dist}(x, S)$  for all  $x \in \{x \in H \mid h(x) > 0\}$ , then  $\varphi'(h(x))\text{dist}(0, \partial h(x)) \geq c$  for all  $x \in \{x \in H \mid h(x) > 0\}$ .*

Recall that we denote by  $f^*$  the optimal value of the TRS ( $P_0$ ). With Lemma 4.1, we are ready to prove the KL inequality for the convex TRS. As one can observe from Theorem 4.2, the KL inequality for the convex TRS holds globally.

**Theorem 4.2.** *For the TRS ( $P_0$ ) with  $\lambda_1 \geq 0$ , there exists some constant  $\tau > 0$  such that the KL inequality holds*

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1-\rho} \leq \tau \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad (31)$$

where  $\rho = \frac{1}{4}$  if  $\lambda_1 = 0$ ,  $\mathbf{b} \in \text{Range}(\mathbf{A})$  and  $\|\mathbf{A}^\dagger \mathbf{b}\| = 1$ ; otherwise  $\rho = \frac{1}{2}$ .

*Proof.* By Theorem 3.13, we know that there exists a constant  $\tau > 0$  such that for all  $\mathbf{x} \in \mathbb{R}^n$

$$\text{dist}(\mathbf{x}, S_0) \leq \tau (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^\rho \quad (32)$$

with  $\rho = \frac{1}{4}$  if  $\lambda_1 = 0$ ,  $\mathbf{b} \in \text{Range}(\mathbf{A})$  and  $\|\mathbf{A}^\dagger \mathbf{b}\| = 1$ ; and  $\rho = \frac{1}{2}$  otherwise. Now define  $\varphi(s) = \tau s^\rho$  for all  $s \geq 0$ . Obviously,  $\varphi \in \mathcal{K}(0, +\infty)$  and we also note that

$$s\varphi'(s) = \tau \rho s^{\rho-1} = \tau \rho s^\rho = \rho \varphi(s), \quad \forall s > 0.$$

Hence, by (32), Lemma 4.1 and the fact that  $\text{dist}(0, \partial[f(\mathbf{x}) + \delta_B(\mathbf{x})]) = \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x}))$ , we have that for all  $\mathbf{x} \notin S_0$ ,

$$\tau \rho (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{\rho-1} \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})) \geq \rho.$$

Thus, by noting that (31) holds trivially if  $\mathbf{x} \in S_0$ , we complete the proof for the theorem.  $\square$

Next, we consider the nonconvex case. In this case, the analysis is more complicated. We note that in [16], the authors show that under the prox-regularity [53] assumption, for extended-real-valued lower semicontinuous functions, the error bound condition with modulus  $1/2$  (coined as the quadratic growth condition in [16]), implies the metric subregularity [14] property. However, for the TRS, the relations between the metric subregularity property and the KL inequality remain largely unknown. Hence, the results in [16] cannot be directly used here.

In the following, we give a comprehensive analysis of the KL inequality for the nonconvex TRS. For ease of reading, we first provide here a roadmap of the analysis. Apart from some general discussions and computations, our proof can be divided into mainly three steps. Specifically, in the first step, we establish the KL inequality for testing points chosen from the intersection of the interior of the unit norm ball and a neighborhood of the given optimal reference point. Then, we restrict our attentions on the testing points chosen from the boundary of the unit norm ball. We discuss the easy case and hard case 1 in the second step. The analysis for the hard case 2 is presented in the last step where certain parts of the proofs are presented in the Appendix.

We begin our analysis with some general discussions. Let  $\mathbf{x}^*$  be an optimal solution to the TRS ( $P_0$ ) with  $\lambda_1 < 0$ . We know from the KKT condition in Section 2 that there exists  $\lambda^* \geq 0$  such that

$$\nabla f(\mathbf{x}^*) + 2\lambda^* \mathbf{x}^* = 2\mathbf{A}\mathbf{x}^* - 2\mathbf{b} + 2\lambda^* \mathbf{x}^* = \mathbf{0}, \lambda^* \geq -\lambda_1 > 0, \lambda^*(1 - \|\mathbf{x}^*\|) = 0. \quad (33)$$

Thus,  $\|\mathbf{x}^*\| = 1$  and  $\|\nabla f(\mathbf{x}^*)\| = 2\lambda^*$ . For any given  $\mathbf{x} \in \mathbb{R}^n$ , since

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{x}^*)\| = \|2\mathbf{A}(\mathbf{x} - \mathbf{x}^*)\| \leq 2\|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{x}^*\|,$$

it holds that

$$\|\nabla f(\mathbf{x})\| \geq \|\nabla f(\mathbf{x}^*)\| - 2\|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{x}^*\|. \quad (34)$$

Meanwhile since  $f(\mathbf{x})$  is a quadratic function, we have

$$f(\mathbf{x}) - f(\mathbf{x}^*) = \langle \nabla f(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \langle \mathbf{x} - \mathbf{x}^*, \mathbf{A}(\mathbf{x} - \mathbf{x}^*) \rangle. \quad (35)$$

Now, for any  $\mathbf{x}$  satisfying  $\|\mathbf{x} - \mathbf{x}^*\| \leq \frac{\lambda^*}{\|\mathbf{A}\|_2 + \lambda^*}$  and  $\|\mathbf{x}\| \leq 1$ , we know from (34) and (35) that

$$\begin{cases} \|\nabla f(\mathbf{x})\| \geq 2\lambda^* - \frac{2\lambda^* \|\mathbf{A}\|_2}{\|\mathbf{A}\|_2 + \lambda^*} = \frac{2(\lambda^*)^2}{\|\mathbf{A}\|_2 + \lambda^*} = \frac{\lambda^*}{\|\mathbf{A}\|_2 + \lambda^*} \|\nabla f(\mathbf{x}^*)\|, \\ f(\mathbf{x}) - f(\mathbf{x}^*) \leq \frac{2(\lambda^*)^2}{\|\mathbf{A}\|_2 + \lambda^*} + \frac{\|\mathbf{A}\|_2 (\lambda^*)^2}{(\|\mathbf{A}\|_2 + \lambda^*)^2} \leq 2\lambda^* = \|\nabla f(\mathbf{x}^*)\|, \end{cases}$$

and consequently,

$$(f(\mathbf{x}) - f(\mathbf{x}^*))^{1/2} \leq \|\nabla f(\mathbf{x}^*)\|^{1/2} \leq \frac{\|\mathbf{A}\|_2 + \lambda^*}{\sqrt{2}(\lambda^*)^{3/2}} \|\nabla f(\mathbf{x})\|. \quad (36)$$

Then we have the following results associated with the case where the testing points are chosen from the intersection of the interior of the unit norm ball and a neighborhood of the given optimal reference point  $\mathbf{x}^*$ .

**Lemma 4.3.** *Suppose  $\lambda_1 < 0$ . Then there exist some constant  $\tau > 0$  and sufficiently small  $\epsilon > 0$  such that*

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2} \leq \tau \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \epsilon) \cap \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| < 1\}.$$

*Proof.* Note that for all  $\mathbf{x}$  satisfying  $\|\mathbf{x}\| < 1$ , it holds that  $N_B(\mathbf{x}) = \{\mathbf{0}\}$  and  $\text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})) = \|\nabla f(\mathbf{x})\|$ . Hence, we know from (36) that

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2} \leq \frac{\|\mathbf{A}\|_2 + \lambda^*}{\sqrt{2}(\lambda^*)^{3/2}} \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})),$$

whenever  $\|\mathbf{x} - \mathbf{x}^*\| \leq \frac{\lambda^*}{\|\mathbf{A}\|_2 + \lambda^*}$  and  $\|\mathbf{x}\| < 1$ . □

In the subsequent discussions, we will focus on the boundary of the unit norm ball, denoted by  $\mathbf{bd}(B) := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$ . For all  $\mathbf{x} \in \mathbf{bd}(B)$ , it follows from (35) that

$$f(\mathbf{x}) - f(\mathbf{x}^*) = -\langle 2\lambda^* \mathbf{x}^*, \mathbf{x} - \mathbf{x}^* \rangle + \langle \mathbf{x} - \mathbf{x}^*, \mathbf{A}(\mathbf{x} - \mathbf{x}^*) \rangle = \langle \mathbf{x} - \mathbf{x}^*, (\mathbf{A} + \lambda^* \mathbf{I})(\mathbf{x} - \mathbf{x}^*) \rangle, \quad (37)$$



where the second equality holds since  $\|\mathbf{x}\| = \|\mathbf{x}^*\| = 1$ . Meanwhile, for any given  $\mathbf{x} \in \mathbf{bd}(B)$ , we have  $N_B(\mathbf{x}) = \{\nu\mathbf{x} \in \mathbb{R}^n \mid \nu \geq 0\}$  and thus

$$\text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})) = \min_{\nu \geq 0} \|2(\mathbf{A}\mathbf{x} - \mathbf{b}) + \nu\mathbf{x}\|. \quad (38)$$

Let  $\nu^*(\mathbf{x})$  be the optimal solution to problem (38). It can be verified that  $\nu(\mathbf{x}) = \max\{\langle -2(\mathbf{A}\mathbf{x} - \mathbf{b}), \mathbf{x} \rangle, 0\}$ . Moreover, it holds that

$$\nu(\mathbf{x}) \rightarrow \nu(\mathbf{x}^*) = \langle -2(\mathbf{A}\mathbf{x}^* - \mathbf{b}), \mathbf{x}^* \rangle = 2\lambda^* > 0 \text{ as } \mathbf{bd}(B) \ni \mathbf{x} \rightarrow \mathbf{x}^*. \quad (39)$$

Hence, we know that there exists a constant  $\epsilon_0 \in (0, 1)$  such that

$$\nu(\mathbf{x}) = \langle -2(\mathbf{A}\mathbf{x} - \mathbf{b}), \mathbf{x} \rangle > 0, \quad \forall \mathbf{x} \in \mathbf{bd}(B) \cap \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_0\}. \quad (40)$$

Now, we are ready to prove the desired KL inequality for the easy case and hard case 1.

**Lemma 4.4** (KL inequality for the easy case and hard case 1). *Suppose  $\lambda_1 < 0$ . In the easy case and hard case 1, there exist some constant  $\tau > 0$  and sufficiently small  $\epsilon > 0$  such that*

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2} \leq \tau \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \epsilon). \quad (41)$$

*Proof.* Note that in the easy case and hard case 1, it holds that  $\lambda^* > -\lambda_1 > 0$  and  $\|(\mathbf{A} - \lambda_1\mathbf{I})^\dagger\mathbf{b}\| \neq 1$ . Since  $\lambda^* > -\lambda_1 > 0$ , from the discussion of (39), we know that there exists a positive constant  $\epsilon_1 \leq \epsilon_0$  such that  $\nu(\mathbf{x}) + 2\lambda_1 > 0$  whenever  $\|\mathbf{x}\| = 1$  and  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1$ . Recall the objective function  $\tilde{f}$  in the convex problem (P<sub>1</sub>), i.e.,  $\tilde{f}(\mathbf{x}) := \mathbf{x}^T(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} - 2\mathbf{b}^T\mathbf{x} + \lambda_1$ . We note for all  $\mathbf{x}$  satisfying  $\|\mathbf{x}\| = 1$  and  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1$  that

$$\begin{aligned} \text{dist}(-\nabla \tilde{f}(\mathbf{x}), N_B(\mathbf{x})) &= \min_{\mu \geq 0} \|2(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} - 2\mathbf{b} + \mu\mathbf{x}\| \\ &= \|2(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} - 2\mathbf{b} + (-\langle 2(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} - 2\mathbf{b}, \mathbf{x} \rangle)\mathbf{x}\| \\ &= \|2(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} - 2\mathbf{b} + (\nu(\mathbf{x}) + 2\lambda_1)\mathbf{x}\| \\ &= \|2(\mathbf{A}\mathbf{x} - \mathbf{b}) + \nu(\mathbf{x})\mathbf{x}\| \\ &= \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \end{aligned}$$

where the second equality is due to  $-\langle 2(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} - 2\mathbf{b}, \mathbf{x} \rangle = \nu(\mathbf{x}) + 2\lambda_1 > 0$ . Note that  $\tilde{f}(\mathbf{x}) = f(\mathbf{x})$ ,  $\forall \mathbf{x} \in \mathbf{bd}(B)$  and from Lemma 2.1 that  $f^*$ , the optimal value of the TRS (P<sub>0</sub>), is also the optimal value of the convex problem (P<sub>1</sub>). Since  $\|(\mathbf{A} - \lambda_1\mathbf{I})^\dagger\mathbf{b}\| \neq 1$ , we obtain by Theorem 4.2 that there exists a positive constant  $\tau_1$  such that for all  $\mathbf{x} \in \mathbf{bd}(B)$  with  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1$ ,

$$\begin{aligned} (f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2} &= (\tilde{f}(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2} \\ &\leq \tau_1 \text{dist}(-\nabla \tilde{f}(\mathbf{x}), N_B(\mathbf{x})) = \tau_1 \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})). \end{aligned}$$

