ON CONVERGENCE OF THE GENERALIZED LANCZOS TRUST-REGION METHOD FOR TRUST-REGION SUBPROBLEMS

BO FENG AND GANG WU

Abstract. The generalized Lanczos trust-region (GLTR) method is one of the most popular approaches for solving large-scale trust-region subproblem (TRS). Recently, Z. Jia et al. [Z. Jia and F. Wang, SIAM J. Optim., 31 (2021), pp. 887–914] considered the convergence of this method and established some a priori error bounds on the residual, the solution and the Lagrange multiplier. In this paper, we revisit the convergence of the GLTR method and try to improve these bounds. First, we establish a sharper upper bound on the residual. Second, we give a new bound on the distance between the approximation and the exact solution, and show that the convergence of the approximation has nothing to do with the spectral separation. Third, we present some non-asymptotic bounds for the convergence of the Lagrange multiplier, and define a factor that plays an important role on the convergence of the Lagrange multiplier. Fourth, we revisit the convergence of the Krylov subspace method for the cubic regularization variant of trust-region subproblem, and substantially improve the convergence result established in [X. Jia, X. Liang, C. Shen and L. Zhang, SIAM J. Matrix Anal. Appl. 43 (2022), pp. 812–839] on the multiplier. Numerical experiments demonstrate the effectiveness of our theoretical results.

1. Introduction

In this paper, we are interested in the following large-scale trust-region subproblem (TRS) [7, 25]

\[
\min_{\|x\|_2 \leq \Delta} \left\{ f(x) = \frac{1}{2} x^T A x + x^T g \right\},
\]

where \( A \in \mathbb{R}^{n \times n} \) is a symmetric matrix, \( g \in \mathbb{R}^n \) is a nonzero vector, and \( \Delta > 0 \) is the trust-region radius.

The TRS (1) arises from many problems such as Tikhonov regularization of the ill-posed problem [27, 28, 29], graph partitioning problems [16], and the Levenberg–Marquardt algorithm for solving nonlinear least squares problems [25]. Moreover, solving TRS is a key step in trust-region methods for dealing with general nonlinear optimization problems [7, 25]. Some regularization variations of the TRS such as
the cubic regularization variant of TRS [24] were consider in [2, 3, 4, 5, 6, 18, 21, 24], which can be applied to solve unconstrained minimization problems.

There are many methods have been proposed for solving (1). For instance, the Moré–Sorensen [23] method is an efficient approach for small and dense TRS. For large-scale problem, Gould et al. modified the Moré–Sorensen method by using the Taylor series [12], and Steihaug et al. [34, 37] solved the TRS (1) via a truncated conjugate gradient (tCG) method. The generalized Lanczos trust-region method (GLTR) [11] solves (1) by using a projection method. A nested Lanczos method was proposed in [39], which can be viewed as an accelerated GLTR method. A matrix-free method was presented in [27, 28], and a semi-definite programming (SDP) based method was proposed in [26].

In [9], Gander, Golub, and von Matt solved (1) from an eigenproblem point of view. More precisely, the TRS (1) was rewritten as a standard eigen value problem of size $2n$. In [1], Adachi et al. generalized this strategy to a weighted-norm TRS which reduces to a generalized eigenvalue problem of size $2n$. Moreover, Adachi et al. showed that the Lagrange multiplier corresponding to the optimal solution is the rightmost eigenvalue of a matrix pair of size $2n$-by-$2n$, and the optimal solution can be computed from the eigenvector associated with the rightmost eigenvalue. One refers to [7, 8, 15, 25, 33, 36] and the references therein for more methods on the TRS (1).

Indeed, the generalized Lanczos trust-region method (GLTR) [11] is one of the most popular approaches for solving large-scale TRS (1). The convergence of GLTR was investigated in [3, 4, 20, 38]. In this paper, we denote by $x_{opt}$ an optimal solution to TRS (1), and by $\lambda_{opt}$ the Lagrangian multiplier associated with the solution $x_{opt}$. Let $x_k$ and $\lambda_k$ be the approximate solutions of $x_{opt}$ and $\lambda_{opt}$ obtained from the $k$-th step of the GLTR method, respectively. Recently, some upper bounds on $f(x_k) - f(x_{opt})$ and $\| (A + \lambda_k I)x_k + g \|$ were established in [3, 4] and [13]. In [38], Zhang et al. presented some upper bounds on $f(x_k) - f(x_{opt})$ and $\| x_{opt} - x_k \|_2$. The results showed that the convergence of the GLTR method is similar to that of the conjugate gradient (CG) method. Recently, Z. Jia et al. [38] established some a priori error bounds on $|\lambda_{opt} - \lambda_k|$, $(A + \lambda_k I)x_k + g$, $\sin \angle(x_{opt}, x_k)$, as well as $f(x_k) - f(x_{opt})$. The convergence of the Krylov subspace method for cubic regularization variant of TRS was analyzed in [3, 4, 13, 18].

In this paper, we revisit the convergence of the GLTR method, and try to refine the bounds due to Z. Jia et al. [38]. Moreover, we substantially improve some results of X. Jia et al. [18] and Gould et al. [13] on the Krylov subspace method for cubic regularization variant of TRS.

The contributions of this work are as follows:

- First, for the residual $\| (A + \lambda_k I)x_k + g \|$, we establish a sharper upper bound than those of Z. Jia et al. [20, Theorem 5.7] and Golud et al. [13, Theorem 3.4]. Our bound is $O(\kappa^{\frac{1}{2}})$ times smaller than that of Z. Jia et al., where $\kappa$ is the spectral condition number of $A + \lambda_{opt} I$.
- Second, for the approximate solution $x_k$, we give a new bound on $\sin \angle(x_{opt}, x_k)$, and show that the convergence of $x_k$ has nothing to do with the spectral separation $\text{sep}(\lambda_{opt}, C_k)$. As $\text{sep}(\lambda_{opt}, C_k)$ can be very small or even close to zero in practice, our new upper bound is much better than some existing ones.
• Third, for the approximate multiplier $\lambda_k$, the error bound on $\lambda_{opt} - \lambda_k$ is asymptotic [20, Theorem 4.11], i.e., it holds when $k$ is sufficiently large. In this paper, we present some non-asymptotic bounds that hold for all $k \geq 1$. In particular, we show that the factor $s(\lambda_{opt})$ defined in (60) plays an important role on the convergence of $\lambda_k$. To the best of our knowledge, it seems that this is the first analysis from this point of view.

• Fourth, we consider the convergence of the Krylov subspace method for solving the cubic regularization variant (70) of the TRS (1). We substantially improve the convergence result on the multiplier $\mu_{opt}$ due to X. Jia et al. [18].

This paper is organized as follows. In Section 2, we present some preliminaries. In Section 3 and Section 4, we give insight into the convergence of the GLTR method, and establish some refined upper bounds on $\| (A + \lambda_k I)x_k + g \|$, $\sin \angle(x_{opt}, x_k)$ and $\lambda_{opt} - \lambda_k$, respectively. In Section 5, we revisit the convergence of the Krylov subspace method for the cubic regularization variant of the TRS (1). Numerical experiments are performed in Section 6, which illustrate the effectiveness and sharpness of the proposed new upper bounds. Some concluding remarks are given in Section 7.

Throughout this paper, we denote by $(\cdot)^T$ the transpose of a matrix or vector, by $\mathcal{N}(A)$ the null space of a matrix $A$, by $\mathcal{R}(A)$ the range space of a matrix $A$, by $\| \cdot \|$ the Euclidean norm of a matrix or vector. Let $\mathcal{W}$ be a linear subspace of $\mathbb{R}^n$, and let $W$ be an orthonormal basis of $\mathcal{W}$. Then the distance between a nonzero vector $p$ and $W$ is defined as [31]

$$
\sin \angle(p, W) = \min_{x \in \mathcal{W}} \frac{\| p \| - \| x \|}{\| p \|},
$$

and the sine of the angle between a nonzero vector $p$ and $q$ is defined as [31]

$$
\sin \angle(p, q) = \min_{\omega \in \mathbb{C}} \left\| \frac{p}{\| p \|} - \omega q \right\|.
$$

Let $0$, $O$ and $I$ be the zero vector, zero matrix and identity matrix, respectively, whose sizes are clear from the context. Let $e_i$ be the $i$-th column of the identity matrix $I$. In this paper, $A \succeq O$ ($A \succ O$) implies that $A$ is symmetric semi-positive definite (positive definite), and $A \succeq B$ ($A \succ B$) stands for $A - B$ is a symmetric semi-positive definite (positive definite) matrix.

2. Preliminaries

In this section, we briefly introduce the solutions to the subproblem (1), the generalized Lanczos trust-region (GLTR) method [11], and some existing convergence results on the GLTR method.

2.1. Solutions of the subproblems. Let $A = U\Lambda U^T$ be the eigendecomposition, where $U = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^{n \times n}$ is orthonormal, and

$$
\Lambda = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)
$$

with $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ being the eigenvalues of $A$ in decreasing order. The following result provides a necessary and sufficient condition for a global optimal solution of TRS (1).
Theorem 2.1. [7, 23, 32] The vector $x_{opt}$ is a global optimal solution of the trust-region problem (1), if and only if $\|x_{opt}\| \leq \Delta$ and there exists Lagrange multiplier $\lambda_{opt} \geq 0$ such that

$$
(A + \lambda_{opt}I)x_{opt} = -g, \quad \lambda_{opt}(\Delta - \|x_{opt}\|) = 0 \quad \text{and} \quad A + \lambda_{opt}I \succ O.
$$

In particular, there are two situations for the TRS (1) [15, 23, 25]:

Case 1. Easy case: $\lambda_{opt} > -\alpha_n$ and $\lambda_{opt} \geq 0$. In this case, $A + \lambda_{opt}I \succ O$, the solution $x_{opt}$ for TRS (1) is unique and $x_{opt} = -(A + \lambda_{opt}I)^{-1}g$.

Case 2. Hard case: $\lambda_{opt} = -\alpha_n$. In this case, we have

$$
g \perp N(A - \alpha_nI) \quad \text{and} \quad \|(A - \alpha_nI)^{\dagger}g\| \leq \Delta,
$$

where the superscript $\dagger$ denotes the Moore-Penrose inverse of a matrix.

2.2. The generalized Lanczos trust-region (GLTR) method. As $A$ is a symmetric matrix, one can use the $k$-step Lanczos process to generate an orthonormal basis $Q_k = [q_1, q_2, \ldots, q_k]$ for the Krylov subspace $K_k(A, g) = \text{span} \{g, Ag, \ldots, A^{k-1}g\}$. Moreover, the following Lanczos relation holds [30, 31]

$$
AQ_k = Q_kT_k + \beta_k \cdot q_{k+1}e_k^T,
$$

where $e_k$ is the $k$-th column of the identity matrix, and

$$
T_k = Q_k^T AQ_k = \begin{pmatrix} 
\delta_1 & \beta_1 & \beta_2 & \cdots & \beta_{k-1} & \beta_k \\
\beta_1 & \delta_1 & \beta_2 & \cdots & \beta_{k-1} & \beta_k \\
\beta_2 & \beta_1 & \delta_1 & \cdots & \beta_{k-1} & \beta_k \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\beta_{k-1} & \beta_{k-1} & \beta_{k-1} & \cdots & \delta_1 & \beta_k \\
\beta_k & \beta_k & \beta_k & \cdots & \delta_1 & \delta_k 
\end{pmatrix} \in \mathbb{R}^{k \times k},
$$

with $g = \|g\|q_1$ and $Q_k^T g = \|g\|e_1$.

The GLTR method due to Gould et al. [11] is one of the most popular approaches for solving the large-scale TRS problem (1). The GLTR method solves the following TRS problem

$$
\min_{\|x\| \leq \Delta, \ x \in K_k(A, g)} \left\{ f(x) = \frac{1}{2}x^T Ax + x^T g \right\},
$$

which leads to the following reduced TRS problem

$$
\min_{\|h\| \leq \Delta} \left\{ f_k(h) = \frac{1}{2}h^T T_k h + \|g\| \cdot h^T e_1 \right\}.
$$

Let $h_k$ be the minimizer of (6). It can be verified that

$$
x_k = Q_k h_k = \min_{\|x\| \leq \Delta, \ x \in K_k(A, g)} f(x),
$$

which can be used as an approximation to $x_{opt}$.

In [20, 38], it was shown that $\|x_k - x_{opt}\| \to 0$ as $k$ increase. Since (6) is also a trust-region subproblem, by Theorem 2.1, we have that $\|h_k\| \leq \Delta$ and there exists a Lagrange multiplier $\lambda_k \geq 0$, such that

$$
(T_k + \lambda_k I)h_k = -\|g\| e_1, \quad \lambda_k (\Delta - \|h_k\|) = 0 \quad \text{and} \quad T_k + \lambda_k I \succ O.
$$

So we have from (4) that

$$
(A + \lambda_k I)x_k = (A + \lambda_k I)Q_k h_k = Q_k(T_k + \lambda_k I)h_k + \beta_k q_{k+1}e_k^T h_k
$$

$$
= -\|g\| \cdot Q_k e_1 + (\beta_k e_k^T h_k) \cdot q_{k+1} = -g + (\beta_k e_k^T h_k) \cdot q_{k+1}.
$$
As a result, GLTR is an orthogonal projection method satisfying

\[ \begin{align*}
& x_k \in K_k(A, g), \\
& r_k = (A + \lambda_k I)x_k + g \perp K_k(A, g),
\end{align*} \]

which is a key property utilized in our analysis; see Section 3. Notice that \( T_k \) is a tridiagonal and irreducible matrix, the TRS (6) is also in the easy case [11, Theorem 5.3], and the solution is unique

\[ h_k = -\|g\|(T_k + \lambda_k I)^{-1}e_1 \] and \( x_k = -\|g\|Q_k(T_k + \lambda_k I)^{-1}e_1. \]

Let \( k_{\text{max}} \) be the first number of iteration at which the symmetric Lanczos process breaks down, then we have \( x_{k_{\text{max}}} = x_{\text{opt}}, \lambda_{k_{\text{max}}} = \lambda_{\text{opt}} \) and \( f(x_{k_{\text{max}}}) = f(x_{\text{opt}}) \) [11, 20, 38]. We notice that the sequence \( \{\lambda_k\}_{k=0}^{k_{\text{max}}} \) of Lagrangian multipliers associated with (6) is monotonically nondecreasing [22]

\[ 0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{k_{\text{max}}} = \lambda_{\text{opt}}. \]

2.3. Some existing results. Denote by

\[ A_{\text{opt}} = A + \lambda_{\text{opt}} I, \quad \kappa = \frac{\alpha_1 + \lambda_{\text{opt}}}{\alpha_n + \lambda_{\text{opt}}} \quad \text{and} \quad \epsilon_k = \min_{x \in K_k(A, g)} \|x - x_{\text{opt}}\|. \]

Then we have from [20, Theorem 4.10] that

\[ \epsilon_k = \| (I - Q_kQ_k^T) x_{\text{opt}} \| \leq 2\Delta \cdot \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k. \]

Notice that this bound is slightly worse but is more concise than the one given in [38, eq. (4.27)]. Recently, Z. Jia et al. [20] established the following bound on \( \|(A + \lambda_k I)x_k + g\|\):

**Theorem 2.2.** [20, Theorem 5.7] Suppose that \( \|x_{\text{opt}}\| = \|x_k\| = \Delta, k = 1, 2, \ldots, k_{\text{max}}, \) and define \( r_k = (A + \lambda_k I)x_k + g. \) Then asymptotically

\[ \|r_k\| \leq 4\sqrt{\kappa}\|A_{\text{opt}}\|\Delta \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k + \left( \frac{4\eta_1 \Delta^2}{\|g\|} + 8\|A_{\text{opt}}\|\eta_2 \Delta^3 \right) \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} \]

with

\[ \eta_1 = \frac{\|g\|^2}{\Delta^2 + \|g\|^2 \|e_1\|^2 (T_k + \lambda_{\text{opt}} I)^{-2} e_1} \leq \frac{\|g\|^2\|A_{\text{opt}}\|^2}{\|g\|^2 + \|A_{\text{opt}}\|^2 \Delta^2}, \]

\[ \eta_2 = \frac{2}{\Delta^2 + \|g\|^2 \|e_1\|^2 (T_k + \lambda_{\text{opt}} I)^{-2} e_1} \leq \frac{2\|A_{\text{opt}}\|^2}{\|g\|^2 + \|A_{\text{opt}}\|^2 \Delta^2}. \]

In addition, Golub et al. [13] established another upper bound on \( \|r_k\|:

**Theorem 2.3.** [13, Theorem 3.4] The residual \( r_k = (A + \lambda_k I)x_k + g \) for the \( k \)-th iterate, \( x_k, \) generated by the TRS (5) satisfies the bound

\[ \|r_k\| \leq \|g\| \left( \frac{2\beta_k \kappa_k}{\|T_k + \lambda_k I\|} \right) \left( \frac{\sqrt{\kappa_k} - 1}{\sqrt{\kappa_k} + 1} \right)^{k-1}, \]

where \( \kappa_k \) is the 2-condition number of \( T_k + \lambda_k I \) and \( \beta_k \) is the \((k, k+1)\)-st entry of \( T_{k+1}. \)

Z. Jia et al. established the following asymptotic bound on the convergence of \( \lambda_k. \) That is, the bound holds only when \( k \) is sufficiently large.
Theorem 2.4. [20, Theorem 4.11] Suppose that $\|x_k\| = \|x_{opt}\| = \Delta$, $k = 1, 2, \ldots, k_{\text{max}}$. Then asymptotically

$$
0 \leq \lambda_{opt} - \lambda_k \leq \left( \frac{4\eta_1 \Delta}{\|g\|} + 8\|A_{opt}\| \eta_2 \Delta^2 \right) \left( \frac{\sqrt{k} - 1}{\sqrt{k} + 1} \right)^{2k},
$$

where the constants $\eta_1$ and $\eta_2$ are defined by (14).

Remark 2.1. Theorem 2.4 provides an “asymptotic” bound, whose proof requires $k$ is sufficiently large [20, pp. 902]. Therefore, it is interesting to establish “non-asymptotic” bounds that hold in each step. In Section 4, we establish a “non-asymptotic” bound which are closely related to the value of $s(\lambda_{opt})$ defined in (60).

It was shown that (1) can be treated by solving the eigenvalue problem on the $2n$-by-$2n$ matrix [1, 9, 20]:

$$
M = \begin{pmatrix} -A & \frac{2n^2 T}{\Delta^2} \\ I & -A \end{pmatrix} \in \mathbb{R}^{2n \times 2n}.
$$

The following theorem establishes an important relationship between the TRS solution $(\lambda_{opt}, x_{opt})$ and the rightmost eigenpair of $M$.

Theorem 2.5. [20] Let $(\lambda_{opt}, x_{opt})$ satisfy Theorem 2.1 with $\|x_{opt}\| = \Delta$. Then the rightmost eigenvalue $\varpi_1$ of $M$ is real and simple, and $\lambda_{opt} = \varpi_1$. Let $y = (y_1^T, y_2^T)^T$ be the corresponding unit length eigenvector of $M$ with $y_1, y_2 \in \mathbb{R}^n$, and suppose that $g^Ty_2 \neq 0$. Then the unique TRS solution is

$$
x_{opt} = -\frac{\Delta^2}{g^Ty_2} y_1.
$$

Remark 2.2. It was shown that $g^Ty_2 = 0$ only if TRS is in hard case [1, Proposition 4.1]. In other words, we have $g^Ty_2 \neq 0$ in easy case.

In [1], Adachi et al. proved that $\lambda_{opt}$ is the rightmost real eigenvalue of $M$. However, the proof of the uniqueness of $\lambda_{opt}$ was not given. For completeness, we give a proof here.

Theorem 2.6. Suppose that $\|x_{opt}\| = \Delta$. Then trust-region subproblem (1) is in easy case if and only if $\lambda_{opt}$ is a simple eigenvalue of $M$.

Proof. On one hand, if TRS (1) is in easy case, then $A_{opt} = A + \lambda_{opt} I \succ O$, i.e., $\lambda_{opt} \in (-\alpha_n, \infty)$. From [1, eq. (14)–eq. (16)], we have

$$
\det(M - \lambda I) = \det \begin{pmatrix} -I & A + \lambda I \\ A + \lambda I & -\frac{2n^2 T}{\Delta^2} \end{pmatrix} = (-1)^n \det(A + \lambda I)^2 \left( 1 - \frac{\|A + \lambda I\|^{-1} g\|^2}{\Delta^2} \right),
$$

where $\lambda \in (-\alpha_n, \infty)$. Notice that $A_{opt} \succ O$ and

$$
\|(A + \lambda I)^{-1} g\|^2 = g^T(A + \lambda I)^{-2} g = \sum_{i=1}^{n} \frac{(u_i^T g)^2}{(\alpha_i + \lambda)^2}.
$$

Therefore,

$$
\frac{d\det(M - \lambda I)}{d\lambda} \bigg|_{\lambda = \lambda_{opt}} = (-1)^{n+2} \cdot \frac{2 \det(A_{opt})^2}{\Delta^2} \cdot \left( \sum_{i=1}^{n} \frac{(u_i^T g)^2}{(\alpha_i + \lambda_{opt})^3} \right) \neq 0.
$$

That is, $\lambda_{opt}$ is a simple root of $\det(M - \lambda I) = 0$, i.e., $\lambda_{opt}$ is a simple eigenvalue of $M$.
On the other hand, if \( \lambda_{\text{opt}} \) is a simple eigenvalue of \( M \), we next prove that TRS (1) is in easy case. Recall that \( \alpha_n \) is the smallest eigenvalue of \( A \). Suppose that the multiplicity of \( \alpha_n \) is \( s \), with \( s > 1 \). Let

\[
U_1 = [u_1, u_2, \ldots, u_{n-s}] \quad \text{and} \quad U_2 = [u_{n-s+1}, u_{n-s}, \ldots, u_n].
\]

If TRS (1) is in hard case, i.e., \( \lambda_{\text{opt}} = -\alpha_n \), then it follows that \( U_2^T g = 0 \) and \( \| (A - \alpha_n I) g \| \leq \Delta \). Define

\[
p(\lambda) = - \sum_{i=1}^{n-s+1} \frac{u_i u_i^T g}{\alpha_i + \lambda}, \quad \forall \lambda \in [-\alpha_n, \infty).
\]

By [1, eq. (14)–eq. (16)],

\[
\begin{align*}
\det(M - \lambda I) &= \frac{(-1)^n}{\Delta^2} \det \begin{pmatrix}
\Delta^2 - ||p(\lambda)||^2 & -p(\lambda)^T & g^T + p(\lambda)^T (A + \lambda I) \\
p(\lambda) & -I & A + \lambda I \\
g + (A + \lambda I)p(\lambda) & A + \lambda I & O
\end{pmatrix}.
\end{align*}
\]

As \( U_2^T g = 0 \), we have \( U_1 U_1^T g = g \), and

\[
g + (A + \lambda I)p(\lambda) = g - \sum_{i=1}^{n-s+1} u_i u_i^T g = g - U_1 U_1^T g = U_2 U_2^T g = 0.
\]

In addition, if \( \lambda > -\alpha_n \), then \( p(\lambda) \in \mathcal{R}(A + \lambda I) = \mathbb{R}^n \). If \( \lambda = -\alpha_n \), then

\[
p(\lambda) = p(-\alpha_n) \in \mathcal{R}(U_1) = \mathcal{R}(A - \alpha_n I) = \mathcal{R}(A + \lambda I).
\]

Thus, for any \( \lambda \in [-\alpha_n, \infty) \), there is a vector \( q(\lambda) \) such that \( p(\lambda) = (A + \lambda I)q(\lambda) \).

As a result,

\[
\begin{align*}
\det(M - \lambda I) &= \det(A + \lambda I)^2 \left[ 1 - \frac{||p(\lambda)||^2}{\Delta^2} \right] = (\alpha_n + \lambda)^2 \prod_{i=1}^{n-s+1} (\alpha_i + \lambda)^2 \left[ 1 - \frac{||p(\lambda)||^2}{\Delta^2} \right].
\end{align*}
\]

Therefore, \( \lambda_{\text{opt}} = -\alpha_n \) is a multiple root of \( \det(M - \lambda I) = 0 \). In other words, \( \lambda_{\text{opt}} \) is a multiple eigenvalue of \( M \), a contradiction.