This, together with Lemma 4.3, implies that (41) holds with some positive constant  $\epsilon$  and  $\tau$ .  $\square$

Now, we are ready for presenting the analysis for the hard case 2. Recall  $\lambda^* = -\lambda_1$  in the hard case. First, we discuss the case where  $\mathbf{A} - \lambda_1\mathbf{I} = \mathbf{0}$ . In this case, we know from the optimality condition (33) that  $\mathbf{b} = \mathbf{0}$  and thus  $f(\mathbf{x}^*) = f(\mathbf{x}) = \lambda_1\mathbf{x}^T\mathbf{x} = \lambda_1$  whenever  $\|\mathbf{x}\| = 1$ . This, together with Lemma 4.3, implies that the KL inequality holds with the exponent 1/2. Note that this case belongs to the hard case 2 (ii) as  $\mathbf{b} \perp \text{Null}(\mathbf{A} - \lambda_1\mathbf{I})$  and  $\|(\mathbf{A} - \lambda_1\mathbf{I})^\dagger\mathbf{b}\| = 0 < 1$ .

Hence, in the subsequent analysis, we focus on the nontrivial case where  $\mathbf{A} - \lambda_1\mathbf{I} \neq \mathbf{0}$ . Recall the spectral decomposition of  $\mathbf{A}$  as  $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T$  in Section 2. Then,  $\mathbf{A} - \lambda_1\mathbf{I} = \mathbf{P}(\mathbf{\Lambda} - \lambda_1\mathbf{I})\mathbf{P}^T$  is the spectral

decomposition of  $\mathbf{A} - \lambda_1 \mathbf{I}$ . Denote the diagonal matrix  $\mathbf{D} = \mathbf{A} - \lambda_1 \mathbf{I} = \text{Diag}(\mathbf{d})$ . Since  $\mathbf{A} - \lambda_1 \mathbf{I} \neq \mathbf{0}$ , there exists some integer  $1 \leq K < n$  such that

$$\begin{cases} \mathbf{d}_i = 0 & \text{for } i = 1, \dots, K, \\ \mathbf{d}_i = \lambda_i - \lambda_1 & \text{for } i = K + 1, \dots, n, \end{cases} \quad (42)$$

and  $\mathbf{d}_n \geq \dots \geq \mathbf{d}_{K+1} > 0$ . For all  $\mathbf{x} \in \mathbf{bd}(B)$ , we define

$$\alpha(\mathbf{x}) := \langle \mathbf{x}^*, (\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*) \rangle \text{ and } \beta(\mathbf{x}) := \langle \mathbf{x} - \mathbf{x}^*, (\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*) \rangle. \quad (43)$$

For simplicity in the following discussion, we sometimes suppress the dependence on  $\mathbf{x}$  in our notation. For example, we often write  $\alpha(\mathbf{x})$  and  $\beta(\mathbf{x})$  as  $\alpha$  and  $\beta$ , respectively. However, this should not cause any confusion because the reference vector  $\mathbf{x}$  will always be clear from the context. Let  $\mathbf{s} = \mathbf{P}^T \mathbf{x}^*$  and  $\mathbf{z} = \mathbf{P}^T(\mathbf{x} - \mathbf{x}^*)$  and we verify from (43) that

$$\begin{cases} \alpha = \langle \mathbf{x}^*, (\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*) \rangle & = \langle \mathbf{s}, \mathbf{Dz} \rangle, \\ \beta = \langle \mathbf{x} - \mathbf{x}^*, (\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*) \rangle \stackrel{(37)}{=} f(\mathbf{x}) - f^* & = \langle \mathbf{z}, \mathbf{Dz} \rangle. \end{cases} \quad (44)$$

Recall the definition of  $\epsilon_0$  in (40). It holds that for any  $\mathbf{x} \in \mathbf{bd}(B)$  with  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_0$ ,

$$\begin{aligned} & \text{dist}^2(-\nabla f(\mathbf{x}), N_B(\mathbf{x})) \\ &= \|2(\mathbf{Ax} - \mathbf{b}) - 2\lambda_1 \mathbf{x} + (2\lambda_1 + \nu(\mathbf{x}))\mathbf{x}\|^2 \\ &= \|2(\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*) + 2[(\mathbf{x}^*)^T(\mathbf{Ax}^* - \mathbf{b}) - \mathbf{x}^T(\mathbf{Ax} - \mathbf{b})]\mathbf{x}\|^2 \\ &= \|2(\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*) + 2[f(\mathbf{x}^*) - f(\mathbf{x}) + (\mathbf{x}^* - \mathbf{x})^T(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{x}^*]\mathbf{x}\|^2 \\ &= \|2(\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*) - 2(\beta + \alpha)\mathbf{x}\|^2 \\ &= \|2(\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*)\|^2 - 8\mathbf{x}^T(\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*)(\beta + \alpha) + \|2\beta\mathbf{x} + 2\alpha\mathbf{x}\|^2 \\ &= \|2(\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*)\|^2 - 4(\beta + \alpha)^2 \\ &= 4(\|\mathbf{Dz}\|^2 - (\alpha + \beta)^2), \end{aligned} \quad (45)$$

where the second equality follows from the fact that  $2\mathbf{Ax}^* - 2\mathbf{b} - 2\lambda_1 \mathbf{x}^* = \mathbf{0}$  (which is (33) with  $\lambda^* = -\lambda_1$ ) and the definition of  $\nu(\mathbf{x})$ , the third equality follows from  $2\mathbf{Ax}^* - 2\mathbf{b} - 2\lambda_1 \mathbf{x}^* = \mathbf{0}$ , (37) and  $\|\mathbf{x}\| = \|\mathbf{x}^*\| = 1$ , the fourth equality holds due to (43), the sixth equality holds since  $\mathbf{x}^T(\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{x} - \mathbf{x}^*) = \alpha + \beta$  and  $\|\mathbf{x}\| = 1$ , and the last equality holds as  $\mathbf{z} = \mathbf{P}^T(\mathbf{x} - \mathbf{x}^*)$ .

Define function  $H : \mathbf{bd}(B) \rightarrow \mathbb{R}$  as:

$$H(\mathbf{x}) := \|(\mathbf{Ax} - \mathbf{b}) - \lambda_1 \mathbf{x} + (\lambda_1 + \nu(\mathbf{x})/2)\mathbf{x}\|^2 \stackrel{(45)}{=} \|\mathbf{Dz} - (\alpha + \beta)\mathbf{x}\|^2 = \|\mathbf{Dz}\|^2 - (\alpha + \beta)^2, \quad (46)$$

and thus we have

$$\text{dist}^2(-\nabla f(\mathbf{x}), N_B(\mathbf{x})) = 4H(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbf{bd}(B) \cap \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_0\}. \quad (47)$$

In the next lemma, we show that the KL inequality holds with an exponent 1/2 for the hard case 2 (ii).

**Lemma 4.5** (KL inequality for the hard case 2 (ii)). *Suppose that  $\lambda_1 < 0$ . In the hard case 2 (ii), there exist some constant  $\tau > 0$  and sufficiently small  $\epsilon > 0$  such that*

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2} \leq \tau \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \epsilon).$$

*Proof.* We have discussed that the results hold under the case  $\mathbf{A} - \lambda_1 \mathbf{I} = \mathbf{0}$ . In the remaining, we consider the case where  $\mathbf{A} - \lambda_1 \mathbf{I} \neq \mathbf{0}$ . In the hard case 2 (ii), since  $\mathbf{x}^* \notin \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$ , i.e.,  $\mathbf{s} \notin \text{Null}(\mathbf{D})$ ,

we know that  $\sum_{j=1}^K \mathbf{s}_j^2 > 0$ , and from Lemma 7.1 in Appendix that there exists  $0 < \epsilon' < \epsilon_0$  such that for any  $\mathbf{x} \in \mathbf{bd}(B)$  with  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon'$ ,

$$H(\mathbf{x}) \geq \frac{1}{2} \mathbf{d}_{K+1} \left( \sum_{j=1}^K \mathbf{s}_j^2 \right) \beta,$$

i.e.,

$$\text{dist}^2(-\nabla f(\mathbf{x}), N_B(\mathbf{x})) = 4H(\mathbf{x}) \geq 2\mathbf{d}_{K+1} \left( \sum_{j=1}^K \mathbf{s}_j^2 \right) (f(\mathbf{x}) - f^*).$$

This, together with Lemma 4.3, completes the proof.  $\square$

Before discussing the general results for the hard case 2 (i), we state in the following lemma that in the hard case 2 (i) if the test point  $\mathbf{x}$  is restricted in the subspace  $\text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$ , the desired KL inequality can be derived by combining Lemma 4.4 with a reduction strategy. To avoid the tedious reduction arguments, we present the proof for the lemma in the Appendix.

**Lemma 4.6.** *Suppose  $\lambda_1 < 0$ . In the hard case 2 (i), there exist some constant  $\tau > 0$  and sufficiently small  $\epsilon > 0$  such that*

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{1/2} \leq \tau \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \epsilon) \cap \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I}).$$

After these preparations, we are ready to show in the next lemma that the KL inequality holds with the exponent  $3/4$  for the hard case 2 (i), which is a subcase of the TRS-ill case (30).

**Lemma 4.7** (KL inequality for the hard case 2 (i)). *Suppose that  $\lambda_1 < 0$ . In the hard case 2 (i), there exist some constant  $\tau > 0$  and sufficiently small  $\epsilon > 0$  such that*

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^{3/4} \leq \tau \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \epsilon). \quad (48)$$

*Proof.* Note that in the hard case 2 (i), it holds that  $\mathbf{x}^* \perp \text{Null}(\mathbf{A} - \lambda_1 \mathbf{I})$  (or equivalently,  $\mathbf{x}^* \in \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$  or  $\mathbf{s} \in \text{Range}(\mathbf{D})$ ). Recall the definition of  $\epsilon_0$  in (40) and denote  $\epsilon_1 := \min\{\epsilon_0, 1/2\}$ . For any  $\mathbf{x} \in \mathbf{bd}(B)$  with  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1$ , we know from  $\mathbf{z} = \mathbf{P}^T(\mathbf{x} - \mathbf{x}^*)$  that  $\gamma := \|\mathbf{z}\|^2 \leq \epsilon_1^2 \leq 1/4$ . Recall the definitions of  $\alpha$  and  $\beta$  in (43). In the following, we consider the following two cases, i.e., **Case I:** with  $|\alpha + \beta| \leq \frac{1}{2} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2}$  and **Case II:** with  $|\alpha + \beta| > \frac{1}{2} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2}$ .