Denote by

\[
\bar{Q}_k = \begin{pmatrix} Q_k \\ Q_k^T \end{pmatrix} \in \mathbb{R}^{2n \times 2k} \quad \text{and} \quad M_k = \bar{Q}_k^T M \bar{Q}_k = \begin{pmatrix} -T_k & -\frac{q_k^2}{\Delta} e_1 e_1^T \\ I_k & -T_k \end{pmatrix}.
\]

Suppose that \( ||x_k|| = \Delta \). As \( \lambda_k \) is the Lagrange multiplier of TRS (6), it follows from Theorem 2.5 that \( \lambda_k \) is the rightmost eigenvalue of \( M_k \). Let \( (\lambda_k, z_k) \) be the rightmost eigenpair of \( M_k \), where \( z_k = ((z_1^{(k)})^T, (z_2^{(k)})^T)^T \), with \( z_1^{(k)}, z_2^{(k)} \in \mathbb{R}^k \) and \( ||z_k|| = 1 \). Then

\[
y_k = \bar{Q}_k z_k = (Q_k z_1^{(k)})^T, (Q_k z_2^{(k)})^T)^T = ((y_1^{(k)})^T, (y_2^{(k)})^T)^T
\]

is a Ritz vector of \( M \) in the subspace \( \bar{S}_k = \mathcal{R}(\bar{Q}_k) \), which is an approximation to \( y \).
Lemma 3.1. \( \text{span} \{ z_k \} \), and let \( C_k = Z_k^T M_k Z_k \). If \( \lambda_{\text{opt}} \) is not an eigenvalue of \( C_k \), we can define
\[
\text{sep}(\lambda_{\text{opt}}, C_k) = \| (\lambda_{\text{opt}} I - C_k)^{-1} \|^{-1}.
\]
The following theorem depicts the distance between the approximation \( x_k \) and the optimal solution \( x_{\text{opt}} \):

**Theorem 2.7.** [20, Theorem 5.4] Suppose that \( \| x_{\text{opt}} \| = \| x_k \| = \Delta \) and \( \text{sep}(\lambda_{\text{opt}}, C_k) > 0 \). Then
\[
\sin \angle(x_{\text{opt}}, x_k) \leq c_k \cdot \left( 1 + \frac{\| M \|}{\sqrt{1 - \sin^2 \angle(y, z_k) \cdot \text{sep}(\lambda_{\text{opt}}, C_k)}} \right) \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k,
\]
where
\[
c_k = 2 + \frac{16\| A_{\text{opt}} \|}{(\alpha_1 - \alpha_n)^2(1-t^2)} \left( 1 + \frac{k + 2}{\ln t} \right) \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^2 \text{ with } t = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}.
\]

**Remark 2.3.** From the differential mean value theorem, there is a constant \( \nu \in [\sqrt{\kappa} - 1, \sqrt{\kappa} + 1] \), such that
\[
(1 - t) \ln t = \frac{2}{\sqrt{\kappa} + 1} \left( \ln(\sqrt{\kappa} + 1) - \ln(\sqrt{\kappa} - 1) \right) = \frac{4}{(\sqrt{\kappa} + 1) \nu} \leq \frac{4}{\kappa - 1},
\]
and thus \( \frac{1}{(1-t) \ln t} = O(\kappa) \). Moreover, we have that
\[
\frac{1}{(\alpha_1 - \alpha_n)} \cdot \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} = \frac{\kappa}{\| A_{\text{opt}} \| (\sqrt{\kappa} + 1)^2} = O\left( \frac{1}{\| A_{\text{opt}} \|} \right).
\]
Therefore, it follows from (27) and (28) that \( c_k = O\left( \frac{k \kappa}{\| A_{\text{opt}} \| \text{sep}(\lambda_{\text{opt}}, C_k)} \right) \). In summary,
\[
\tilde{c}_k = c_k \cdot \left( 1 + \frac{\| M \|}{\sqrt{1 - \sin^2 \angle(y, z_k) \cdot \text{sep}(\lambda_{\text{opt}}, C_k)}} \right) = O\left( \frac{k \kappa}{\text{sep}(\lambda_{\text{opt}}, C_k)} \right),
\]
i.e.,
\[
\sin \angle(x_{\text{opt}}, x_k) \leq O\left( \frac{k \kappa}{\text{sep}(\lambda_{\text{opt}}, C_k)} \right) \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k.
\]
In other words, (26) shows that the upper bound of \( \sin \angle(x_{\text{opt}}, x_k) \) is closely related to \( \text{sep}(\lambda_{\text{opt}}, C_k), \kappa \) and \( k \). However, there is no guarantee that \( \text{sep}(\lambda_{\text{opt}}, C_k) \) is uniformly lower bounded, both in theory and in practice. Consequently, the righthand side of (26) can be a poor estimation to the distance between \( x_k \) and \( x_{\text{opt}} \).

3. Improved Upper Bounds on \( \|(A + \lambda_k I)x_k + g\| \) and \( \sin \angle(x_k, x_{\text{opt}}) \)

In this section, we give some new upper bounds on \( \|(A + \lambda_k I)x_k + g\| \) and \( \sin \angle(x_k, x_{\text{opt}}) \). Let \( P_i \) be the set of polynomials whose degrees are no higher than the positive integer \( i \). Let \( C_i(x) \in P_i \) be the Chebyshev polynomial define as
\[
C_i(x) = \begin{cases} 
\cos(i \arccos x) & \text{ for } |x| < 1, \\
\frac{1}{2}((x + \sqrt{x^2 - 1})^i + (x + \sqrt{x^2 - 1})^{-i}) & \text{ for } |x| \geq 1.
\end{cases}
\]
To establish our bounds, we need the following result.

**Lemma 3.1.** Given a matrix \( H > O \). Let \( \kappa_H = \frac{\lambda_{\text{max}}(H)}{\lambda_{\text{min}}(H)} \) be the 2-condition number of \( H \).
Theorem 3.1. Suppose that $\varphi_i \in \mathcal{P}_i$ satisfies
\begin{equation}
1 - x_{\varphi_i}(x) = 2(r^{i+1} + r^{-(i+1)})^{-1} \cdot (C_{i+1} \left( \frac{\kappa_H + 1}{\kappa_H - 1} \right) - \frac{2x}{\lambda_{\max}(H) - \lambda_{\min}(H)}),
\end{equation}
where $r = \sqrt{\frac{\kappa_H - 1}{\kappa_H + 1}}$. Then, we have that $\|I - H_{\varphi_i}(H)\| \leq 2r^{i+1}$.

(ii) Suppose that $H$ is symmetric and $[\lambda_{\min}(H), \lambda_{\max}(H)] \subseteq [\lambda_{\min}(H), \lambda_{\max}(H)]$, then $\|I - H_{\varphi_i}(H)\| \leq 2r^{i+1}$.

Proof. We only need to prove (ii). Let $H = P \Gamma P^T$ be the eigendecomposition, where $P$ is orthonormal, and $\Gamma = \text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_m)$ with $\lambda_{\max}(\tilde{H}) = \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_m = \lambda_{\min}(\tilde{H})$ being the eigenvalues of $\tilde{H}$. Thus,
\begin{equation}
\|I - \tilde{H}_{\varphi_i}(\tilde{H})\| = \|I - \Gamma_{\varphi_i}(\Gamma)\| = \max_{\ell=1,\ldots,m} |1 - \gamma_{\ell}\varphi_i(\gamma_{\ell})| = 2(r^{i+1} + r^{-(i+1)})^{-1} \cdot \max_{\ell=1,\ldots,m} \left| C_{i+1} \left( \frac{\kappa_H + 1}{\kappa_H - 1} \right) - \frac{2\gamma_{\ell}}{\lambda_{\max}(H) - \lambda_{\min}(H)} \right|,
\end{equation}
As $\gamma_{\ell} \in [\lambda_{\min}(H), \lambda_{\max}(H)]$, we have that
\begin{equation}
\left| \frac{\kappa_H + 1}{\kappa_H - 1} \right| - \frac{2\gamma_{\ell}}{\lambda_{\max}(H) - \lambda_{\min}(H)} \right| \leq 1,
\end{equation}
and it follows from (30) that $\varphi \leq 1$. Hence, $\|I - \tilde{H}_{\varphi_i}(\tilde{H})\| \leq 2(r^{i+1} + r^{-(i+1)})^{-1} \leq 2r^{i+1}$, which completes the proof. \qed

We are ready to establish some improved bounds on $\| (A + \lambda_k I) x_k + g \|$ and $\sin \angle(x_{\text{opt}}, x_k)$:

Theorem 3.1. Suppose that $\|x_{\text{opt}}\| = \|x_k\| = \Delta$. Then
\begin{equation}
\| (A + \lambda_k I) x_k + g \| \leq \min \{ \xi_1, \xi_2 \},
\end{equation}
\begin{equation}
\sin \angle(x_{\text{opt}}, x_k) \leq 2 \sqrt{\xi_1} \cdot \left( \frac{\sqrt{\kappa - 1}}{\sqrt{\kappa + 1}} \right)^k,
\end{equation}
where
\begin{equation}
\xi_1 = 2\|A_{\text{opt}}\| \sqrt{\Delta^2 + \epsilon^2} \cdot \left( \frac{\sqrt{\kappa - 1}}{\sqrt{\kappa + 1}} \right)^k \quad \text{and} \quad \xi_2 = 2\beta_k \Delta \cdot \left( \frac{\sqrt{\kappa_k - 1}}{\sqrt{\kappa_k + 1}} \right)^{k-1}.
\end{equation}

Proof. First, for any $\|x\| = \Delta$, we have that
\begin{align}
0 & \leq f(x) - f(x_{\text{opt}}) = \frac{1}{2} (x^T A x - x_{\text{opt}}^T A x_{\text{opt}}) + g^T (x - x_{\text{opt}}) \\
& = \frac{1}{2} (x^T A_{\text{opt}} x - x_{\text{opt}}^T A_{\text{opt}} x_{\text{opt}}) - x_{\text{opt}}^T A_{\text{opt}} (x - x_{\text{opt}}) \\
& = \frac{1}{2} (x_{\text{opt}}^T x_{\text{opt}} - x^T A_{\text{opt}} x_{\text{opt}}).
\end{align}
Let $y_k = \Delta \cdot \frac{Q_k x_{\text{opt}}}{\|Q_k x_{\text{opt}}\|} \in \mathcal{K}_k(A, g)$, then it follows from (7) and (34) that
\begin{equation}
f(x_k) - f(x_{\text{opt}}) \leq f(y_k) - f(x_{\text{opt}}) \leq \frac{1}{2} \|A_{\text{opt}}\| \|x_{\text{opt}} - y_k\|^2.
\end{equation}
Moreover,
\[
\|x_{\text{opt}} - y_k\|^2 = \|(I - Q_kQ_k^T)x_{\text{opt}} + Q_kQ_k^T x_{\text{opt}} - y_k\|^2
\]
\[
= \|(I - Q_kQ_k^T)x_{\text{opt}}\|^2 + \|Q_kQ_k^T x_{\text{opt}} - y_k\|^2
\]
\[
= \|(I - Q_kQ_k^T)x_{\text{opt}}\|^2 + \|Q_kQ_k^T x_{\text{opt}}\|^2 + \|y_k\|^2 - 2y_k^T Q_kQ_k^T x_{\text{opt}}
\]
\[
= \|(I - Q_kQ_k^T)x_{\text{opt}}\|^2 + (\Delta - \|Q_kQ_k^T x_{\text{opt}}\|)^2 = \varepsilon_k^2 + (\Delta - \|Q_kQ_k^T x_{\text{opt}}\|)^2.
\]

Now we consider \(\Delta - \|Q_kQ_k^T x_{\text{opt}}\|.\) As \(\|x_{\text{opt}}\| = \Delta\) and \(\|Q_kQ_k^T x_{\text{opt}}\| \leq \|x_{\text{opt}}\| = \Delta\), we have that
\[
\Delta - \|Q_kQ_k^T x_{\text{opt}}\| = \frac{\Delta^2 - \Delta \|Q_kQ_k^T x_{\text{opt}}\|^2}{\Delta} \leq \frac{\Delta^2 - \|Q_kQ_k^T x_{\text{opt}}\|^2}{\Delta} = \frac{\|(I - Q_kQ_k^T)x_{\text{opt}}\|^2}{\Delta} = \frac{\varepsilon_k^2}{\Delta}.
\]
As a result,
\[
(36) \quad 0 \leq f(x_k) - f(x_{\text{opt}}) \leq \frac{1}{2}\|A_{\text{opt}}\| \left( 1 + \frac{\varepsilon_k^2}{\Delta^2} \right) \|x_{\text{opt}}\|^2 \leq 2\|A_{\text{opt}}\| \left( \Delta^2 + \varepsilon_k^2 \right) \left( \frac{\sqrt{k} - 1}{\sqrt{k} + 1} \right)^{2k}.
\]

Notice that
\[
(37) \quad A_{\text{opt}}(x_k - x_{\text{opt}}) = (A + \lambda_{\text{opt}}I)x_k + g = (A + \lambda_kI)x_k + g + (\lambda_{\text{opt}} - \lambda_k)x_k.
\]
By (9),
\[
\|A_{\text{opt}}(x_k - x_{\text{opt}})\| = \left\| (A + \lambda_{\text{opt}}I)x_k + g \right\| + (\lambda_{\text{opt}} - \lambda_k)^2 \left\| x_k \right\|^2,
\]
then
\[
\|(A + \lambda_{\text{opt}}I)x_k + g\| \leq \|A_{\text{opt}}(x_k - x_{\text{opt}})\|^2 \leq \|A_{\text{opt}}\|^2 \cdot \|A_{\text{opt}}(x_k - x_{\text{opt}})\|^2
\]
\[
(38) = \|A_{\text{opt}}\| \cdot \|(x_k - x_{\text{opt}})^T A_{\text{opt}}(x_k - x_{\text{opt}})\| \leq 2\|A_{\text{opt}}\| \cdot (f(x_k) - f(x_{\text{opt}})) \leq \xi_k^2.
\]
Second, for any \(\varphi_{k-2} \in P_{k-2}\), we have from (8) that
\[
\|(A + \lambda_{\text{opt}}I)x_k + g\| = \beta_k \|e_k^T h_k\| \leq \beta_k \left( \|e_k^T (I - (T_k + \lambda_{\text{opt}}I)\varphi_{k-2}(T_k + \lambda_{\text{opt}}I))h_k\| + \|e_k^T (T_k + \lambda_{\text{opt}}I)\varphi_{k-2}(T_k + \lambda_{\text{opt}}I)e_1\| \right).
\]
Recall that \(\varphi_{k-2} \in P_{k-2}\) and \(T_k + \lambda_{\text{opt}}I\) is a symmetric and tridiagonal matrix. Thus, \(e_k^T \varphi_{k-2}(T_k + \lambda_{\text{opt}}I)e_1 = 0\) for any \(\varphi_{k-2} \in P_{k-2}\). Consequently,
\[
e_k^T (T_k + \lambda_{\text{opt}}I) \cdot \varphi_{k-2}(T_k + \lambda_{\text{opt}}I)h_k = -\gamma_k \|e_k^T (I - (T_k + \lambda_{\text{opt}}I)\varphi_{k-2}(T_k + \lambda_{\text{opt}}I)h_k\| + \|e_k^T (T_k + \lambda_{\text{opt}}I)\varphi_{k-2}(T_k + \lambda_{\text{opt}}I)e_1\| = -\|g\| \|e_k^T \varphi_{k-2}(T_k + \lambda_{\text{opt}}I)e_1\| = 0.
\]
Specially, if we choose \(\varphi_{k-2}\) that satisfies (31), it then follows from Lemma 3.1 that
\[
(39) \quad \|A + \lambda_{\text{opt}}I\| x_k + g \| \leq \beta_k \|h_k\| \|(I - (T_k + \lambda_{\text{opt}}I)\varphi_{k-2}(T_k + \lambda_{\text{opt}}I))\| \leq \xi_2.
\]
A combination of the above equation with (38) gives (32).

Finally, we notice that
\[
(\sin \angle(x_{\text{opt}}, x_k))^2 = \min_{\chi \in \mathbb{C}} \frac{\|x_{\text{opt}} - \chi \cdot x_k\|^2}{\Delta^2} = \frac{x_{\text{opt}} - x_k}{\Delta} \frac{x_{\text{opt}} - x_k}{\Delta} = \frac{\|x_{\text{opt}} - x_k\|^2}{\Delta^2} \leq \frac{(x_{\text{opt}} - x_k)^T A_{\text{opt}}(x_{\text{opt}} - x_k)}{\Delta^2 (\alpha_n + \lambda_{\text{opt}})} \leq \frac{(f(x_k) - f(x_{\text{opt}}))}{\Delta^2 (\alpha_n + \lambda_{\text{opt}})} \leq \frac{4k \left( 1 + \frac{\varepsilon_k^2}{\Delta^2} \right) \left( \frac{\sqrt{k} - 1}{\sqrt{k} + 1} \right)^{2k}}{\Delta^2 (\alpha_n + \lambda_{\text{opt}})}.
\]
which completes the proof.

Remark 3.1. First, the upper bound established in (13) is about \( \sqrt{\kappa} \) times larger than \( \xi_1 \), and our new upper bound (32) improves the one given in (13), especially when the condition number \( \kappa \) is large. Moreover, we notice that

\[
\Delta = \|x_k\| = \|h_k\| = \|g\| \| (T_k + \lambda_k I)^{-1} e_1 \| \leq \| (T_k + \lambda_k I)^{-1} \| \| g \| = \frac{\| g \|^\kappa}{\| T_k + \lambda_k I \|},
\]

and our bound (32) is no worse than (15). Indeed, compared with the bounds (13) and (15), in our new upper bound (32), there are no condition numbers such as \( \kappa \) or \( \kappa_k \) in the coefficients before \( \left( \frac{\sqrt{\kappa}}{\sqrt{\kappa + 1}} \right)^{k-1} \) and \( \left( \frac{\sqrt{\kappa}}{\sqrt{\kappa + 1}} \right)^k \), respectively. Hence, we improved the upper bounds of Z. Jia et al. and Gould et al. substantially.

Second, our new upper bound (33) improved (26) substantially. More precisely, compared with (26), (33) indicates that the factor before \( \left( \frac{\sqrt{\kappa}}{\sqrt{\kappa + 1}} \right)^k \) is in the order of \( \sqrt{\kappa} \), which is free of \( \text{sep}(\lambda_{\text{opt}}, C_k) \) and \( k \). In other words, Theorem 2.7 shows that the convergence of \( x_k \) is influenced by the factor \( \text{sep}(\lambda_{\text{opt}}, C_k) \) [20], while our bound (33) reveals that it has nothing to do with \( \text{sep}(\lambda_{\text{opt}}, C_k) \) at all.

4. Non-asymptotic error estimates for \( \lambda_{\text{opt}} - \lambda_k \)

In this section, we focus on the convergence of \( \lambda_k \). To do this, we first need the following three lemmas.

Lemma 4.1. Let \( t = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa + 1}} \), and denote by

\[
g_i(x) = 2 \left( t^i + t^{-i} \right)^{-1} C_i \left( \frac{\kappa + 1}{\kappa - 1} - \frac{2t}{\alpha_1 - \alpha_n} \right),
\]

where \( C_i(\cdot) \) is defined in (30). Then we have that

\[
0 \leq -\frac{dg_i}{dx} \bigg|_{x=0} \leq \frac{i \sqrt{\kappa}}{\| A_{\text{opt}} \|} \quad \text{and} \quad \frac{d^2 g_i}{dx^2} \bigg|_{x=0} \geq 0, \quad i = 1, 2, \ldots
\]

Proof. Let \( \ell = \frac{\alpha_1 - \alpha_n}{2} \), then

\[
\frac{dg_i}{dx} \bigg|_{x=0} = \frac{4 \left( t^i + t^{-i} \right)^{-1}}{\alpha_1 - \alpha_n} \cdot \frac{dC_i}{dx} \bigg|_{x=\ell}.
\]

Since

\[
\left( x - \sqrt{x^2 - 1} \right)^i \left( 1 - \frac{x}{\sqrt{x^2 - 1}} \right) \leq \left( x + \sqrt{x^2 - 1} \right)^i \left( 1 + \frac{x}{\sqrt{x^2 - 1}} \right), \quad \forall x > 1,
\]

we have that

\[
0 \leq \frac{dC_i}{dx} = \frac{i}{2} \left( x + \sqrt{x^2 - 1} \right)^{i-1} \left( 1 + \frac{x}{\sqrt{x^2 - 1}} \right) + \left( x - \sqrt{x^2 - 1} \right)^{i-1} \left( 1 - \frac{x}{\sqrt{x^2 - 1}} \right)
\]

\[
\leq \frac{i}{2} \left( x + \sqrt{x^2 - 1} \right)^{i-1} \left( 1 + \frac{x}{\sqrt{x^2 - 1}} \right), \quad \forall x > 1.
\]

Moreover,

\[
\ell + \sqrt{\ell^2 - 1} = \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1}, \quad 1 + \frac{\ell}{\sqrt{\ell^2 - 1}} = \frac{(\sqrt{\kappa} + 1)^2}{2\sqrt{\kappa}} \quad \text{and} \quad \left( t^i + t^{-i} \right)^{-1} \leq t^i.
\]
So it follows from (42) that
\[ 0 \leq (t^i + t^{-i})^{-1} \left. \frac{dC_i}{dx} \right|_{x=t} \leq \frac{i}{4} \left( \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa}} \right)^2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right). \]

By (41),
\[ 0 \leq \left. \frac{dg_i}{dx} \right|_{x=0} \leq \frac{i}{\alpha_1 - \alpha_n} \left( \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa}} \right)^2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right) = \frac{i \sqrt{\kappa}}{\|A_{\text{opt}}\|}, \]
where the last equality is from (28).

In addition, it follows from (41) that
\[ \left. \frac{d^2 C_i}{dx^2} \right|_{x=0} = \frac{8}{(\alpha_1 - \alpha_n)^2} \left( t^i + t^{-i} \right)^{-1} \left. \frac{dC_i}{dx} \right|_{x=t}. \]

From [19, eq. (13)], we obtain
\[ \left. \frac{d^2 C_i}{dx^2} \right|_{x=t} = \sum_{s=0}^{i-2} \vartheta_s C_s(t) \text{ with } \vartheta_s \geq 0, \ s = 0, 1, \ldots, \ i - 2. \]

As \( \ell > 1 \), we have \( C_i(\ell) \geq 0 \), and \( \left. \frac{d^2 C_i}{dx^2} \right|_{x=0} \geq 0. \) \hfill \Box

**Lemma 4.2.** For \( j = 1, 2, \ldots, 2k - 2 \), we have that
\[ g^T A_{\text{opt}}^j g = \|g\|^2 \cdot e_1^T (T_k + \lambda_{\text{opt}} I)^j e_1. \]

**Proof.** We denote by \( T_{k,\text{opt}} = T_k + \lambda_{\text{opt}} I \) for the sake of simplicity. For \( j = 1, 2, \ldots, k - 1 \), we have from [13, eq. (A.9)] that
\[ A_{\text{opt}}^j Q_k = Q_k T_{k,\text{opt}}^j + \beta_k \sum_{i=0}^{j-1} A_{\text{opt}}^{j-i-1} q_{k+1} \cdot e_k^T T_{k,\text{opt}}^i. \]

As \( T_{k,\text{opt}} \) is a symmetric and tridiagonal matrix, \( e_k^T T_{k,\text{opt}}^i e_1 = 0 \) for \( 0 \leq i \leq k - 2 \). Recall that \( Q_k e_1 = \frac{g}{\|g\|} \), so we obtain
\[
\begin{align*}
Q_k T_{k,\text{opt}}^j e_1 & = A_{\text{opt}}^j Q_k e_1 = A_{\text{opt}}^j \frac{g}{\|g\|}, \\
e_k^T T_{k,\text{opt}}^j e_1 & = e_1^T Q_k^T A_{\text{opt}}^j Q_k e_1 = \frac{g^T A_{\text{opt}}^j g}{\|g\|^2},
\end{align*}
\]

For \( j = k, k+1, \ldots, 2k - 2 \), we have \( j - k + 1 = 1, 2, \ldots, k - 1 \), and
\[
\frac{g^T A_{\text{opt}}^j g}{\|g\|^2} = e_1^T Q_k^T A_{\text{opt}}^j Q_k e_1 = (A_{\text{opt}}^{j-k} Q_k e_1)^T A_{\text{opt}}^{k-1} Q_k e_1 = (Q_k T_{k,\text{opt}}^{j-k-1} e_1)^T Q_k T_{k,\text{opt}}^{k-1} e_1 = e_1^T T_{k,\text{opt}} e_1,
\]
where the third equality follows from the first equation of (44). \hfill \Box