**Case I.** If  $|\alpha + \beta| \leq \frac{1}{2} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2}$ , from (42), (44) and (46), we have

$$H(\mathbf{x}) = \|\mathbf{Dz}\|^2 - (\alpha + \beta)^2 \geq \mathbf{d}_{K+1} \beta - \frac{1}{4} \mathbf{d}_{K+1} \beta = \frac{3}{4} \mathbf{d}_{K+1} \beta, \quad \forall \mathbf{x} \in \mathbf{bd}(B) \text{ with } \|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1. \quad (49)$$

**Case II.** Now suppose  $|\alpha + \beta| > \frac{1}{2} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2}$ . Note that  $\alpha = \langle \mathbf{s}, \mathbf{Dz} \rangle \rightarrow 0$  and  $\beta = \langle \mathbf{z}, \mathbf{Dz} \rangle \rightarrow 0$  as  $\mathbf{z} \rightarrow \mathbf{0}$  or equivalently  $\mathbf{x} \rightarrow \mathbf{x}^*$ . Then, by reducing  $\epsilon_0$  if necessary, we have  $|\alpha| \geq \beta$  and thus

$$|\alpha| \geq \frac{1}{4} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2} \quad \forall \mathbf{x} \in \mathbf{bd}(B) \text{ with } \|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1. \quad (50)$$

Then if  $\alpha \leq 0$ , we know that for any  $\mathbf{x} \in \mathbf{bd}(B)$  with  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1$ ,

$$H(\mathbf{x}) = \|\mathbf{Dz}\|^2 - \alpha^2 - 2\alpha\beta - \beta^2 \geq 0 - 2\alpha\beta + \alpha\beta = -\alpha\beta \geq \frac{1}{4} \sqrt{\mathbf{d}_{K+1}} \beta^{3/2}, \quad (51)$$

where the first inequality holds as  $\|\mathbf{Dz}\|^2 - \alpha^2 \geq (\mathbf{s}^T \mathbf{Dz})^2 - \alpha^2 = 0$  and  $|\alpha| \geq \beta$ .

Next consider the case with  $\alpha > 0$ . Recall that in the hard case 2 (i), we have  $\mathbf{s} \in \text{Range}(\mathbf{D})$ . Hence, by reducing  $\epsilon_0$  if necessary, we know from Lemma 4.6, (44), (46) and (47) that there exists some constant  $c_1 \in (0, 1)$  such that for any  $\mathbf{x} \in \mathbf{bd}(B) \cap \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$  with  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1$ ,

$$\frac{1}{4} \text{dist}^2(-\nabla f(\mathbf{x}), N_B(\mathbf{x})) = H(\mathbf{x}) = \|\mathbf{Dz} - (\alpha + \beta)(\mathbf{s} + \mathbf{z})\|^2 \geq c_1^2 \mathbf{d}_{K+1} \beta. \quad (52)$$

For any given  $\mathbf{x} \in \mathbf{bd}(B)$  satisfying  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1$  and  $\mathbf{x} \notin \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$ , consider the orthogonal decomposition of  $\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2$  such that  $\mathbf{z}_1 \in \text{Range}(\mathbf{D})$  and  $\mathbf{z}_2 \in \text{Null}(\mathbf{D})$ . Since  $\mathbf{x} \notin \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$ ,  $\mathbf{x}^* \in \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$  and  $\mathbf{z} = P^T(\mathbf{x} - \mathbf{x}^*)$ , it holds that  $\mathbf{z} \notin \text{Range}(D)$ . Moreover, the fact that  $\alpha = \langle \mathbf{s}, \mathbf{D}\mathbf{z} \rangle = \langle \mathbf{s}, \mathbf{D}\mathbf{z}_1 \rangle > 0$  further implies  $\mathbf{z} \notin \text{Null}(D)$ . Hence, we have that  $\mathbf{z}_1 \neq \mathbf{0}$  and  $\mathbf{z}_2 \neq \mathbf{0}$ . Let

$$c = \frac{\mathbf{d}_{K+1}c_1}{16\sqrt{\mathbf{d}_n^2 + \mathbf{d}_{K+1}^2}} < \frac{c_1}{16} < \frac{1}{16}, \quad (53)$$

where  $c_1$  is the constant in (52). We consider two subcases:

1. Suppose that  $\|\mathbf{z}_2\| < c\sqrt{\gamma}$ . Then, we have

$$\beta = \langle \mathbf{D}\mathbf{z}, \mathbf{z} \rangle = \langle \mathbf{D}\mathbf{z}_1, \mathbf{z}_1 \rangle \geq \mathbf{d}_{K+1}\|\mathbf{z}_1\|^2 \geq \mathbf{d}_{K+1}(1 - c^2)\gamma \geq \frac{\mathbf{d}_{K+1}}{4}\gamma. \quad (54)$$

In our proof, the dimension of subspace  $\text{Range}(\mathbf{D})$  needs to be taken into consideration. Specifically, we use different strategies to handle cases with  $\dim(\text{Range}(\mathbf{D})) \leq 2$  and  $\dim(\text{Range}(\mathbf{D})) \geq 3$  (which implicitly requires  $n \geq 4$ ). For ease of reading, we defer the proof for the case with  $\dim(\text{Range}(\mathbf{D})) \leq 2$  to Appendix 7.3.

Now we focus on the case with  $\dim(\text{Range}(\mathbf{D})) \geq 3$ . Let  $\tilde{\mathbf{z}} = \mathbf{z}_1 + \mathbf{z}_3$  with  $\mathbf{z}_3 \in \text{Range}(\mathbf{D})$ ,  $\mathbf{z}_3 \perp \mathbf{s}$ ,  $\mathbf{z}_3 \perp \mathbf{z}_1$ ,  $\|\mathbf{z}_3\| = \|\mathbf{z}_2\|$  and  $\langle \mathbf{z}_3, \mathbf{D}\mathbf{z}_1 \rangle \leq 0$ . Note that  $\mathbf{z}_3$ , and consequently  $\tilde{\mathbf{z}}$ , is well-defined because  $\dim(\text{Range}(\mathbf{D})) \geq 3$ . From  $\tilde{\mathbf{z}}$ , we can further construct  $\tilde{\mathbf{x}}$ , and the corresponding  $\tilde{\alpha}$  and  $\tilde{\beta}$  as in (44), in the following way:

$$\tilde{\mathbf{x}} = \mathbf{x}^* + \mathbf{P}\tilde{\mathbf{z}}, \quad \tilde{\alpha} = \mathbf{s}^T \mathbf{D}\tilde{\mathbf{z}} \quad \text{and} \quad \tilde{\beta} = \tilde{\mathbf{z}}^T \mathbf{D}\tilde{\mathbf{z}}.$$

Since  $\|\mathbf{s}\| = 1$ ,  $\mathbf{D}\mathbf{z}_2 = 0$ ,  $\langle \mathbf{z}_3, \mathbf{D}\mathbf{z}_1 \rangle \leq 0$  and  $\|\mathbf{z}_3\| = \|\mathbf{z}_2\|$ , we have

$$|\tilde{\alpha} - \alpha| = |\mathbf{s}^T \mathbf{D}(\tilde{\mathbf{z}} - \mathbf{z})| = |\mathbf{s}^T \mathbf{D}(\mathbf{z}_3 - \mathbf{z}_2)| = |\mathbf{s}^T \mathbf{D}\mathbf{z}_3| \leq \|\mathbf{s}\| \|\mathbf{D}\mathbf{z}_3\| \leq \mathbf{d}_n \|\mathbf{z}_2\|, \quad (55)$$

and

$$|\tilde{\beta} - \beta| = 2\langle \mathbf{z}_3, \mathbf{D}\mathbf{z}_1 \rangle + \langle \mathbf{z}_3, \mathbf{D}\mathbf{z}_3 \rangle \leq \mathbf{d}_n \|\mathbf{z}_3\|_2^2 = \mathbf{d}_n \|\mathbf{z}_2\|_2^2 \leq \mathbf{d}_n c^2 \gamma \stackrel{(54)}{\leq} \frac{4\mathbf{d}_n c^2}{\mathbf{d}_{K+1}} \beta.$$

Meanwhile, we have from (53) that  $4\mathbf{d}_n c^2 / \mathbf{d}_{K+1} \leq 1/4$ , and consequently,

$$\frac{1}{4}\beta \leq \left(1 - \frac{4\mathbf{d}_n c^2}{\mathbf{d}_{K+1}}\right)\beta \leq \tilde{\beta} \leq \left(1 + \frac{4\mathbf{d}_n c^2}{\mathbf{d}_{K+1}}\right)\beta \leq 2\beta. \quad (56)$$

Furthermore, it holds that  $\|\tilde{\mathbf{x}} - \mathbf{x}^*\|^2 = \|\tilde{\mathbf{z}}\|^2 = \|\mathbf{z}\|^2 = \gamma$  and  $\|\tilde{\mathbf{x}}\|^2 = \|\mathbf{s} + \tilde{\mathbf{z}}\|^2 = \|\mathbf{s} + \mathbf{z}\|^2 = \|\mathbf{x}\|^2 = 1$ . Since  $\tilde{\mathbf{x}} \in \mathbf{bd}(B) \cap \text{Range}(\mathbf{A} - \lambda_1 \mathbf{I})$  and  $\|\tilde{\mathbf{x}} - \mathbf{x}^*\| = \sqrt{\gamma} \leq \epsilon_1$ , we know from (52) that

$$\frac{1}{2} \text{dist}(-\nabla f(\tilde{\mathbf{x}}), N_B(\tilde{\mathbf{x}})) = \|\mathbf{D}\tilde{\mathbf{z}} - (\tilde{\alpha} + \tilde{\beta})(\mathbf{s} + \tilde{\mathbf{z}})\| \geq c_1 \sqrt{\mathbf{d}_{K+1}} \tilde{\beta}^{1/2}. \quad (57)$$

Next, we bound the difference between  $\text{dist}(-\nabla f(\tilde{\mathbf{x}}), N_B(\tilde{\mathbf{x}}))/2$  and  $\text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x}))/2$  by

$$\begin{aligned} & \|[\mathbf{D}\tilde{\mathbf{z}} - (\tilde{\alpha} + \tilde{\beta})(\mathbf{s} + \tilde{\mathbf{z}})] - [\mathbf{D}\mathbf{z} - (\alpha + \beta)(\mathbf{s} + \mathbf{z})]\| \\ &= \|\mathbf{D}(\tilde{\mathbf{z}} - \mathbf{z}) - (\tilde{\alpha} - \alpha)\mathbf{s} - \tilde{\alpha}\tilde{\mathbf{z}} + \alpha\mathbf{z} - \tilde{\beta}(\mathbf{s} + \tilde{\mathbf{z}}) + \beta(\mathbf{s} + \mathbf{z})\| \\ &\leq \|\mathbf{D}\mathbf{z}_3\| + |\tilde{\alpha} - \alpha| + |\tilde{\alpha}|\|\tilde{\mathbf{z}}\| + |\alpha|\|\mathbf{z}\| + \tilde{\beta} + \beta \\ &\leq 2\mathbf{d}_n \|\mathbf{z}_2\| + 2\mathbf{d}_n \gamma + 3\beta, \end{aligned} \quad (58)$$

where the first inequality follows from  $\mathbf{D}(\tilde{\mathbf{z}} - \mathbf{z}) = \mathbf{D}(\mathbf{z}_3 - \mathbf{z}_2) = \mathbf{D}\mathbf{z}_3$ ,  $\|\mathbf{s}\| = 1$  and  $\|\mathbf{s} + \tilde{\mathbf{z}}\| = \|\mathbf{s} + \mathbf{z}\| = 1$ , and the last inequality follows from  $\|\mathbf{D}\mathbf{z}_3\| \leq \mathbf{d}_n \|\mathbf{z}_3\| = \mathbf{d}_n \|\mathbf{z}_2\|$ , (55),  $|\alpha|\|\mathbf{z}\| =$