Let
\[ h_k = -\|g\|(T_k + \lambda_{\text{opt}} I)^{-1} e_1 \text{ and } \bar{x}_k = Q_k h_k. \]

Notice that
\[ A_{\text{opt}} \bar{x}_k + g = A_{\text{opt}} Q_k h_k + \|g\| Q_k e_1 \quad (4) = (\beta_k e_1^T h_k) \cdot q_{k+1} \perp K_k(A, g), \]
then we have that
Lemma 4.3. Let \( \bar{x}_k \) be defined in (45), \( \|x_{opt}\|_{A_{opt}} = \sqrt{x_{opt}^T A_{opt} x_{opt}} \) and \( k = 1, 2, \ldots, k_{\text{max}} \). Then

\[
\|x_{opt}\|^2 - \|\bar{x}_k\|^2 \leq \frac{4}{\Delta^2} \left( 1 + \frac{2k \sqrt{\kappa}}{\|A_{opt}\|} \left( \frac{\|x_{opt}\|_{A_{opt}}}{\Delta} \right)^2 \right) \left( \sqrt{\kappa} - 1 \right)^{2k}. \tag{47} \]

Proof. From (59) and the facts that \( \lambda_{opt} \geq \lambda_k \) and \( T_k + \lambda_{opt} I = Q_k^T A_{opt} Q_k \geq O \), we obtain

\[
0 \leq \|x_{opt}\|^2 - \|\bar{x}_k\|^2 = T_{opt} x_{opt} - \bar{x}_k^T \bar{x}_k = g^T A_{opt}^{-2} g - \|g\|^2 + e_1^T T_{k_{opt}}^{-2} e_1, \tag{48} \]

where \( T_{k_{opt}} = T_k + \lambda_{opt} I \). Suppose that \( \varphi_{2k-1}(x) \in P_{2k-1} \) satisfies (31), and let

\[
p_{2k-2}(x) = \frac{\varphi_{2k-1}(x) - \tilde{a}_0}{x} \in P_{2k-2}, \text{ with } x \in [\alpha_n + \lambda_{opt}, \alpha_1 + \lambda_{opt}],
\]

where \( \tilde{a}_0 \) is the constant term of \( \varphi_{2k-1}(x) \). According to Lemma 4.2,

\[
g^T p_{2k-2}(A_{opt}) g = \|g\|^2 \cdot e_1^T p_{2k-2}(T_{k_{opt}}) e_1,
\]

so we obtain

\[
\frac{g^T A_{opt}^{-2} g}{\|g\|^2} - e_1^T T_{k_{opt}}^{-2} e_1 = \frac{g^T [A_{opt}^{-2} - p_{2k-2}(A_{opt})] g}{\|g\|^2} - e_1^T [T_{k_{opt}}^{-2} - p_{2k-2}(T_{k_{opt}})] e_1. \tag{49} \]

Now we consider \( g^T [A_{opt}^{-2} - p_{2k-2}(A_{opt})] g \). By the definition of \( p_{2k-2}(x) \), we get

\[
= x_{opt}^T [I - A_{opt}^{-2} p_{2k-2}(A_{opt})] x_{opt} - \tilde{a}_0 g^T A_{opt}^{-2} g
\]

\[
= x_{opt}^T [I - A_{opt} \varphi_{2k-1}(A_{opt})] x_{opt} - \tilde{a}_0 g^T x_{opt}, \tag{50} \]

Similarly, we can prove that

\[
e_1^T [T_{k_{opt}}^{-2} - p_{2k-2}(T_{k_{opt}})] e_1 = \tilde{h}_k^T [I - T_{k_{opt}} \varphi_{2k-1}(T_{k_{opt}})] \tilde{h}_k - \tilde{a}_0 e_1^T \tilde{h}_k \tag{51} \]

where \( \tilde{h}_k \) is defined in (45).

From the Poincaré separation theorem [17, Corollary 4.3.37], we have \( \alpha_n + \lambda_{opt} \leq \lambda_{\min}(T_k) + \lambda_{opt} \leq \lambda_{\max}(T_k) + \lambda_{opt} \leq \alpha_1 + \lambda_{opt} \). Hence,

\[
\|I - T_{k_{opt}} \varphi_{2k-1}(T_{k_{opt}})\| \leq 2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k}. \tag{52} \]

It follows from (49)–(52) and Lemma 3.1 (i) that

\[
g^T A_{opt}^{-2} g - \|g\|^2 \cdot e_1^T (T_k + \lambda_{opt} I)^{-2} e_1
\]

\[
= x_{opt}^T [I - A_{opt} \varphi_{2k-1}(A_{opt})] x_{opt} - \|g\|^2 \cdot \tilde{h}_k^T [I - T_{k_{opt}} \varphi_{2k-1}(T_{k_{opt}})] \tilde{h}_k
\]

\[+ \tilde{a}_0 (\|g\| \cdot e_1^T \tilde{h}_k - g^T x_{opt}) \]

\[
\leq 4 \Delta^2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} + \tilde{a}_0 (\|g\| \cdot e_1^T \tilde{h}_k - g^T x_{opt}), \tag{53} \]

where the last inequality is from (52) and \( \|\tilde{h}_k\| = \|\bar{x}_k\| \leq \Delta \); refer to (48).
Moreover, by [14, Theorem 3.1.1],
\[
\|g\|e_1^T \tilde{T}_k - g^T x_{opt} = \|g\|e_1^T Q_k^T Q_k \tilde{T}_k - g^T x_{opt} = g^T \bar{x}_k - g^T x_{opt} = -x_{opt}^T (A_{opt} \bar{x}_k + g)
\]
(54a)
\[
\Leftrightarrow (\bar{x}_k - x_{opt})^T (A_{opt} \bar{x}_k + g) = (\bar{x}_k - x_{opt})^T A_{opt} x_{opt} - x_{opt}^T x_{opt}
\]
(54b)
\[
\leq 4 \|x_{opt}\|^2_{A_{opt}} \cdot \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k}.
\]
Recall that $\bar{a}_0$ is the constant of $\varphi_{2k-1}(x)$ satisfying (31), then $-\bar{a}_0$ is the coefficient on the $x$ term of
\[
g_{2k}(x) = 1 - x \varphi_{2k-1}(x) = (t^{2k} + t^{-2k})^{-1} \cdot C_{2k} \left( \frac{\kappa + 1}{\kappa - 1} - \frac{2x}{\alpha_1 - \alpha_0} \right),
\]
and $\bar{a}_0 = -\frac{dg_{2k}}{dx}|_{x=0}$. Combining (48), (53), (54b), Lemma 4.1, we complete the proof.

We are ready to establish the following nonasymptotic bound on $\lambda_{opt} - \lambda_k$.

**Theorem 4.1.** Suppose that $\|x_{opt}\| = \|x_k\| = \Delta$, $k = 1, 2, \ldots, k_{max}$. Then
\[
0 \leq \lambda_{opt} - \lambda_k \leq 2s_k \cdot \left( 1 + \frac{2k\sqrt{\Delta}}{\|x_{opt}\| \|A_{opt}\|} \left( \frac{\|x_{opt}\|}{\|A_{opt}\|} \Delta \right)^2 \right) \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k},
\]
where
\[
s_k = \frac{2}{\|g\|^2 \cdot e_1^T (T_k + \lambda_{opt} I)^{-3} e_1}.
\]

**Proof.** It follows from (10) and (45) that
\[
\|x_k\|^2 - \|\bar{x}_k\|^2 = \|g\|^2 \cdot e_1^T \left( (T_k + \lambda_k I)^{-2} - (T_k + \lambda_{opt} I)^{-2} \right) e_1
\]
(57)
\[
= \|g\|^2 \cdot e_1^T \left( (T_k + \lambda_{opt} I)^{-1} (T_k + \lambda_{opt} I)^{-1} \right) \left( (T_k + \lambda_k I)^{-1} - (T_k + \lambda_{opt} I)^{-1} \right) e_1.
\]
Notice that $\lambda_k \leq \lambda_{opt}$, so we have
\[
\begin{aligned}
& (T_k + \lambda_{opt} I)^{-1} - (T_k + \lambda_{opt} I)^{-1} = (\lambda_{opt} - \lambda_k) (T_k + \lambda_k I)^{-1} (T_k + \lambda_{opt} I)^{-1}, \\
& (T_k + \lambda_k I)^{-1} (T_k + \lambda_{opt} I)^{-1} = (T_k + \lambda_{opt} I)^{-1} (T_k + \lambda_{opt} I)^{-1} = 2 (T_k + \lambda_{opt} I)^{-3}.
\end{aligned}
\]
Therefore,
\[
2 (\lambda_{opt} - \lambda_k) \cdot e_1^T (T_k + \lambda_{opt} I)^{-3} e_1
\]
(58)
\[
\leq e_1^T \left( (T_k + \lambda_{opt} I)^{-1} - (T_k + \lambda_{opt} I)^{-1} \right) \left( (T_k + \lambda_k I)^{-1} + (T_k + \lambda_{opt} I)^{-1} \right) e_1.
\]
As $\|x_{opt}\| = \|x_k\|$, from (57)-(58), we obtain
\[
0 \leq \lambda_{opt} - \lambda_k \leq \frac{\|x_k\|^2 - \|\bar{x}_k\|^2}{2\|g\|^2 \cdot e_1^T (T_k + \lambda_{opt} I)^{-3} e_1} = \frac{\|x_{opt}\|^2 - \|\bar{x}_k\|^2}{2\|g\|^2 \cdot e_1^T (T_k + \lambda_{opt} I)^{-3} e_1}.
\]
A combination of Lemma 4.3 yields (55).\[ \square \]

Next, we focus on $s_k$ and prove that
\[
s_k = \frac{\Delta^2}{\|g\|^2 \cdot e_1^T (T_k + \lambda_{opt} I)^{-3} e_1} \rightarrow \frac{\Delta^2}{g^T A_{opt}^{-1} g} = s(\lambda_{opt})
\]
as $k$ increase, moreover, we have that $s_{k_{max}} = s(\lambda_{opt})$. It is only necessary to consider the denominator $\|g\|^2 \cdot e_1^T (T_k + \lambda_{opt} I)^{-3} e_1$.\[ \square \]
Theorem 4.2. For $k = 1, 2, \ldots, k_{\max}$, we have

$$0 \leq \frac{g^T A_{opt}^{-3} g - \|g\|^2 \cdot e_1^T (T_k + \lambda_{opt} I)^{-3} e_1}{g^T A_{opt}^{-3} g} \leq 4 (1 + \tau_k) \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k}$$

and $g^T A_{opt}^{-3} g = \|g\|^2 \cdot e_1^T (T_{k_{\max}} + \lambda_{opt} I)^{-3} e_1$, where

$$\tau_k = \frac{\Delta^2 (2k + 1) \sqrt{\kappa}}{\|g\|^2 \cdot \|A_{opt}\|} \left( 1 + \frac{2k \sqrt{\kappa}}{\|A_{opt}\|} \left( \frac{\|x_{opt}\| \cdot \|A_{opt}\|}{\Delta} \right)^2 \right).$$

Consequently, as $k$ increases, we have that

$$s_k \to s(\lambda_{opt}) \text{ and } s_{k_{\max}} = s(\lambda_{opt}).$$

Proof. From $A_{opt} Q_{k_{\max}} = Q_{k_{\max}} (T_{k_{\max}} + \lambda_{opt} I)$, we have $A_{opt}^{-3} Q_{k_{\max}} = Q_{k_{\max}} (T_{k_{\max}} + \lambda_{opt} I)^{-3}$. As $Q_{k_{\max}} e_1 = \frac{g}{\|g\|}$, we have from [10, Theorem 4.1] that

$$g^T A_{opt}^{-3} g - \|g\|^2 \cdot e_1^T (T_{k_{\max}} + \lambda_{opt} I)^{-3} e_1 = \|g\|^2 \cdot \left( e_1^T (T_{k_{\max}} + \lambda_{opt} I)^{-3} e_1 - e_1^T (T_{k_{\max}} + \lambda_{opt} I)^{-3} e_1 \right) \geq 0,$$

and $g^T A_{opt}^{-3} g = \|g\|^2 \cdot e_1^T (T_{k_{\max}} + \lambda_{opt} I)^{-3} e_1$, i.e., $s(\lambda_{max}) = s(\lambda_{opt})$, where $T_{k_{opt}} = T_k + \lambda_{opt} I$.

Suppose that $\varphi_{2k}(x) \in P_{2k}$ satisfies (31). Let

$$q_{2k-2}(x) = \frac{\varphi_{2k}(x) - \tilde{a}_1 x - \tilde{a}_0}{x^2} \in P_{2k-2}, \text{ with } x \in [\alpha_n + \lambda_{opt}, \alpha_1 + \lambda_{opt}],$$

where $\tilde{a}_0$ and $\tilde{a}_1$ are the constant term and the coefficient on the $x$ term of $\varphi_{2k}(x)$, respectively. By Lemma 4.2,

$$g^T A_{opt}^{-3} g - \|g\|^2 = \frac{g^T (A_{opt}^{-3} - q_{2k-2}(A_{opt})) g}{\|g\|^2} e_1^T (T_{k_{\max}} - q_{2k-2}(T_{k_{\max}})) e_1.$$

Now we consider $g^T (A_{opt}^{-3} - q_{2k-2}(A_{opt})) g$. From the definition of $q_{2k-2}(x)$, we obtain

$$\begin{align*}
g^T (A_{opt}^{-3} - q_{2k-2}(A_{opt})) g &= g^T A_{opt}^{-\frac{3}{2}} \left( I - A_{opt} q_{2k-2}(A_{opt}) \right) A_{opt}^{-\frac{3}{2}} g \\
&= g^T A_{opt}^{-\frac{3}{2}} \left( I - A_{opt} q_{2k-2}(A_{opt}) \right) A_{opt}^{-\frac{3}{2}} g + a_1 g^T A_{opt}^{-1} g + a_0 g^T A_{opt}^{-2} g \\
&= g^T A_{opt}^{-\frac{3}{2}} \left( I - A_{opt} \varphi_{2k}(A_{opt}) \right) A_{opt}^{-\frac{3}{2}} g - a_1 g^T x_{opt} + a_0 \|x_{opt}\|^2.
\end{align*}$$

Similarly, let $\tilde{h}_k$ be defined in (45), we arrive at

$$\begin{align*}
e_1^T (T_{k_{\max}} - q_{2k-2}(T_{k_{\max}})) e_1 \\
&= e_1^T (I - T_{k_{\max}} \varphi_{2k}(T_{k_{\max}})) T_{k_{\max}}^{-\frac{3}{2}} e_1 - a_1 \frac{e_1^T \tilde{h}_k}{\|g\|^2} + a_0 \frac{\|\tilde{h}_k\|^2}{\|g\|^2}.
\end{align*}$$
A combination of (64)–(66) yields
\[ g^T A_{opt}^{-3} g - \|g\|^2 \cdot e^T (T_k + \lambda_{opt} I)^{-1} e_1 = g^T A_{opt}^{-3} [I - A_{opt} \varphi_{2k}(A_{opt})] A_{opt}^{-\frac{3}{2}} g - \|g\|^2 e^T (T_{k, opt} - \frac{\hat{x}^T}{\hat{x}^T} I - T_{k, opt} \varphi_{2k}(T_{k, opt})) (T_{k, opt} - \frac{\hat{x}^T}{\hat{x}^T} I)^{-\frac{3}{2}} e_1 + \tilde{\alpha}_1 (\|g\| \cdot e^T \hat{h}_k - g^T x_{opt}) + \tilde{\alpha}_0 (\|x_{opt}\|^2 - \|\hat{h}_k\|^2) \]
\[ \leq g^T A_{opt}^{-3} g \cdot (\|I - A_{opt} \varphi_{2k}(A_{opt})\| + \|I - T_{k, opt} \varphi_{2k}(T_{k, opt})\|) + \tilde{\alpha}_1 (\|g\| \cdot e^T \hat{h}_k - g^T x_{opt}) + \tilde{\alpha}_0 (\|x_{opt}\|^2 - \|\hat{h}_k\|^2) \]
\[ \leq g^T A_{opt}^{-3} g \cdot (\|I - A_{opt} \varphi_{2k}(A_{opt})\| + \|I - T_{k, opt} \varphi_{2k}(T_{k, opt})\|) + \tilde{\alpha}_1 (\|g\| \cdot e^T \hat{h}_k - g^T x_{opt}) + \tilde{\alpha}_0 (\|x_{opt}\|^2 - \|\hat{h}_k\|^2) \]
\[ \leq 4 g^T A_{opt}^{-3} \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} + \tilde{\alpha}_1 (\|g\| \cdot e^T \hat{h}_k - g^T x_{opt}) + \tilde{\alpha}_0 (\|x_{opt}\|^2 - \|\hat{h}_k\|^2), \]
where we used \( \|\hat{h}_k\| = \|x_k\| \); see (45).
In terms of the definition of \( \varphi_{2k}(x) \), we see that \( -\tilde{\alpha}_0 \) and \( -\tilde{\alpha}_1 \) are the coefficients on the terms \( x \) and \( x^2 \) of
\[ \tilde{g}_{2k+1}(x) = 1 - x \varphi_{2k}(x) = (t^{2k+1} + t^{-(2k+1)})^{-1} C_{2k+1} \left( \frac{\kappa + 1}{\kappa - 1} - \frac{2x}{\alpha_1 - \alpha_n} \right), \]
respectively. Thus,
\[ \tilde{\alpha}_0 = -\frac{d \tilde{g}_{2k+1}}{dx} \bigg|_{x=0} \text{ and } \tilde{\alpha}_1 = -\frac{1}{2} \frac{d^2 \tilde{g}_{2k+1}}{dx^2} \bigg|_{x=0}. \]
Moreover, it follows from Lemma 4.1 and (54a) that
\[ 0 \leq \tilde{\alpha}_0 \leq \left( 2k+1 \right) \frac{\sqrt{\kappa}}{\|A_{opt}\|} \text{ and } \tilde{\alpha}_1 (\|g\| \cdot e^T \hat{h}_k - g^T x_{opt}) \leq 0. \]
A combination of (67), (68), and Lemma 4.3 yields (61).

**Remark 4.1.** In [20, eq. (4.9)], the condition number of \( \lambda_{opt} \) is defined as
\[ \text{cond}(\lambda_{opt}) = \frac{1}{2} \|y_1^T y_2\|^{-1}. \]
As \( x_{opt} = -\text{sign}(g^T y_2) \frac{\Delta y}{\|y_1\|} \) [1, eq. (21)] and \( M y = \lambda_{opt} y \), we have \( y_1 = A_{opt} y_2 \) and
\[ s(\lambda_{opt}) = \frac{\Delta^2}{g^T A_{opt}^{-1} g} = \frac{\Delta^2}{\|x_{opt}\|^2 A_{opt}^{-1} x_{opt}} = \frac{\|y_1\|^2}{\|y_1\|^2} = \frac{\|y_1\|^2}{\|y_1\|^2} \text{cond}(\lambda_{opt}). \]
Thus, \( s(\lambda_{opt}) \) is closely related to the condition number \( \text{cond}(\lambda_{opt}) \). However, we point out that \( s(\lambda_{opt}) \) can be much smaller than \( \text{cond}(\lambda_{opt}) \) in practice. This can happen, say, when \( \|y_1\| \ll 1 \) in the “nearly hard cases” [1]; see the numerical results in Table 2. That is, \( s(\lambda_{opt}) \) can be much different from the condition number \( \text{cond}(\lambda_{opt}) \), and it is more appropriate for depicting the convergence of \( \lambda_k \).

**Remark 4.2.** In summary, Theorem 4.1 and Theorem 4.2 show the importance of \( s(\lambda_{opt}) \) on the convergence of \( \lambda_k \). More precisely, \( \lambda_k \) may converge slow if \( s(\lambda_{opt}) \) is very large; see Figure 4. Compared with the result given in Theorem 2.4, our bound (55) is non-asymptotic, and there is no need to assume that \( k \) is sufficiently large.
5. On the convergence of Krylov subspace method for the cubic regularization problem

In this section, we consider the following cubic regularization problem, which is a regularization variant of the TRS (1) [2, 3, 4, 5, 6, 13, 18, 21, 24]:

\[
\min_{x \in \mathbb{R}^n} f_\sigma(x) \quad \text{with} \quad f_\sigma(x) = \frac{1}{2} x^T Ax + x^T g + \frac{\sigma}{3} \|x\|^3,
\]

where \(\sigma > 0\). Denote by \(\mu_{\text{opt}} = \sigma \|x_{\text{opt}}^\sigma\|\), and by

\[
\varepsilon_{\sigma,k} = \| (I - Q_k Q_k^T) x_{\text{opt}}^\sigma \| = \| (I - Q_k Q_k^T) x_{\text{opt}} \| \leq 2 \|x_{\text{opt}}^\sigma\| \cdot \left( \frac{\sqrt{\kappa_\sigma} - 1}{\sqrt{\kappa_\sigma} + 1} \right)^k,
\]

whose proof is similar to that of (12). We omit the details and refer to [20, 38].

The following theorem provides a necessary and sufficient condition for determining a global optimal solution of the subproblem (70).

**Theorem 5.1.** [5, Theorem 3.1] Any \(x_{\text{opt}}^\sigma\) is a global minimizer of \(f_\sigma(x)\) over \(\mathbb{R}^n\) if and only if it satisfies the system of equation

\[
(A + \mu_{\text{opt}} I) x_{\text{opt}}^\sigma = -g \quad \text{and} \quad A + \mu_{\text{opt}} I \succ 0.
\]

If \(A + \mu_{\text{opt}} I\) is positive definite, \(x_{\text{opt}}^\sigma\) is unique.

Suppose that \(A + \mu_{\text{opt}} I \succ 0\), then according to Theorem 5.1, \(x_{\text{opt}}^\sigma\) is unique and the cubic regularization problem (70) is in easy case [18, Definition 2.3]. Otherwise, one fails to get a good approximation to \(x_{\text{opt}}^\sigma\) from the Krylov subspace \(K_k(A, g)\) [5, 13]. In the Krylov subspace method, we solve the following problem [18]

\[
\min_{x \in K_k(A, g)} \left\{ f_\sigma(x) = \frac{1}{2} x^T Ax + x^T g + \frac{\sigma}{3} \|x\|^3 \right\}
\]

for the cubic regularized quadratic subproblem.

Let \(Q_k\) be an orthonormal basis for the Krylov subspace \(K_k(A, g)\), and let \(T_k\) be the triangular matrix obtained from the \(k\)-step Lanczos process; refer to (4). Let

\[
h_k^\sigma = \arg \min_{h \in \mathbb{R}^n} \left\{ \frac{1}{2} h^T T_k h + \|g\| h^T e_1 + \frac{\sigma}{3} \|h\|^3 \right\},
\]

which satisfies

\[
(T_k + \sigma \|h_k^\sigma\||I) h_k^\sigma = -\|g\| e_1 \quad \text{and} \quad T_k + \sigma \|h_k^\sigma\||I \succ 0.
\]

As \(T_k\) is an irreducible triangular matrix, by [11, Theorem 5.3] and [18, Definition 2.3], the cubic regularized quadratic subproblem (74) is also in easy case, where \(T_k + \sigma \|h_k^\sigma\||I \succ 0\) and \(h_k^\sigma = -\|g\| \cdot (T_k + \sigma \|h_k^\sigma\||I)^{-1} e_1\) is unique. Denote by \(x_k^\sigma = Q_k h_k^\sigma\), then it can be utilized as an approximation to \(x_{\text{opt}}^\sigma\). By [6, Theorem 1], we see that

\[
\|x_k^\sigma\| \leq \|x_2^\sigma\| \leq \cdots \leq \|x_{\text{opt}}^\sigma\| = \|x_{\text{opt}}^\sigma\|.
\]

Similar to (9), we have

\[
\begin{align*}
x_k^\sigma & \in K_k(A, g) \\
r_k^\sigma & = (A + \sigma \|x_k^\sigma\||I) x_k^\sigma + g \perp K_k(A, g).
\end{align*}
\]
In this section, we are interested in the convergence of the Krylov subspace method for the subproblem (70). Some upper bounds on \(|(A + \mu_k I)x_k^\sigma + g|\) and \(|\mu_k - \mu_{opt}|\) were established in \([13, 18]\), where \(\mu_k = \sigma\|x_k^\sigma\|\). We try to establish some sharper error bounds in this section.