$|\mathbf{s}^T \mathbf{Dz}| \|\mathbf{z}\| \leq \mathbf{d}_n \|\mathbf{z}\|^2 = \mathbf{d}_n \gamma$ ,  $|\tilde{\alpha}| \|\tilde{\mathbf{z}}\| = |\mathbf{s}^T \mathbf{D}\tilde{\mathbf{z}}| \|\tilde{\mathbf{z}}\| \leq \mathbf{d}_n \|\tilde{\mathbf{z}}\|^2 = \mathbf{d}_n \gamma$  and (56). The right-hand-side of (58) can be further bounded by

$$\begin{aligned}
2\mathbf{d}_n \|\mathbf{z}_2\| + 2\mathbf{d}_n \gamma + 3\beta &\stackrel{(54)}{\leq} 2\mathbf{d}_n \|\mathbf{z}_2\| + \left(3 + \frac{8\mathbf{d}_n}{\mathbf{d}_{K+1}}\right)\beta \\
&\leq 2\mathbf{d}_n c\sqrt{\gamma} + \left(3 + \frac{8\mathbf{d}_n}{\mathbf{d}_{K+1}}\right)\beta \\
&\stackrel{(53)}{\leq} \frac{c_1}{8} \mathbf{d}_{K+1} \sqrt{\gamma} + \left(3 + \frac{8\mathbf{d}_n}{\mathbf{d}_{K+1}}\right)\beta \\
&\stackrel{(54)}{\leq} \frac{c_1}{4} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2} + \left(3 + \frac{8\mathbf{d}_n}{\mathbf{d}_{K+1}}\right)\beta \\
&= \left(\frac{c_1}{4} \sqrt{\mathbf{d}_{K+1}} + \left(3 + \frac{8\mathbf{d}_n}{\mathbf{d}_{K+1}}\right)\beta^{1/2}\right)\beta^{1/2}.
\end{aligned}$$

Since  $\beta \rightarrow 0$  as  $\mathbf{x} \rightarrow \mathbf{x}^*$ , by reducing  $\epsilon_0$  if necessary, we have

$$\left(3 + \frac{8\mathbf{d}_n}{\mathbf{d}_{K+1}}\right)\beta^{1/2} \leq \frac{c_1}{8} \sqrt{\mathbf{d}_{K+1}}$$

and consequently,

$$2\mathbf{d}_n \|\mathbf{z}_2\| + 2\mathbf{d}_n \gamma + 3\beta \leq \frac{3c_1}{8} \sqrt{\mathbf{d}_{K+1}}. \quad (59)$$

Therefore, we know that

$$\begin{aligned}
H(\mathbf{x}) &= \|\mathbf{Dz} - (\alpha + \beta)(\mathbf{s} + \mathbf{z})\| \\
&\geq c_1 \sqrt{\mathbf{d}_{K+1}} \tilde{\beta}^{1/2} - 2\mathbf{d}_n \|\mathbf{z}_2\| - 2\mathbf{d}_n \gamma - 3\beta \\
&\geq \frac{c_1}{2} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2} - \frac{3c_1}{8} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2} \\
&\geq \frac{c_1}{8} \sqrt{\mathbf{d}_{K+1}} \beta^{1/2},
\end{aligned} \quad (60)$$

where the first inequality is due to (57) and (58), the second equality is due to (56) and (59).

2. On the other hand, suppose that  $\|\mathbf{z}_2\| \geq c\sqrt{\gamma}$ . Then,  $\|\mathbf{z}_1\| \leq \sqrt{(1-c^2)\gamma}$ . Let  $\mathbf{z}' = \rho\mathbf{z}_1$  and  $\mathbf{x}' = \mathbf{x}^* + \mathbf{Pz}' = \mathbf{s} + \mathbf{z}'$ , where  $\rho = \gamma/\|\mathbf{z}_1\|^2 \geq 1/\sqrt{1-c^2} > 1$ . Then we have

$$\|\mathbf{x}'\| = \|\mathbf{s} + \mathbf{z}'\|^2 = 1 + 2\mathbf{s}^T \mathbf{z}' + \|\mathbf{z}'\|^2 = 1 - \rho\gamma + \rho^2 \|\mathbf{z}_1\|^2 = 1,$$

where the third equality is due to

$$2\mathbf{s}^T \mathbf{z}' = 2\rho\mathbf{s}^T \mathbf{z}_1 = 2\rho\mathbf{s}^T \mathbf{z} = \rho(\|\mathbf{s} + \mathbf{z}\|^2 - \|\mathbf{s}\|^2 - \|\mathbf{z}\|^2) = \rho(1 - 1 - \gamma) = -\rho\gamma.$$

Similar as in (44), define  $\alpha'$  and  $\beta'$  corresponding to  $\mathbf{x}'$  as  $\alpha' = \mathbf{s}^T \mathbf{Dz}' = \rho\alpha$  and  $\beta' = \mathbf{z}'^T \mathbf{Dz}' = \rho^2\beta$ . Then, recalling the definition of  $H$  in (46), we have

$$\begin{aligned}
\frac{H(\mathbf{x}')}{(\beta')^{3/2}} &= \frac{\|\mathbf{Dz}'\|^2 - \alpha'^2 - 2\alpha'\beta' - \beta'^2}{(\beta')^{3/2}} \\
&= \frac{(\|\mathbf{Dz}\|^2 - \alpha^2)/\rho - 2\alpha\beta - \rho\beta^2}{\beta^{3/2}} \\
&= \frac{(\|\mathbf{Dz}\|^2 - \alpha^2 - 2\alpha\beta - \beta^2)/\rho + 2(1/\rho - 1)\alpha\beta + (1/\rho - \rho)\beta^2}{\beta^{3/2}} \\
&\leq \frac{H(\mathbf{x})}{\rho\beta^{3/2}} + \frac{2(1/\rho - 1)\alpha\beta}{\beta^{3/2}},
\end{aligned}$$

where the last inequality is due to  $1/\rho - \rho < 0$ . Note that since  $\|\mathbf{x}'\| = 1$ , it holds that  $H(\mathbf{x}') = \|\mathbf{Dz}'\| - (\alpha' + \beta')^2 = \|\mathbf{Dz}' - (\alpha' + \beta')(\mathbf{x}')\|^2 \geq 0$ . Hence we have

$$\frac{H(\mathbf{x})}{\beta^{3/2}} \geq \frac{2\rho(1-1/\rho)\alpha\beta}{\beta^{3/2}} \stackrel{(50)}{\geq} \frac{\rho-1}{2} \sqrt{\mathbf{d}_{K+1}} \geq \frac{1/\sqrt{1-c^2}-1}{2} \sqrt{\mathbf{d}_{K+1}}. \quad (61)$$

Combining (44–47), (49), (51), (60) and (61) (and (78) in Appendix 7.3 for the case with  $\dim(\text{Range}(\mathbf{D})) \leq 2$ ) with Lemma 4.3, and noting that, by reducing  $\epsilon_0$  if necessary,  $\beta^{3/4} \leq \beta^{1/2}$ , we know that (48) holds with some positive constant  $\tau$  and sufficient small  $\epsilon > 0$ .  $\square$

We summarize the results of Theorems 4.2 and Lemmas 4.3–4.7 in the following theorem.

**Theorem 4.8.** *For any optimal solution  $\mathbf{x}^*$  to the TRS  $(P_0)$ , there exist some constant  $\tau_{KL} > 0$  and sufficiently small  $\epsilon > 0$  such that the KL inequality holds*

$$(f(\mathbf{x}) - f^* + \delta_B(\mathbf{x}))^\varrho \leq \tau_{KL} \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \epsilon), \quad (62)$$

where  $\varrho = \begin{cases} 3/4, & \text{for the TRS-ill case (30),} \\ 1/2, & \text{otherwise.} \end{cases}$

The above theorem in fact shows us that the KL exponent 3/4 holds for all cases, due to the fact that the function value  $f(\mathbf{x})$  is bounded in the unit ball and the inequality  $t^{3/4} \leq t^{1/2}$  for all  $0 \leq t \leq 1$ .

**Remark 4.9.** *From the proofs for Lemmas 4.3–4.7, we see that, in addition to the problem data, the KL constant  $\tau_{KL}$  also depends on the reference optimal point  $\mathbf{x}^*$  and the optimal Lagrange multiplier  $\lambda^*$ . Here, we conjecture that, similar as shown in [30, 8, 37], the KL constant  $\tau_{KL}$  should have some connection with the error bound constant  $\tau_{EB}$  in Theorem 3.13. Unfortunately, our current approach, relying heavily on the special structures of the TRS especially for the nonconvex case, cannot reveal the explicit relation between  $\tau_{EB}$  and  $\tau_{KL}$ . This relation will be further explored in our future work.*

**Remark 4.10.** *We should point out that our result for the KL inequality only holds locally around global optimal solutions. It can be shown the KL inequality in Definition 1.2 fails to hold if the test point  $\mathbf{x}$  is chosen to be any non-optimal stationary point of the TRS (see [3, Example 9.6] for discussions of stationary points of the TRS, and [48, 58] for the existence of non-optimal stationary points of the TRS).*

## 5 Convergence analysis of projected gradient methods

Recently, [4] demonstrated that with a proper initialization, projected gradient methods (in fact, more general first order conic methods) for solving the TRS converge to a global optimal solution. However, the rate of convergence for these algorithms are not studied in their paper. Meanwhile, it is well known that projected gradient methods converge to a stationary point in rate  $O(1/\sqrt{k})$  for minimizing a nonconvex Lipschitz continuous differentiable function over a closed convex set with  $k$  being the iteration index; see, e.g., [51, Chapter 1] for unconstrained minimization and [24, Corollary 1] for general composite minimization. Hence, a straightforward conclusion will be that projected gradient methods for the TRS achieve a sublinear iteration complexity of  $O(1/\sqrt{k})$ . Here, we improve this result by showing that the local convergence rate of projected gradient methods for solving the TRS can be improved to at least  $O(1/k^2)$ . In fact, in most cases, it even enjoys local linear convergence. As one can observe in the subsequent analysis in this section, the cornerstone for these superior improvements is the obtained KL inequality for the TRS.

Let  $\Pi_B : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be the Euclidean projector onto the unit norm ball  $B$ , i.e., for any  $\mathbf{z} \in \mathbb{R}^n$ ,

$$\Pi_B(\mathbf{z}) = \begin{cases} \frac{\mathbf{z}}{\|\mathbf{z}\|}, & \text{if } \|\mathbf{z}\| \geq 1, \\ \mathbf{z}, & \text{otherwise.} \end{cases}$$

Our main contribution in this section is summarized in the following theorem for constant step size projected gradient methods. Note that the Lipschitz constant of the gradient of the objective function in the TRS is  $L = 2\|\mathbf{A}\|_2$  and the step size we will use is  $t \in (0, \frac{2}{L})$ , i.e.,  $t \in (0, \frac{1}{\|\mathbf{A}\|_2})$ . Specifically, we show that projected gradient methods achieve a locally sublinear rate of  $O(1/k^2)$  in the *TRS-ill case*; otherwise, the local convergence rate can be further improved to linear.

**Theorem 5.1** (Constant step size projected gradient methods). *Let the step size  $t \in (0, 2/L)$ , i.e.,  $t \in (0, 1/\|\mathbf{A}\|_2)$ . Suppose for all  $k \geq 0$  that  $\mathbf{x}_{k+1} = \Pi_B(\mathbf{x}_k - t\nabla f(\mathbf{x}_k))$  and the sequence  $\{\mathbf{x}_k\}$  converge to an optimal solution  $x^*$  to the TRS  $(\mathbf{P}_0)$ . Then, there exists a sufficiently large positive integer  $N$  such that  $\{\mathbf{x}_k\}_{k \geq N} \subseteq B(\mathbf{x}^*, \epsilon)$  where  $\epsilon > 0$  is the same constant in Theorem 4.8.*

Let  $M = \frac{\tau_{KL}(2\|\mathbf{A}\|_2 + 1/t)}{\sqrt{1/t - \|\mathbf{A}\|_2}}$  with  $\tau_{KL} > 0$  being the constant in Theorem 4.8. Then, in the TRS-ill case (30), it holds that

$$f(\mathbf{x}_k) - f^* \leq \frac{1}{\left(\frac{k-N}{2M^2 + \frac{3}{2}\sqrt{r_N}} + \frac{1}{\sqrt{f(\mathbf{x}_N) - f^*}}\right)^2}, \quad \forall k \geq N; \quad (63)$$

otherwise, it holds that

$$f(\mathbf{x}_k) - f^* \leq \left(\frac{M^2}{M^2 + 1}\right)^k (f(\mathbf{x}_N) - f^*), \quad \forall k \geq N.$$

Proof. Since  $\mathbf{x}_k \rightarrow \mathbf{x}^*$  as  $k \rightarrow \infty$ , for the given constant  $\epsilon > 0$  in Theorem 4.8, there exists  $N > 0$  such that  $\mathbf{x}_k \in B(\mathbf{x}^*, \epsilon)$  for all  $k \geq N$ .

For all  $k \geq 0$ , one can rewrite the updating rule  $\mathbf{x}_{k+1} = \Pi_B(\mathbf{x}_k - t\nabla f(\mathbf{x}_k))$  in the following manner

$$\mathbf{x}_{k+1} = \operatorname{argmin}_{\mathbf{u}} \left\{ \nabla f(\mathbf{x}_k)^T (\mathbf{u} - \mathbf{x}_k) + \frac{1}{2t} \|\mathbf{u} - \mathbf{x}_k\|^2 + \delta_B(\mathbf{u}) \right\},$$

whose optimality condition asserts:

$$0 \in \mathbf{x}_{k+1} - (\mathbf{x}_k - t\nabla f(\mathbf{x}_k)) + N_B(\mathbf{x}_{k+1})$$

with  $N_B(\mathbf{x}_{k+1})$  being the normal cone of  $B$  at  $\mathbf{x}_{k+1}$ . Denote for all  $k \geq 0$ ,  $\mathbf{v}_{k+1} = -(\mathbf{x}_{k+1} - \mathbf{x}_k)/t - \nabla f(\mathbf{x}_k)$ . Then, it holds for all  $k \geq 0$  that  $\mathbf{v}_{k+1} \in N_B(\mathbf{x}_{k+1})$  and

$$\|\mathbf{v}_{k+1} + \nabla f(\mathbf{x}_k)\| = \|\mathbf{x}_{k+1} - \mathbf{x}_k\|/t. \quad (64)$$

Since  $1/t > \|\mathbf{A}\|_2 = L/2$ , from the proof in Proposition 2.3.2 in [6], we know that

$$f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k) \leq (L/2 - 1/t)\|\mathbf{x}_k - \mathbf{x}_{k+1}\|^2 = (\|\mathbf{A}\|_2 - 1/t)\|\mathbf{x}_k - \mathbf{x}_{k+1}\|^2, \quad \forall k \geq 0. \quad (65)$$

Meanwhile, from the KL inequality (62), it holds for all  $k \geq N$  that

$$\begin{aligned} (f(\mathbf{x}_{k+1}) - f^*)^\varrho &\leq \tau_{KL} \operatorname{dist}(-\nabla f(\mathbf{x}_{k+1}), N_B(\mathbf{x}_{k+1})) \\ &\leq \tau_{KL} \|\nabla f(\mathbf{x}_{k+1}) + \mathbf{v}_{k+1}\| \\ &= \tau_{KL} \|\nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k) + \mathbf{v}_{k+1}\| \\ &\leq \tau_{KL} (\|\nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)\| + \|\nabla f(\mathbf{x}_k) + \mathbf{v}_{k+1}\|), \end{aligned}$$

where  $\varrho$  is the KL exponent and  $\tau_{KL} > 0$  is the constant in Theorem 4.8. The above inequality, together with the Lipschitz continuity of  $\nabla f$  (note that the Lipschitz constant is  $L = 2\|\mathbf{A}\|_2$ ) and (64), gives

$$(f(\mathbf{x}_{k+1}) - f^*)^\varrho \leq \tau_{KL}(2\|\mathbf{A}\|_2 + 1/t)\|\mathbf{x}_k - \mathbf{x}_{k+1}\|, \quad \forall k \geq N.$$

Substituting this to (65) implies

$$(f(\mathbf{x}_{k+1}) - f^*)^\varrho \leq \frac{\tau_{KL}(2\|\mathbf{A}\|_2 + 1/t)}{\sqrt{1/t - \|\mathbf{A}\|_2}} (f(\mathbf{x}_k) - f(\mathbf{x}_{k+1}))^{\frac{1}{2}}, \quad \forall k \geq N.$$

Defining  $r_k = f(\mathbf{x}_k) - f^*$  and  $M = \frac{\tau_{KL}(2\|\mathbf{A}\|_2 + 1/t)}{\sqrt{1/t - \|\mathbf{A}\|_2}}$ , we have

$$r_{k+1}^\varrho \leq M(r_k - r_{k+1})^{\frac{1}{2}}, \quad \forall k \geq N. \quad (66)$$

We divide our discussions into two cases:

- In the *TRS-ill case* (30), from Theorem 4.8 and (66), we have  $r_{k+1}^{\frac{3}{2}} \leq M^2(r_k - r_{k+1})$ . Hence, for all  $k \geq N$ , we have from (65) that  $r_{k+1} \leq r_k$  and

$$\begin{aligned}
\frac{1}{\sqrt{r_{k+1}}} - \frac{1}{\sqrt{r_k}} &\geq \frac{1}{\sqrt{r_{k+1}}} - \frac{1}{\sqrt{r_{k+1} + \frac{1}{M^2}r_{k+1}^{3/2}}} \\
&= \frac{\sqrt{r_{k+1} + \frac{1}{M^2}r_{k+1}^{3/2}} - \sqrt{r_{k+1}}}{\sqrt{r_{k+1}}\sqrt{r_{k+1} + \frac{1}{M^2}r_{k+1}^{3/2}}} \\
&= \frac{r_{k+1}^{3/2}}{M^2\sqrt{r_{k+1}}\sqrt{r_{k+1} + \frac{1}{M^2}r_{k+1}^{3/2}}(\sqrt{r_{k+1} + \frac{1}{M^2}r_{k+1}^{3/2}} + \sqrt{r_{k+1}})} \\
&= \frac{1}{M^2\sqrt{1 + \frac{1}{M^2}\sqrt{r_{k+1}}}(\sqrt{1 + \frac{1}{M^2}\sqrt{r_{k+1}}} + 1)} \\
&= \frac{1}{M^2(\sqrt{1 + \frac{1}{M^2}\sqrt{r_{k+1}}} + 1 + \frac{1}{M^2}\sqrt{r_{k+1}})} \\
&\geq \frac{1}{M^2(2 + \frac{3}{2M^2}\sqrt{r_{k+1}})} \\
&\geq \frac{1}{M^2(2 + \frac{3}{2M^2}\sqrt{r_N})} \\
&\geq \frac{1}{2M^2 + \frac{3}{2}\sqrt{r_N}},
\end{aligned}$$

where the second inequality follows from  $\sqrt{1 + \frac{1}{M^2}\sqrt{r_{k+1}}} \leq 1 + \frac{1}{2M^2}\sqrt{r_{k+1}}$ . Hence, for all  $k \geq N$ , we have

$$\frac{1}{\sqrt{r_k}} \geq \frac{k - N}{2M^2 + \frac{3}{2}\sqrt{r_N}} + \frac{1}{\sqrt{r_N}}$$

and thus

$$r_k \leq \frac{1}{\left(\frac{k-N}{2M^2 + \frac{3}{2}\sqrt{r_N}} + \frac{1}{\sqrt{r_N}}\right)^2}.$$

- Otherwise, from Theorem 4.8 and (66), we have  $r_{k+1} \leq M^2(r_k - r_{k+1})$  for all  $k \geq N$ . This implies

$$r_k \leq \left(\frac{M^2}{M^2 + 1}\right)^{k-N} r_N, \quad \forall k \geq N.$$

We thus complete the proof for the theorem.  $\square$

We remark that the sublinear convergence rate (63) holds in all cases as the KL inequality holds with exponent  $3/4$  for all cases. Note that the assumption in Theorem 5.1 that projected gradient methods converge to a global optimal solution is not restrictive at all. As is mentioned in the introduction and the beginning of this section, the assumption can be guaranteed as long as the starting point of the projection gradient method is properly chosen [4]. Moreover, it is also noted in [4] that the initial point can be obtained without much difficulty. Although the iteration complexity results derived in Theorem 5.1 only holds locally around the optimal solution  $\mathbf{x}^*$ , by using similar ideas in [26], one can directly extend these results to a global version. To see this, let  $N$  be the positive integer in Theorem 5.1. Note from the proof



of Theorem 5.1 that  $\{f(\mathbf{x}_k)\}_{k \geq 0}$  is non-increasing. Then, we have that for all  $1 \leq k \leq N$ ,

$$0 \leq f(\mathbf{x}_k) - f^* \leq f(\mathbf{x}_0) - f^* \leq \min \left\{ \frac{N^2(f(\mathbf{x}_0) - f^*)}{k^2}, (f(\mathbf{x}_0) - f^*) \left( \frac{M^2}{M^2 + 1} \right)^{k-N} \right\}.$$

Hence it follows from Theorem 5.1 that there exist constants  $C_1, C_2 > 0$  such that

$$f(\mathbf{x}_k) - f(\mathbf{x}^*) \leq \begin{cases} \frac{C_1}{k^2}, & \text{for the TRS-ill case (30),} \\ C_2 \left( \frac{M^2}{M^2 + 1} \right)^k, & \text{otherwise,} \end{cases} \quad \forall k \geq 1.$$

Note that constants  $C_1, C_2$  depend on the choice of the initial point and can be hard to estimate explicitly. In light of Theorem 3.13, we can further derive the convergence rate associated with  $\text{dist}(\mathbf{x}_k, S_0)$ :

$$\text{dist}(\mathbf{x}_k, S_0) \leq \begin{cases} \frac{\tau_{EB} C_1^{1/4}}{\sqrt{k}}, & \text{for the TRS-ill case (30),} \\ \tau_{EB} \sqrt{C_2} \left( \frac{M^2}{M^2 + 1} \right)^{k/2}, & \text{otherwise,} \end{cases} \quad \forall k \geq 1,$$

where  $\tau_{EB} > 0$  is the error bound modulus in Theorem 3.13. We also note that the projected gradient method with backtracking line search can also be analyzed in a similar way, as (65) still holds with a little more conservative constant, and thus we omit its analysis for simplicity.

**Remark 5.2.** *It would be interesting to compare Theorem 5.1 with the results obtained in [65] and [10], which demonstrated that the generalized Lanczos trust-region (GLTR) method has a linear convergence rate for the easy case. While we have proved that projected gradient methods converge sublinearly at rate  $O(1/k^2)$  for the TRS-ill case, and linearly otherwise (including the easy case).*

**Remark 5.3.** *Based on the established KL inequality for the TRS in Theorem 4.8, the same order of the convergence rate for the gap  $f(\mathbf{x}_k) - f(\mathbf{x}^*)$  and  $\text{dist}(\mathbf{x}_k, S_0)$  can also be obtained by using the arguments in [22]. However, our proof here is simpler and has an explicit dependence of the constants.*

## 6 Conclusion

In this paper, we conducted a thorough analysis about Hölderian error bounds and the KL inequality associated with the TRS. Specifically, we showed that for the TRS, a Hölderian error bound holds with modulus 1/4 and the KL inequality holds with exponent 3/4. Moreover, we demonstrated that the Hölderian error bound modulus and the KL exponent in fact are both 1/2 unless in the *TRS-ill case (30)*. As a byproduct, we further proved that projected gradient methods for solving the TRS enjoy a local sublinear convergence rate  $O(1/k^2)$ , which can be further improved to a linear rate unless in the *TRS-ill case*.

## Acknowledgements

We would like to thank Professor Jong-Shi Pang at University of Southern California for his helpful comments on an early version of this paper and the two anonymous referees for helpful suggestions to improve the quality of this paper.

This research was supported by National Key R&D Program of China, 2020YFA0711900, 2020YFA0711901; the National Natural Science Foundation of China, 12171100; the ‘‘Chenguang Program’’ by the Shanghai Education Development Foundation and Shanghai Municipal Education Commission, 19CG02; and the Shanghai Science and Technology Program, 21JC1400600.

## 7 Appendix

### 7.1 Complementary proof for Lemma 4.5

Here, we provide a supporting lemma that helps prove Lemma 4.5. With the same notation  $\alpha, \beta$  in (44) and  $H(\mathbf{x})$  in (46), recall that  $\|\mathbf{s}\| = \|\mathbf{P}^T \mathbf{x}^*\| = 1$  and  $\{\mathbf{d}_i\}_{i=1}^n$  in (42). We have the following lemma.