5.1. Some existing results. The following result is due to Gould and Simoncini [13]:

**Theorem 5.2.** [13, Theorem 3.4] The residual \((A + \mu_k I)x_k^\sigma + g\) for the \(k\)-th iterate, \(x_k^\sigma\), generated by the regularization subproblem (73) satisfies the bound

\[
\| (A + \mu_k I)x_k^\sigma + g \| \leq \|g\| \left( \frac{2\beta_k \kappa_{\sigma,k}}{\|T_k + \mu_k I\|} \right) \left( \frac{\sqrt{\kappa_{\sigma,k} - 1}}{\sqrt{\kappa_{\sigma,k} + 1}} \right)^{k-1},
\]

where \(\kappa_{\sigma,k}\) is the 2-condition number of \(T_k + \mu_k I\) and \(\beta_k\) is the \((k,k+1)\)-st entry of \(T_k+1\).

Recently, X. Jia et al. [18] established the following bound on \(|\mu_k^3 - \mu_{opt}^3|\):

**Theorem 5.3.** [18, Theorem 4.6] Suppose that the cubic regularization problem (70) is in easy case. Then

\[
|\mu_k^3 - \mu_{opt}^3| \leq 6\sigma^2 \|g\| \sqrt{\kappa} \cdot \Pi_k + 12\sigma^2 \|A_{\sigma,\opt}\| \cdot \Pi_k^2,
\]

where

\[
\Pi_k = \frac{2\|g\| (\zeta_{\sigma} + \sqrt{\zeta_{\sigma}^2 - 1})^{-k}}{(\alpha_{\sigma} - \alpha_\sigma) (\zeta_{\sigma}^2 - 1)} \quad \text{with} \quad \zeta_{\sigma} = \frac{\kappa_{\sigma} + 1}{\kappa_{\sigma} - 1}.
\]

5.2. Improved upper bounds on \(|(A + \mu_k I)x_{\sigma,k} + g|\) and \(\mu_{opt} - \mu_k\). We are in a position to establish a new upper bound on \(|(A + \mu_k I)x_k^\sigma + g|\).

**Theorem 5.4.** Suppose that the cubic regularization problem (70) is in easy case. Then

\[
\| (A + \mu_k I)x_k^\sigma + g \| \leq \min \{\xi_{\sigma_1}, \xi_{\sigma_2}\},
\]

where

\[
\xi_{\sigma_1} = 2\|A_{\sigma,\opt}\| \sqrt{\|x_{\sigma,\opt}^\sigma\| \|x_{\sigma,k}^\sigma \| \left( \frac{\sqrt{\zeta_{\sigma} - 1}}{\sqrt{\zeta_{\sigma} + 1}} \right)^{k-1}} \quad \text{and} \quad \xi_{\sigma_2} = 2\beta_k \|x_k^\sigma\| \left( \frac{\sqrt{\kappa_{\sigma,k} - 1}}{\sqrt{\kappa_{\sigma,k} + 1}} \right)^{k-1},
\]

with \(\beta_k\) being the \((k,k+1)\)-st entry of \(T_k+1\) and \(\kappa_{\sigma,k}\) being the 2-condition number of \(T_k + \mu_k I\).

**Proof.** Without loss of generality, we suppose that \(x_{\sigma,\opt}^\sigma \neq 0\). Otherwise, we have from (76) that \(x_k^\sigma = 0\). Denote by

\[
\rho(x) = \frac{\mu_{opt}^2}{2} \cdot (\|x_{\sigma,\opt}^\sigma\|^2 - \|x\|^2) + \frac{\sigma^2}{3} (\|x\|^3 - \|x_{\sigma,\opt}^\sigma\|^3).
\]

Recall that \(\mu_{opt} = \sigma\|x_{\sigma,\opt}^\sigma\|\). Then

\[
\rho(x_k^\sigma) = \sigma \|x_{\sigma,\opt}^\sigma\| \left( \|x_{\sigma,\opt}^\sigma\|^2 - \|x_k^\sigma\|^2 \right) + \frac{\sigma^2}{3} (\|x_k^\sigma\|^3 - \|x_{\sigma,\opt}^\sigma\|^3) = \frac{\sigma}{6} (\|x_{\sigma,\opt}^\sigma\|^2 + \|x_{\sigma,\opt}^\sigma\| \|x_k^\sigma\| - 2 \|x_k^\sigma\|^2) (\|x_{\sigma,\opt}^\sigma\| - \|x_k^\sigma\|) \geq 0.
\]
For any $x$, we have that
\begin{equation}
0 \leq f_s(x) - f_s(x_{\sigma, opt}) = \frac{1}{2} \left( x^T A x - (x_{\sigma, opt})^T A x_{\sigma, opt} \right) + g^T (x - x_{\sigma, opt}) + \frac{\sigma}{3} \left( \|x\|^4 - \|x_{\sigma, opt}\|^4 \right)
= \frac{1}{2} \left( x^T A \sigma, opt x - (x_{\sigma, opt})^T A \sigma, opt x_{\sigma, opt} \right) - (x_{\sigma, opt})^T A \sigma, opt (x - x_{\sigma, opt}) + \rho(x)
\end{equation}
(82)
\begin{align*}
&= \frac{1}{2} (x - x_{\sigma, opt})^T A \sigma, opt (x - x_{\sigma, opt}) + \rho(x).
\end{align*}

Let $y_k = \frac{Q_k Q_k^T x_{\sigma, opt}}{\|Q_k Q_k^T x_{\sigma, opt}\|} \|x_{\sigma, opt}\|$. As $|g^T x_{\sigma, opt}| = (x_{\sigma, opt})^T A \sigma, opt x_{\sigma, opt} > 0$, we have $Q_k Q_k^T x_{\sigma, opt} \neq 0$, and $y_k$ is well-defined. Notice that $\rho(y_k) = 0$, so we have
\begin{align*}
0 &\leq f_s(x_k^\sigma) - f_s(x_{\sigma, opt}) \leq f_s(y_k) - f_s(x_{\sigma, opt}) \leq \frac{1}{2} \|A \sigma, opt\| \|y_k - x_{\sigma, opt}\|^2 \\
&\leq \frac{1}{2} \|A \sigma, opt\| \left( 1 + \frac{\epsilon_{2,k}^2}{\|x_{\sigma, opt}\|^2} \right) \epsilon_{2,k}^2,
\end{align*}
(83)
\leq 2 \|A \sigma, opt\| \left( \|x_{\sigma, opt}\|^2 + \epsilon_{2,k}^2 \right) \left( \frac{\sqrt{\kappa_{\sigma}} - 1}{\sqrt{\kappa_{\sigma}} + 1} \right)^{2k}.

From (37) and (77), we obtain
\begin{align*}
\| (A + \mu_k I) x_k^\sigma + g \| &= \|A \sigma, opt (x_k^\sigma - x_{\sigma, opt}) \|^2 - (\mu_{\sigma, opt} - \mu_k)^2 \|x_{\sigma, opt}\|^2 \\
&\leq \|A \sigma, opt (x_k^\sigma - x_{\sigma, opt})\|^2 \leq \|A \sigma, opt\| \|x_k^\sigma - x_{\sigma, opt}\|^2 \\
&\leq 2 \|A \sigma, opt\| \cdot \left( \frac{1}{2} (x_k^\sigma - x_{\sigma, opt})^T A \sigma, opt (x_k^\sigma - x_{\sigma, opt}) + \rho(x_k^\sigma) \right) \\
&\leq 2 \|A \sigma, opt\| \cdot \left( f_s(x_k^\sigma) - f_s(x_{\sigma, opt}) \right) \leq \epsilon_{1,2}^2.
\end{align*}
(84)

Recall that $T_k + \mu_k I > O$ is a symmetric and tridiagonal matrix, and $h_k^\sigma = -\|g\|(T_k + \mu_k I)^{-1} e_1$. Similar to the proof of Theorem 3.1, we have $\|(A + \mu_k I) x_k^\sigma + g\| \leq \xi_{\sigma,2}$. This completes the proof.

**Remark 5.1.** It follows from (75) that
\begin{align*}
\|x_k^\sigma\| = h_k^\sigma = \|g\| \|T_k + \mu_k I\|^{-1} e_1 \leq \|(T_k + \mu_k I)^{-1} g\| = \frac{\|g\|}{\|T_k + \mu_k I\|},
\end{align*}
so our bound (80) is no worse than (78). Moreover, in (80), there is no $\kappa_{\sigma, k}$ in the coefficient before $\left( \frac{\sqrt{\kappa_{\sigma} - 1}}{\sqrt{\kappa_{\sigma}} + 1} \right)^k$ and $\left( \frac{\sqrt{\kappa_{\sigma} + 1}}{\sqrt{\kappa_{\sigma}} - 1} \right)^k$.

Next, we consider the convergence of $\mu_k$.

**Theorem 5.5.** Suppose that the cubic regularization problem (70) is in easy case. Then
\begin{align*}
0 &\leq \mu_{opt} - \mu_k \leq 2s_{\sigma,k} \left( 1 + \frac{2k \sqrt{\kappa_{\sigma}}}{\|A \sigma, opt\|} \left( \frac{\|x_{\sigma, opt}\|}{\|x_{\sigma, opt}\|} \right)^2 \right) \left( \frac{\sqrt{\kappa_{\sigma} - 1}}{\sqrt{\kappa_{\sigma}} + 1} \right)^{2k},
\end{align*}
where $s_{\sigma,k} = \frac{\|x_{\sigma, opt}\|^2}{\|g\|^2 e_1 \cdot (1 + \mu_{opt} I)^{-1} e_1}$. 


Proof. Recall that $\mu_{\text{opt}} = \sigma \| x_{\text{opt}}^2 \|$ and $\mu_k = \sigma \| x_k^2 \|$ and $\| x_{\text{opt}}^2 \| \geq \| x_k^2 \|$. Hence, $\mu_{\text{opt}} \geq \mu_k$ and $T_k + \mu_{\text{opt}} I \succ T_k + \mu_k I \succ O$. Let $\bar{x}_k^2 = -\| g \| Q_k (T_k + \mu_{\text{opt}} I)^{-1} e_1$. From (77), (59), we have

$$0 \leq \mu_{\text{opt}} - \mu_k \leq \frac{\| x_k^2 \|^2 - \| \bar{x}_k^2 \|^2}{2\| g \|^2 \cdot e_1^T (T_k + \mu_{\text{opt}} I)^{-3} e_1} \leq \frac{\| x_{\text{opt}}^2 \|^2 - \| \bar{x}_k^2 \|^2}{2\| g \|^2 \cdot e_1^T (T_k + \mu_{\text{opt}} I)^{-3} e_1}.$$ 

The remaining proof is similar to that of Theorem 4.1. 

Corollary 5.1. Suppose that the cubic regularization problem (70) is in easy case. Then

$$0 \leq \mu_{\text{opt}} - \mu_k \leq 6s_{\sigma,k} \cdot \left( \sigma^2 \| x_{\text{opt}}^2 \|^2 + \frac{2\sigma^2 \| Q_k \| \| x_{\text{opt}}^2 \|^2}{\| A_{\sigma,\text{opt}} \|} \right) \left( \frac{\sqrt{2} - 1}{\sqrt{2} + 1} \right)^{2k}.$$ 

where $s_{\sigma,k} = \frac{\| x_{\text{opt}}^2 \|^2}{\| g \|^2 \cdot e_1^T (T_k + \mu_{\text{opt}} I)^{-3} e_1}.$

Proof. As $\mu_{\text{opt}} \geq \mu_k \geq 0$ and $\mu_{\text{opt}} = \sigma \| x_{\text{opt}}^2 \|$, we have that

$$\mu_{\text{opt}}^3 - \mu_k^3 = (\mu_{\text{opt}} - \mu_k) (\mu_{\text{opt}}^2 + \mu_{\text{opt}}^2 + \mu_{\text{opt}}^2) \leq 3\mu_{\text{opt}}^2 (\mu_{\text{opt}} - \mu_k) = 3\sigma^2 \| x_{\text{opt}}^2 \|^2 (\mu_{\text{opt}} - \mu_k),$$

and a combination of Theorem 5.5 yields the result. 