**Lemma 7.1.** *It holds that for all  $\mathbf{z} \in \mathbb{R}^n$*

$$H(\mathbf{x}) = \sum_{i=1}^K \mathbf{s}_i^2 \sum_{j=K+1}^n (\mathbf{d}_j \mathbf{z}_j)^2 + \sum_{K+1 \leq i < j \leq n} (\mathbf{d}_i \mathbf{z}_i \mathbf{s}_j - \mathbf{d}_j \mathbf{z}_j \mathbf{s}_i)^2 - 2\alpha\beta - \beta^2.$$

Moreover, if  $\mathbf{s} \notin \text{Null}(\mathbf{D})$ , it holds that

$$H(\mathbf{x}) \geq \frac{1}{2} \mathbf{d}_{K+1} \left( \sum_{j=1}^K \mathbf{s}_j^2 \right) \beta \quad (67)$$

for all  $\mathbf{z}$  with sufficiently small  $\|\mathbf{z}\|$ .

Proof. By some calculations, we see that

$$\begin{aligned} H(\mathbf{x}) &= \sum_{i=1}^n (\mathbf{d}_i \mathbf{z}_i)^2 - \left( \sum_{i=1}^n \mathbf{s}_i \mathbf{d}_i \mathbf{z}_i \right)^2 - 2\alpha\beta - \beta^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n \mathbf{s}_j^2 (\mathbf{d}_i \mathbf{z}_i)^2 - \left( \sum_{i=1}^n \mathbf{s}_i \mathbf{d}_i \mathbf{z}_i \right)^2 - 2\alpha\beta - \beta^2 \\ &= \sum_{1 \leq i < j \leq n} (\mathbf{d}_i \mathbf{z}_i \mathbf{s}_j - \mathbf{d}_j \mathbf{z}_j \mathbf{s}_i)^2 - 2\alpha\beta - \beta^2 \\ &= \sum_{i=1}^K \sum_{j=K+1}^n (\mathbf{d}_i \mathbf{z}_i \mathbf{s}_j - \mathbf{d}_j \mathbf{z}_j \mathbf{s}_i)^2 + \sum_{i=1}^K \sum_{j=i+1}^K (\mathbf{d}_i \mathbf{z}_i \mathbf{s}_j - \mathbf{d}_j \mathbf{z}_j \mathbf{s}_i)^2 \\ &\quad + \sum_{K+1 \leq i < j \leq n} (\mathbf{d}_i \mathbf{z}_i \mathbf{s}_j - \mathbf{d}_j \mathbf{z}_j \mathbf{s}_i)^2 - 2\alpha\beta - \beta^2 \\ &= \sum_{i=1}^K \mathbf{s}_i^2 \sum_{j=K+1}^n (\mathbf{d}_j \mathbf{z}_j)^2 + \sum_{K+1 \leq i < j \leq n} (\mathbf{d}_i \mathbf{z}_i \mathbf{s}_j - \mathbf{d}_j \mathbf{z}_j \mathbf{s}_i)^2 - 2\alpha\beta - \beta^2, \end{aligned}$$

where the second equation holds since  $\|\mathbf{s}\| = 1$ , the third equation follows from the fact that  $\sum_{1 \leq i, j \leq n} a_i^2 b_j^2 - (\sum_{1 \leq i \leq n} a_i b_i)^2 = \sum_{1 \leq i < j \leq n} (a_i b_j - a_j b_i)^2$ , and the last equation is due to  $\mathbf{d}_i = 0$  for  $i = 1, \dots, K$ .

Since  $\sum_{i=K+1}^n \mathbf{d}_i^2 \mathbf{z}_i^2 \geq \sum_{i=K+1}^n \mathbf{d}_{K+1} \mathbf{d}_i \mathbf{z}_i^2 = \mathbf{d}_{K+1} \beta$ , we obtain that

$$H(\mathbf{x}) \geq (\mathbf{d}_{K+1} \sum_{j=1}^K \mathbf{s}_j^2 - 2\alpha - \beta) \beta, \quad \forall \mathbf{z} \in \mathbb{R}^n.$$

If, further,  $\mathbf{s} \notin \text{Null}(\mathbf{D})$ , together with  $\mathbf{d}_{K+1} > 0$ , we know that  $\mathbf{d}_{K+1} \sum_{j=1}^K \mathbf{s}_j^2 > 0$ . This together with the fact that  $\alpha \rightarrow 0$  and  $\beta \rightarrow 0$  as  $\mathbf{z} \rightarrow \mathbf{0}$ , implies that (67) holds when  $\|\mathbf{z}\|$  is sufficiently small. We thus complete the proof.  $\square$

## 7.2 Proof of Lemma 4.6

Proof. Recall the spectral decomposition of  $\mathbf{A}$  and the definition of the diagonal matrix  $\mathbf{D}$  as:

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T = \mathbf{P}\text{Diag}(\lambda)\mathbf{P}^T, \quad \mathbf{A} - \lambda_1\mathbf{I} = \mathbf{P}\mathbf{D}\mathbf{P}^T, \quad \mathbf{D} = \mathbf{\Lambda} - \lambda_1\mathbf{I},$$

where  $\mathbf{P}\mathbf{P}^T = \mathbf{P}^T\mathbf{P} = \mathbf{I}$ . Define linear operators  $\mathcal{L}_K : \mathbb{R}^n \rightarrow \mathbb{R}^K$  and  $\mathcal{U}_K : \mathbb{R}^n \rightarrow \mathbb{R}^{n-K}$  as

$$\mathcal{L}_K(\mathbf{x}) = [\mathbf{x}_1, \dots, \mathbf{x}_K]^T, \quad \mathcal{U}_K(\mathbf{x}) = [\mathbf{x}_{K+1}, \dots, \mathbf{x}_n]^T, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Let  $\mathbf{M} = \text{Diag}(\mathcal{U}_K(\lambda))$  where  $\lambda \in \mathbb{R}^n$  is the vector formed by all the eigenvalues of  $\mathbf{A}$  in ascending order. Define function  $q$  as

$$q(\mathbf{z}) := \langle \mathbf{z}, \mathbf{M}\mathbf{z} \rangle - 2\langle \mathcal{U}_K(\mathbf{P}^T\mathbf{b}), \mathbf{z} \rangle, \quad \forall \mathbf{z} \in \mathbb{R}^{n-K}.$$

Consider the following optimization problem

$$\min \left\{ q(\mathbf{z}) \mid \mathbf{z} \in \widehat{B} \right\} \quad \text{with } \widehat{B} := \{ \mathbf{z} \in \mathbb{R}^{n-K} \mid \|\mathbf{z}\| \leq 1 \}. \quad (68)$$

We show next that  $\mathcal{U}_K(\mathbf{P}^T\mathbf{x}^*)$  is an optimal solution to (68). Since  $\mathbf{b}, \mathbf{x}^* \in \text{Range}(\mathbf{A} - \lambda_1\mathbf{I})$ , we know that  $\mathbf{P}^T\mathbf{b}, \mathbf{P}^T\mathbf{x}^* \in \text{Range}(\mathbf{\Lambda} - \lambda_1\mathbf{I})$ . Note that since the first  $K$  diagonal entries of the diagonal matrix  $\mathbf{\Lambda} - \lambda_1\mathbf{I}$  are zeros, we have  $\mathcal{L}_K(\mathbf{P}^T\mathbf{x}^*) = \mathcal{L}_K(\mathbf{P}^T\mathbf{b}) = \mathbf{0}$  and  $\|\mathcal{U}_K(\mathbf{P}^T\mathbf{x}^*)\| = 1$ . Moreover, since  $(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x}^* = \mathbf{b}$ , we have  $(\mathbf{\Lambda} - \lambda_1\mathbf{I})(\mathbf{P}^T\mathbf{x}^*) = \mathbf{P}^T\mathbf{b}$ . Therefore,  $(\mathbf{M} - \lambda_1\mathbf{I})\mathcal{U}_K(\mathbf{P}^T\mathbf{x}^*) = \mathcal{U}_K(\mathbf{P}^T\mathbf{b})$ , i.e.,  $\mathcal{U}_K(\mathbf{P}^T\mathbf{x}^*)$  and  $-\lambda_1$  solve the optimality conditions (see the KKT conditions in Section 2) associated with the smaller dimensional TRS (68). Since  $\lambda_{K+1}$  is the smallest eigenvalue of  $\mathbf{M}$  and an optimal Lagrange multiplier is  $-\lambda_1 > -\lambda_{K+1}$ , we know that problem (68) falls into the easy case or the hard case 1. Hence, by Lemma 4.4, we have that there exist  $\epsilon_2, \tau > 0$  such that for all  $\mathbf{z} \in B(\mathcal{U}_K(\mathbf{P}^T\mathbf{x}^*), \epsilon_2)$ ,

$$[q(\mathbf{z}) - q(\mathcal{U}_K(\mathbf{P}^T\mathbf{x}^*)) + \delta_{\widehat{B}}(\mathbf{z})]^{1/2} \leq \tau \text{dist}(-\nabla q(\mathbf{z}), N_{\widehat{B}}(\mathbf{z})). \quad (69)$$

Note that for all  $\mathbf{x} \in \text{Range}(\mathbf{A} - \lambda_1\mathbf{I})$ , it holds that  $\mathcal{L}_K(\mathbf{P}^T\mathbf{x}) = \mathbf{0}$  and  $\|\mathcal{U}_K(\mathbf{P}^T\mathbf{x})\| = \|\mathbf{P}^T\mathbf{x}\| = \|\mathbf{x}\|$ . Then, for all  $\mathbf{x} \in \text{Range}(\mathbf{A} - \lambda_1\mathbf{I})$  with  $\|\mathbf{x}\| < 1$ , we know that  $\|\mathcal{U}_K(\mathbf{P}^T\mathbf{x})\| < 1$ ,  $N_{\widehat{B}}(\mathcal{U}_K(\mathbf{P}^T\mathbf{x})) = \mathbf{0}$ , and

$$\begin{aligned} \text{dist}(-\nabla q(\mathcal{U}_K(\mathbf{P}^T\mathbf{x})), N_{\widehat{B}}(\mathcal{U}_K(\mathbf{P}^T\mathbf{x}))) &= \|\nabla q(\mathcal{U}_K(\mathbf{P}^T\mathbf{x}))\| = \|2(\mathbf{M}\mathcal{U}_K(\mathbf{P}^T\mathbf{x}) - \mathcal{U}_K(\mathbf{P}^T\mathbf{b}))\| \\ &= \|2(\mathbf{\Lambda}\mathbf{P}^T\mathbf{x} - \mathbf{P}^T\mathbf{b})\| = \|2(\mathbf{P}\mathbf{\Lambda}\mathbf{P}^T\mathbf{x} - \mathbf{b})\| \\ &= \|2(\mathbf{A}\mathbf{x} - \mathbf{b})\| = \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})). \end{aligned} \quad (70)$$

Meanwhile, for all  $\mathbf{x} \in \text{Range}(\mathbf{A} - \lambda_1\mathbf{I})$  with  $\|\mathbf{x}\| = 1$ , it holds  $\|\mathcal{U}_K(\mathbf{P}^T\mathbf{x})\| = 1$  and

$$\|\mathcal{U}_K(\mathbf{P}^T\mathbf{x}) - \mathcal{U}_K(\mathbf{P}^T\mathbf{x}^*)\| = \|\mathbf{P}^T\mathbf{x} - \mathbf{P}^T\mathbf{x}^*\| = \|\mathbf{x} - \mathbf{x}^*\|.$$

Then for all  $\mathbf{x} \in \mathbf{bd}(B) \cap \text{Range}(\mathbf{A} - \lambda_1\mathbf{I})$  satisfying  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon := \min\{\epsilon_0, \epsilon_2\}$ , we further have

$$\begin{aligned} &\text{dist}(-\nabla q(\mathcal{U}_K(\mathbf{P}^T\mathbf{x})), N_{\widehat{B}}(\mathcal{U}_K(\mathbf{P}^T\mathbf{x}))) \\ &= \|2(\mathbf{M}\mathcal{U}_K(\mathbf{P}^T\mathbf{x}) - \mathcal{U}_K(\mathbf{P}^T\mathbf{b})) + \langle -2(\mathbf{M}\mathcal{U}_K(\mathbf{P}^T\mathbf{x}) - \mathcal{U}_K(\mathbf{P}^T\mathbf{b})), \mathcal{U}_K(\mathbf{P}^T\mathbf{x}) \rangle \mathcal{U}_K(\mathbf{P}^T\mathbf{x})\| \\ &= \|2(\mathbf{\Lambda}\mathbf{P}^T\mathbf{x} - \mathbf{P}^T\mathbf{b}) + \langle -2(\mathbf{\Lambda}\mathbf{P}^T\mathbf{x} - \mathbf{P}^T\mathbf{b}), \mathbf{P}^T\mathbf{x} \rangle \mathbf{P}^T\mathbf{x}\| \\ &= \|2(\mathbf{P}\mathbf{\Lambda}\mathbf{P}^T\mathbf{x} - \mathbf{b}) + \langle -2(\mathbf{P}\mathbf{\Lambda}\mathbf{P}^T\mathbf{x} - \mathbf{b}), \mathbf{x} \rangle \mathbf{x}\| \\ &= \|2(\mathbf{A}\mathbf{x} - \mathbf{b}) + \langle -2(\mathbf{A}\mathbf{x} - \mathbf{b}), \mathbf{x} \rangle \mathbf{x}\| \\ &= \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})), \end{aligned} \quad (71)$$

where  $\epsilon_0$  is given in (40), the first and last equations follow from (38), (39), (40) and

$$\langle -2(\mathbf{M}\mathcal{U}_K(\mathbf{P}^T\mathbf{x}) - \mathcal{U}_K(\mathbf{P}^T\mathbf{b})), \mathcal{U}_K(\mathbf{P}^T\mathbf{x}) \rangle = \langle -2(\mathbf{A}\mathbf{x} - \mathbf{b}), \mathbf{x} \rangle > 0,$$

and the third equation holds since  $\mathbf{P}$  is an orthogonal matrix. Since  $q(\mathcal{U}_K(\mathbf{P}^T\mathbf{x}^*)) = f(\mathbf{x}^*) = f^*$  and  $q(\mathcal{U}_K(\mathbf{P}^T\mathbf{x})) = f(\mathbf{x})$  for all  $\mathbf{x} \in \text{Range}(\mathbf{A} - \lambda_1\mathbf{I})$ , we know from (69), (70) and (71) that

$$(f(\mathbf{x}) - f(\mathbf{x}^*) + \delta_B(\mathbf{x}))^{1/2} \leq \tau \text{dist}(-\nabla f(\mathbf{x}), N_B(\mathbf{x})),$$

for all  $\mathbf{x} \in B(\mathbf{x}^*, \epsilon) \cap \text{Range}(\mathbf{A} - \lambda_1\mathbf{I})$ .  $\square$

### 7.3 Complementary proof for Lemma 4.7

As a complementary proof for Lemma 4.7, we use the same notation as in the main proof. Here, we consider **Case II** with  $\dim(\text{Range}(\mathbf{D})) \leq 2$ . Recall that we are in case where  $\|\mathbf{s}\| = \|\mathbf{s} + \mathbf{z}\| = 1$ ,  $\alpha = \langle \mathbf{s}, \mathbf{D}\mathbf{z} \rangle > 0$ ,  $\|\mathbf{z}\|^2 = \gamma < 1/4$ ,  $\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2$  with  $\mathbf{z}_1 \in \text{Range}(\mathbf{D})$  and  $\mathbf{z}_2 \in \text{Null}(\mathbf{D})$  and  $\|\mathbf{z}_2\| \leq c\sqrt{\gamma}$  with constant  $c \leq 1/16$  given in (52). It can also be verified that  $\langle \mathbf{s}, \mathbf{z}_1 \rangle = \langle \mathbf{s}, \mathbf{z} \rangle = -\gamma/2$ .

We first claim that  $\dim(\text{Range}(\mathbf{D})) > 1$ . Indeed, if  $\dim(\text{Range}(\mathbf{D})) = 1$ , since  $\mathbf{s} \in \text{Range}(\mathbf{D})$ , it holds that  $\text{Range}(\mathbf{D}) = \text{span}\{\mathbf{s}\}$ , i.e., the space spanned by  $\{\mathbf{s}\}$ . Since  $\mathbf{z}_1 \in \text{Range}(\mathbf{D})$  and  $\langle \mathbf{s}, \mathbf{z}_1 \rangle = -\gamma/2$ , we have  $\mathbf{z}_1 = -\gamma\mathbf{s}/2$ . Hence,

$$\langle \mathbf{s}, \mathbf{D}\mathbf{z} \rangle = \langle \mathbf{s}, \mathbf{D}\mathbf{z}_1 \rangle = -\frac{\gamma}{2} \langle \mathbf{s}, \mathbf{D}\mathbf{s} \rangle \stackrel{(42)}{<} 0.$$

However, this contradicts our assumption that  $\alpha = \langle \mathbf{s}, \mathbf{D}\mathbf{z} \rangle > 0$ . Hence,  $\dim(\text{Range}(\mathbf{D})) = 2$ . Therefore, (42) reduces to

$$\mathbf{d}_n \geq \mathbf{d}_{n-1} > 0 = \mathbf{d}_{n-2} = \dots = \mathbf{d}_1, \quad \text{i.e.,} \quad K = n - 2.$$

Let  $\mathbf{e}_i$ ,  $i = 1, \dots, n$ , be the standard basis of  $\mathbb{R}^n$ . Then, we can represent  $\mathbf{s}$  and  $\mathbf{z}$  by the linear combinations of  $\mathbf{e}_n$  and  $\mathbf{e}_{n-1}$  as  $\mathbf{s} = u_1\mathbf{e}_{n-1} + u_2\mathbf{e}_n$  and  $\mathbf{z}_1 = w_1\mathbf{e}_{n-1} + w_2\mathbf{e}_n$ , respectively. Since  $\|\mathbf{s}\| = 1$ ,  $\|\mathbf{z}_1\|^2 = \gamma$  and  $\|\mathbf{z}_2\| \leq c\sqrt{\gamma}$ , we know that

$$u_1^2 + u_2^2 = 1, \quad \text{and} \quad \|\mathbf{z}_1\|^2 = w_1^2 + w_2^2 = \gamma - \|\mathbf{z}_2\|^2 \geq (1 - c^2)\gamma. \quad (72)$$

We discuss two cases here, i.e., cases with  $(u_2\mathbf{d}_n w_2)(u_1\mathbf{d}_{n-1} w_1) > 0$  and  $(u_2\mathbf{d}_n w_2)(u_1\mathbf{d}_{n-1} w_1) \leq 0$ . Suppose that  $(u_2\mathbf{d}_n w_2)(u_1\mathbf{d}_{n-1} w_1) > 0$ . By reducing  $\epsilon_0$  if necessary, we can assume without loss of generality that

$$\sqrt{\gamma} \leq \epsilon_1 = \min(\epsilon_0, 1/2) < \min(|u_1|, |u_2|). \quad (73)$$

Since  $\mathbf{s}^T \mathbf{z} = u_1 w_1 + u_2 w_2 = -\gamma/2 < 0$ ,  $\mathbf{d}_{n-1} > 0$  and  $\mathbf{d}_n > 0$ , we have  $(u_2\mathbf{d}_n w_2) < 0$  and  $(u_1\mathbf{d}_{n-1} w_1) < 0$ . From (72) and  $c \leq 1/16$ , we know that

$$|w_1| + |w_2| \geq \sqrt{w_1^2 + w_2^2} \geq \frac{1}{2}\sqrt{\gamma}.$$

This further implies that

$$\begin{aligned} \mathbf{s}^T \mathbf{z} = u_2 w_2 + u_1 w_1 &= -|u_2||w_2| - |u_1||w_1| \\ &\leq -\min(|u_1|, |u_2|)(|w_2| + |w_1|) \\ &\leq -\frac{\min(|u_1|, |u_2|)\sqrt{\gamma}}{2} \\ &\stackrel{(73)}{<} -\frac{\gamma}{2}. \end{aligned}$$

which contradicts the fact that  $\mathbf{s}^T \mathbf{z} = -\gamma/2$ .

Hence, we only need to consider the case with  $(u_2\mathbf{d}_n w_2)(u_1\mathbf{d}_{n-1} w_1) \leq 0$ . Then, it holds that

$$\begin{aligned} \|\mathbf{D}\mathbf{z}\|^2 - \alpha^2 &= \mathbf{d}_n^2 w_2^2 + \mathbf{d}_{n-1}^2 w_1^2 - (u_2\mathbf{d}_n w_2 + u_1\mathbf{d}_{n-1} w_1)^2 \\ &\geq (1 - u_2^2)\mathbf{d}_n^2 w_2^2 + (1 - u_1^2)\mathbf{d}_{n-1}^2 w_1^2 \\ &= u_1^2 \mathbf{d}_n^2 w_2^2 + u_2^2 \mathbf{d}_{n-1}^2 w_1^2. \end{aligned} \quad (74)$$

Suppose first that  $\min(u_1^2, u_2^2) = 0$ . Without loss of generality, we can assume  $u_1 = 0$ , and thus  $|u_2| = 1$ . Since  $\langle \mathbf{s}, \mathbf{z}_1 \rangle = -\gamma/2 = u_2 w_2$ , we know that  $w_2^2 = \gamma^2/4$ , which, together with (72) and the fact that  $\gamma \leq 1/4$  and  $c \leq 1/16$ , implies that

$$w_1^2 \geq (1 - c^2)\gamma - \gamma^2/4 \geq (1 - 1/16^2 - 1/16)\gamma \geq \frac{7}{8}\gamma. \quad (75)$$

Meanwhile, it holds that

$$\beta = \langle \mathbf{z}, \mathbf{Dz} \rangle = \langle \mathbf{z}_1, \mathbf{Dz}_1 \rangle \leq \mathbf{d}_n \|\mathbf{z}_1\|^2 \leq \mathbf{d}_n \gamma,$$

which, together with (74) and (75), implies that

$$\|\mathbf{Dz}\|^2 - \alpha^2 \geq \frac{7}{8} \mathbf{d}_{n-1}^2 \gamma \geq \frac{7 \mathbf{d}_{n-1}^2}{8 \mathbf{d}_n} \mathbf{d}_n \gamma \geq \frac{7 \mathbf{d}_{n-1}^2}{8 \mathbf{d}_n} \beta. \quad (76)$$

Suppose now that  $\min(u_1^2, u_2^2) \neq 0$ . Then, from (74), we know that

$$\begin{aligned} \|\mathbf{Dz}\|^2 - \alpha^2 &\geq \min(u_1^2, u_2^2)(\mathbf{d}_n^2 w_2^2 + \mathbf{d}_{n-1}^2 w_1^2) \\ &\geq \min(u_1^2, u_2^2) \mathbf{d}_{n-1} (\mathbf{d}_n w_2^2 + \mathbf{d}_{n-1} w_1^2) \\ &= \min(u_1^2, u_2^2) \mathbf{d}_{n-1} \beta. \end{aligned} \quad (77)$$

Combining (76) and (77), we conclude that if  $\dim(\text{Range}(\mathbf{D})) = 2$ , it holds that

$$\|\mathbf{Dz}\|^2 - \alpha^2 \geq c_2 \beta,$$

where  $c_2 = \min\{\frac{7 \mathbf{d}_{n-1}^2}{8 \mathbf{d}_n}, \min(u_1^2, u_2^2) \mathbf{d}_{n-1}\}$ . Since  $\alpha \rightarrow 0$  and  $\beta \rightarrow 0$  as  $\gamma \rightarrow 0$  (or equivalently,  $\mathbf{x} \rightarrow \mathbf{x}^*$ ), by reducing  $\epsilon_0$  if necessary, we know in **Case II** if  $\dim(\text{Range}(\mathbf{D})) = 2$  that

$$\frac{1}{4} \text{dist}^2(-\nabla f(\mathbf{x}), N_B(\mathbf{x})) = H(\mathbf{x}) = \|\mathbf{Dz}\|^2 - \alpha^2 - 2\alpha\beta - \beta^2 \geq \beta(c_2 - 2\alpha - \beta) \geq \frac{c_2}{2} \beta, \quad (78)$$

for all  $\mathbf{x} \in \mathbf{bd}(B)$  with  $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon_1$ .

## References

- [1] F. J. Aragón Artacho and M. H. Geoffroy. Characterization of metric regularity of subdifferentials. *Journal of Convex Analysis*, 15(2):365–380, 2008.
- [2] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality. *Mathematics of Operations Research*, 35(2):438–457, 2010.
- [3] A. Beck. *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*, MOS-SIAM Ser. Optim. 19, SIAM, Philadelphia, PA, 2014.
- [4] A. Beck and Y. Vaisbourd. Globally solving the trust region subproblem using simple first-order methods. *SIAM Journal on Optimization*, 28(3):1951–1967, 2018.
- [5] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009.
- [6] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1999.
- [7] J. Bolte, A. Daniilidis, O. Ley, and L. Mazet. Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. *Transactions of the American Mathematical Society*, 362(6):3319–3363, 2010.
- [8] J. Bolte, T. P. Nguyen, J. Peypouquet, and B. W. Suter. From error bounds to the complexity of first-order descent methods for convex functions. *Mathematical Programming*, 165(2):471–507, 2017.
- [9] J. V. Burke and M. C. Ferris. Weak sharp minima in mathematical programming. *SIAM Journal on Control and Optimization*, 31(5):1340–1359, 1993.

- [10] Y. Carmon and J. C. Duchi. Analysis of Krylov subspace solutions of regularized non-convex quadratic problems. In *Advances in Neural Information Processing Systems*, pages 10705–10715, 2018.
- [11] A. R. Conn, N. I. Gould, and P. L. Toint. *Trust Region Methods*, volume 1. SIAM, Philadelphia, 2000.
- [12] Y. Cui, C. Ding, and X.-Y. Zhao. Quadratic growth conditions for convex matrix optimization problems associated with spectral functions. *SIAM Journal on Optimization*, 27:2332–2355, 2017.
- [13] Y. Cui, D. Sun, and K.-C. Toh. On the R-superlinear convergence of the KKT residuals generated by the augmented Lagrangian method for convex composite conic programming. *Mathematical Programming*, 178(1-2):381–415, 2019.
- [14] A. L. Dontchev and R. T. Rockafellar. *Implicit Functions and Solution Mappings*. Springer Monographs in Mathematics, Springer, 208, 2009.
- [15] D. Drusvyatskiy and A. S. Lewis. Tilt stability, uniform quadratic growth, and strong metric regularity of the subdifferential. *SIAM Journal on Optimization*, 23(1):256–267, 2013.
- [16] D. Drusvyatskiy, B. S. Mordukhovich, T. T. A. Nghia. Second-order growth, tilt stability, and metric regularity of the subdifferential. *Journal of Convex Analysis*, 21:1165–1192, 2014.
- [17] D. Drusvyatskiy and A. S. Lewis. Error bounds, quadratic growth, and linear convergence of proximal methods. *Mathematics of Operations Research*, 43(3):919–948, 2018.
- [18] D. Drusvyatskiy, A. D. Ioffe, and A. S. Lewis. Nonsmooth optimization using Taylor-like models: error bounds, convergence, and termination criteria. *Mathematics Programming*, <https://doi.org/10.1007/s10107-019-01432-w>, 2019.
- [19] M. Fabian, R. Henrion, A. Y. Kruger, and J. V. Outrata. Error bounds: necessary and sufficient conditions. *Set-Valued Variational Analysis*, 18:121–149, 2010.
- [20] O. E. Flippo and B. Jansen. Duality and sensitivity in nonconvex quadratic optimization over an ellipsoid. *European journal of operational research*, 94(1):167–178, 1996.
- [21] C. Fortin and H. Wolkowicz. The trust region subproblem and semidefinite programming. *Optimization methods and software*, 19(1):41–67, 2004.
- [22] P. Frankel, G. Garrigos, and J. Peypouquet. Splitting methods with variable metric for Kurdyka–Lojasiewicz functions and general convergence rates. *Journal of Optimization Theory and Applications*, 165(3):874–900, 2015.
- [23] B. Gao, X. Liu, X. Chen, and Y.-X. Yuan. On the Lojasiewicz exponent of the quadratic sphere constrained optimization problem. *arXiv preprint arXiv:1611.08781*, 2016.
- [24] S. Ghadimi, G. Lan, and H. Zhang. Mini-batch stochastic approximation methods for nonconvex stochastic composite optimization. *Mathematical Programming*, 155:267–305, 2016.
- [25] N. I. Gould, S. Lucidi, M. Roma, and P. L. Toint. Solving the trust-region subproblem using the Lanczos method. *SIAM Journal on Optimization*, 9(2):504–525, 1999.
- [26] D. Han, D. Sun, and L. Zhang. Linear rate convergence of the alternating direction method of multipliers for convex composite programming. *Mathematics of Operations Research*, 43(2):622–637, 2017.

- [27] E. Hazan and T. Koren. A linear-time algorithm for trust region problems. *Mathematical Programming*, 158(1-2):363–381, 2016.
- [28] N. Ho-Nguyen and F. Kilinc-Karzan. A second-order cone based approach for solving the trust-region subproblem and its variants. *SIAM Journal on Optimization*, 27(3):1485–1512, 2017.
- [29] K. Hou, Z. Zhou, A. M.-C. So, and Z.-Q. Luo. On the linear convergence of the proximal gradient method for trace norm regularization. In *Advances in Neural Information Processing Systems*, pages 710–718, 2013.
- [30] A. D. Ioffe. Metric regularity – A survey. Part 1. Theory. *Journal of the Australian Mathematical Society*, 101:188–243, 2016.
- [31] R. Jiang and D. Li. A linear-time algorithm for generalized trust region problems. *SIAM Journal on Optimization*, 30(1): 915–932, 2020.
- [32] R. Jiang and D. Li. Novel reformulations and efficient algorithms for the generalized trust region subproblem. *SIAM Journal on Optimization*, 29(2):1603–1633, 2019.
- [33] R. Jiang, D. Li, and B. Wu. SOCP reformulation for the generalized trust region subproblem via a canonical form of two symmetric matrices. *Mathematical Programming*, 169(2):531–563, 2018.
- [34] G. Li, B. S. Mordukhovich, and T. Pham. New fractional error bounds for polynomial systems with applications to hölderian stability in optimization and spectral theory of tensors. *Mathematical Programming*, 153(2):333–362, 2015.
- [35] G. Li and T. K. Pong. Douglas–Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems. *Mathematical Programming*, 159(1-2):371–401, 2016.
- [36] G. Li and T. K. Pong. Calculus of the exponent of Kurdyka–Lojasiewicz inequality and its applications to linear convergence of first-order methods. *Foundations of computational mathematics*, 18(5):1199–1232, 2018.
- [37] M. H. Li, K. W. Meng, and X. Q. Yang. On error bound moduli for locally Lipschitz and regular functions. *Mathematical Programming*, 171:463–487, 2018.
- [38] W. Li. Error bounds for piecewise convex quadratic programs and applications. *SIAM Journal on Control and Optimization*, 33(5):1510–1529, 1995.
- [39] H. Liu, A. M.-C. So, and W. Wu. Quadratic optimization with orthogonality constraint: explicit Lojasiewicz exponent and linear convergence of retraction-based line-search and stochastic variance-reduced gradient methods. *Mathematical Programming*, 178(1-2):215–262, 2019.
- [40] H. Liu, W. Wu, and A. M.-C. So. Quadratic optimization with orthogonality constraints: explicit Lojasiewicz exponent and linear convergence of line-search methods. In *ICML*, pages 1158–1167, 2016.
- [41] S. Lojasiewicz. Une propriété topologique des sous-ensembles analytiques réels. In *Les Équations aux Dérivées Partielles*, pages 87–89, Éditions du centre National de la Recherche Scientifique, Paris 1963.
- [42] S. Lojasiewicz. Sur la géométrie semi- et sous-analytique. *Annales de l’institut Fourier*, 43:1575–1595, 1993.
- [43] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, 1996.

- [44] Z.-Q. Luo and J. F. Sturm. Error bounds for quadratic systems. In *High performance optimization*, pages 383–404. Springer, 2000.
- [45] Z.-Q. Luo and P. Tseng. On the linear convergence of descent methods for convex essentially smooth minimization. *SIAM Journal on Control and Optimization*, 30(2): 408-425, 1992.
- [46] Z.-Q. Luo and P. Tseng. Error bound and convergence analysis of matrix splitting algorithms for the affine variational inequality problem. *SIAM Journal on Optimization*, 2(1): 43-54, 1992.
- [47] Z.-Q. Luo and P. Tseng. Error bounds and convergence analysis of feasible descent methods: a general approach. *Annals of Operations Research*, 46(1):157–178, 1993.
- [48] J. M. Martínez. Local minimizers of quadratic functions on Euclidean balls and spheres. *SIAM Journal on Optimization*, 4(1):159–176, 1994.
- [49] J. J. Moré. Generalizations of the trust region problem. *Optimization Methods and Software*, 2(3-4):189–209, 1993.
- [50] J. J. Moré and D. C. Sorensen. Computing a trust region step. *SIAM Journal on Scientific and Statistical Computing*, 4(3):553–572, 1983.
- [51] Y. Nesterov. *Introductory Lectures on Convex Programming: a Basic Course*. Kluwer, Boston, 2004.
- [52] J.-S. Pang. Error bounds in mathematical programming. *Mathematical Programming*, 79(1-3):299–332, 1997.
- [53] R. Poliquin and R. T. Rockafellar. Prox-regular functions in variational analysis. *Transactions of the American Mathematical Society*, 348(5):1805–1838, 1996.
- [54] T. K. Pong and H. Wolkowicz. The generalized trust region subproblem. *Computational Optimization and Applications*, 58(2):273–322, 2014.
- [55] F. Rendl and H. Wolkowicz. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Mathematical Programming*, 77(1):273–299, 1997.
- [56] A. M.-C. So and Z. Zhou. Non-asymptotic convergence analysis of inexact gradient methods for machine learning without strong convexity. *Optimization Methods and Software*, 32(4):963–992, 2017.
- [57] J. Wang and Y. Xia. A linear-time algorithm for the trust region subproblem based on hidden convexity. *Optimization Letters*, 11(8):1639–1646, 2017.
- [58] J. Wang and Y. Xia. Closing the gap between necessary and sufficient conditions for local nonglobal minimizer of trust region subproblem. *SIAM Journal on Optimization*, 30(3):1980-1995, 2020.
- [59] T. Wang and J.-S. Pang. Global error bounds for convex quadratic inequality systems. *Optimization*, 31(1):1–12, 1994.
- [60] V. A. Yakubovich. S-Procedure in nonlinear control theory. *Vestnik Leningrad University*, 1:62–77, 1971.
- [61] Y. Yuan. Recent advances in trust region algorithms. *Mathematical Programming*, 151(1):249–281, 2015.
- [62] H. Zhang, A. R. Conn, and K. Scheinberg. A derivative-free algorithm for least-squares minimization. *SIAM Journal on Optimization*, 20(6):3555–3576, 2010.



- [63] H. Zhang, A. Milzarek, Z. Wen, and W. Yin. On the geometric analysis of a quartic-quadratic optimization problem under a spherical constraint. *Mathematical Programming*, <https://doi.org/10.1007/s10107-021-01702-6>, 2021.
- [64] L.-H. Zhang and C. Shen. A nested Lanczos method for the trust-region subproblem. *SIAM Journal on Scientific Computing*, 40(4):2005–2032, 2018.
- [65] L.-H. Zhang, C. Shen, and R.-C. Li. On the generalized Lanczos trust-region method. *SIAM Journal on Optimization*, 27(3):2110–2142, 2017.
- [66] Z. Zhou and A. M.-C. So. A unified approach to error bounds for structured convex optimization problems. *Mathematical Programming*, 165(2):689–728, 2017.
- [67] Z. Zhou, Q. Zhang, and A. M.-C. So.  $\ell_{1,p}$ -norm regularization: Error bounds and convergence rate analysis of first-order methods. In *ICML*, pages 1501–1510, 2015.