Remark 5.2. Theorem 5.3 shows that the upper bound of $| \mu_k^3 - \mu_{\text{opt}}^3 |$ is in the order of $\| k \| = \mathcal{O} \left( \frac{\sqrt{\sqrt{2} - 1}}{\sqrt{\sqrt{2} + 1}} \right)^{2k}$ [18]. As a comparison, Corollary 5.1 indicates that it is in the order of $\| k \| = \mathcal{O} \left( \frac{\sqrt{\sqrt{2} - 1}}{\sqrt{\sqrt{2} + 1}} \right)^{2k}$. Therefore, our new bound is much better than that of X. Jia et al.; see Figure 5.

6. Numerical experiments

In this section, we perform some numerical experiments to illustrate the effectiveness of our theoretical results. All the numerical experiments were run on a Inter(R) Core i5 with CPU 2.60GHz and RAM 4GB under Windows 7 operation system. The experimental results are obtained from using MATLAB R2014b implementation with machine precision $u \approx 2.22 \times 10^{-16}$.

Example 1. In this example, we compare our new bounds on $\| (A + \lambda_k I) x_k + g \|$, $\sin \beta(x_{\text{opt}}, x_k)$, $\lambda_{\text{opt}} - \lambda_k$ with those of Z. Jia et al. [20] and Gould et al. [13]. We consider the diagonal matrix $A = \text{diag}(t_{jn}^{[a,b]})$ [20, 38], with

$$t_{jn}^{[a,b]} = \left( \frac{b - a}{2} \right) \left( j_{jn} + \frac{a + b}{b - a} \right)$$

with $j_{jn} = \cos \left( \frac{2j - 1}{n} \pi \right)$, $j = 1, 2, \ldots, n$.

where $n = 10000$ and $t_{jn}^{[a,b]}$ is the $n$-th translated Chebyshev zero nodes on $[a, b]$. The vector $g$ is generated by the MATLAB build-in function $\text{randn.m}$, and is normalized with its Euclidean norm.

In the upper bounds for comparison, the “exact” values of $\lambda_{\text{opt}}$, $x_{\text{opt}}$ and $\kappa$ are computed by using the GLTR method [11], which are listed in Table 1 with different $[a, b]$ and $\Delta$. Figure 1–Figure 3 plot the curves of $\| (A + \lambda_k I) x_k + g \|$, $\sin \beta(x_{\text{opt}}, x_k)$, and $\lambda_{\text{opt}} - \lambda_k$, as well as the upper bounds for comparison as $k$ increases.

It is obvious to see from the figures that our new upper bounds are much sharper than those of Z. Jia et al. [20] and Gould et al. [13]. First, it is seen from Figure 1 that for the upper bounds $\| (A + \lambda_k I) x_k + g \|$, our bound (32) is about 10 times smaller than (15) (the one due to Gould et al.), and is about 1000 times smaller.
Table 1. Example 1: Approximations of $\lambda_{opt}$, $\|x_{opt}\|$ and $\kappa$ compute by using GLTR.

| $[a, b]$ | $\Delta$ | $\lambda_{opt}$ | $\|x_{opt}\|$ | $\|(A + \lambda_{opt}I)x_{opt} + g\|$ | $\kappa$ |
|---------|---------|----------------|----------------|--------------------------------|--------|
| $[-5, 5]$ | 1      | 5.2813         | 1.0000         | $3.4226 \times 10^{-12}$     | 36.5505 |
| $[-10, 10]$ | 10     | 10.0096        | 10.0000        | $9.0860 \times 10^{-12}$     | 2.0887 $\times 10^4$ |
| $[-50, 50]$ | 15     | 50.0032        | 15.0000        | $9.7802 \times 10^{-12}$     | 3.0794 $\times 10^4$ |
| $[-100, 100]$ | 20    | 100.0018       | 20.0000        | $3.9402 \times 10^{-11}$     | 1.1428 $\times 10^5$ |

Table 2. Example 1: A comparison of $s(\lambda_{opt}) = \frac{\Delta^2}{g^T A^{-1} g}$ and $\text{cond}(\lambda_{opt}) = \frac{1}{2} \|y_1^T y_2\|^{-1}$.

| $[a, b]$ | $\Delta$ | $s(\lambda_{opt})$ | $\text{cond}(\lambda_{opt})$ |
|---------|---------|----------------|----------------------------|
| $[-5, 5]$ | 1      | 0.3811         | 1.6073                     |
| $[-10, 10]$ | 10     | 0.0129         | 43.6256                    |
| $[-50, 50]$ | 15     | 0.0044         | 1.2807 $\times 10^2$      |
| $[-100, 100]$ | 20    | 0.0023         | 2.3780 $\times 10^2$      |

Figure 1. Example 1: A comparison of the upper bounds on $\|(A + \lambda_k I)x_k + g\|$.

than (13) (the one due to Z. Jia et al.), especially when the condition number $\kappa$ is large. Second, Figure 2 demonstrate that the new upper bound of $\sin \angle(x_{opt}, x_k)$
Figure 2. Example 1: A comparison of the upper bounds on $\sin \angle (\mathbf{x}_{\text{opt}}, \mathbf{x}_k)$.

- (a) $A = \text{diag}([t_{jn}^{-5}, 5])$ with $\Delta = 1$.
- (b) $A = \text{diag}([t_{jn}^{-10}, 10])$ with $\Delta = 10$.
- (c) $A = \text{diag}([t_{jn}^{-50}, 50])$ with $\Delta = 15$.
- (d) $A = \text{diag}([t_{jn}^{-100}, 100])$ with $\Delta = 20$.

is much smaller than the one due to Z. Jia et al. This is because $\text{sep}(\lambda_{\text{opt}}, C_k)$ is usually very small in practice, and the new upper bound (33) can be much smaller than (26); refer to (29). Furthermore, we observe from Figure 2 that the curve of the upper bound established by Z. Jia et al. is first up and then down as $k$ increases, especially when $\kappa$ is relatively large. This is because one has to first offset the value of $\tilde{c}_k$ (which is also increased on $k$, see (29)), by using $(\sqrt{\kappa - 1})^k$.

Third, we see from Figure 3 that for $\lambda_{\text{opt}} - \lambda_k$, our bound (55) is much better than (16) due to Z. Jia et al. However, it seems that the new upper bound is a little large at the beginning of the iterations. This is due to the fact that $\|g\|^2 \cdot e_1^T (T_k + \lambda_{\text{opt}} I)^{-3} e_1$, the denominator of $s_k$ (see (60)), can be small at the beginning. Note that this coincides with the trend of the convergence of $\lambda_{\text{opt}} - \lambda_k$.

Furthermore, it is shown from Theorem 4.1 that $s(\lambda_{\text{opt}})$ plays an important role in the convergence of $\lambda_k$. In Remark 4.1, we pointed out $s(\lambda_{\text{opt}})$ is closely related to the condition number $\text{cond}(\lambda_{\text{opt}})$, however, the former can be much smaller than that of the latter. In order to show this more precisely, we list in Table 2 the two values for different $[a, b]$ and $\Delta$. One observes that $s(\lambda_{\text{opt}})$ can be about $10^4$ times smaller than $\text{cond}(\lambda_{\text{opt}})$ in practice. One refers to Example 2 for more details on the importance of $s(\lambda_{\text{opt}})$.

Example 2. In this example, we try to show the importance of $s(\lambda_{\text{opt}})$.
on the convergence of \( \lambda_k \); refer to (60). Consider the matrix

\[ A_1 = \text{diag}(t_{jn}^{[a,b]}), \ j = 1, 2, \ldots, n, \ \text{with} \ a = 1, \ b = 3000, \ \text{and} \ n = 10000, \]

where \( t_{jn}^{[a,b]} \) is defined in (86). Let \( A_1 = \text{diag}(A_1, 0.01) \), and

\[ A = A_1 - 500 I, \ \text{and} \ f_\zeta = (\zeta, \ldots, \zeta, 1)^T \in \mathbb{R}^{10001}, \]

where \( I \) is identity matrix of appropriate size. We set

\[ \Delta = \| A_1 f_\zeta \| \ \text{and} \ g = A_1^2 f_\zeta. \]

Thus, for any \( \zeta \in \mathbb{R} \), we have that

\[ \lambda_{opt} = 500, \ x_{opt} = -A_1 f_\zeta, \ A_{opt} = A_1 \ \text{and} \ \kappa(A_{opt}) = 3 \times 10^5. \]

To illustrate the importance of \( s(\lambda_{opt}) \) on the convergence of \( \lambda_k \), we present in Table 3 the values of \( s(\lambda_{opt}) \) corresponding to different \( \zeta \). In Figure 4, we plot the curves of \( \lambda_{opt} - \lambda_k \) with different \( s(\lambda_{opt}) \). It is obvious to see from the figure that the larger \( s(\lambda_{opt}) \), the slower \( \lambda_k \) converges. Recall from Theorem 4.1 and Theorem 4.2 that \( s(\lambda_{opt}) \) and \( \kappa \) are two key factors for the convergence of \( \lambda_k \). As the condition numbers \( \kappa \) are the same in all the cases, the differences are mainly from those of \( s(\lambda_{opt}) \). This shows the effectiveness of our theoretical results, refer to Remark 4.2.

**Example 3.** In this example, we consider the cubic regularization problem (70). To show the effectiveness of our theoretical results, we compare our upper bound (85) with (79) established by X. Jia et al. Consider the matrix used in Example 1,
where

\[ A = \text{diag}(t^{[a,b]}_{jn}), \quad j = 1, 2, \ldots, n, \text{ with } a = -1, \ b = 1, \text{ and } n = 10000. \]

### Table 4. Example 3: \( \sigma, \mu_{opt}, \kappa_{\sigma,k} \) and residual norms in the cubic regularization problem.

| \( \sigma \) | \( \mu_{opt} \) | \( \kappa_{\sigma,k} \) | \( \| (A + \mu_{opt}I)x_{opt} + g \| \) |
|----------------|-----------------|-----------------|----------------------|
| 1              | 1.3405          | 6.8729          | 1.3439e-16           |
| 0.1            | 1.0214          | 94.5315         | 7.3994e-16           |
| 0.001          | 1.0010          | 2.0914e+03      | 6.1067e-15           |
| 0.0001         | 1.0000          | 4.5269e+04      | 7.9361e-14           |

It was pointed out that (70) can be equivalently rewritten as a large-scale eigenvalue problem [21]. In this example, we make use of the MATLAB built-in function `eigs.m` (with stopping tolerance \( tol = 10^{-12} \)) to compute \( \mu_{opt} \) and \( \mu_k \). As \( A_{opt} \) is a diagonal matrix in this example, the solution \( x_{opt}^{\sigma} \) is computed by using the dot division command \( \backslash \) in MATLAB.

The choices of \( \sigma \), the values of \( \mu_{opt}, \kappa_{\sigma,k} \) and residual norms are listed in Table 4. Figure 5 plots the curves of \( \mu^3_{opt} - \mu^3_k \) and the two upper bounds as \( k \) increases. Two remarks are in order. First, we see from Figure 5 that our result is much sharper than that of X. Jia et al., and the convergence rates of the two upper bounds are different in essence. More precisely, the convergence rate of our new bound is \((\sqrt{\kappa_{\sigma}}-1)^2\), while that of X. Jia et al. is \((\sqrt{\kappa_{\sigma}}-1)^2\); please see (85) and (79), respectively. Second, similar to Figure 3, it is seen from Figure 5 that the new upper bound is relatively large at the initial stage of the iteration, which coincides with the trend of the real values of \( \mu^3_{opt} - \mu^3_k \). This is because the value of \( \|g\|^2 \cdot e^T(3T_k + \mu_{opt}I)^{-3}e_1 \) can be small in the initial stage.

### 7. Conclusion

The GLTR method is a popular approach for solving large-scale TRS. In essence, this method is a projection method in which the original large-scale TRS is projected into a small-sized one. Recently, Z. Jia et al. considered the convergence of the GLTR method [20]. In this paper, we revisit this problem and establish some
Figure 5. Example 3: A comparison of the upper bounds on $\mu_{\text{opt}}^3 - \mu_k^3$.

Refined bounds on the residual norm $\| (A + \lambda_k I) x_k + g \|$, the distance $\sin \angle (x_{\text{opt}}, x_k)$ between the exact solution $x_{\text{opt}}$ and the approximate solution $x_k$, as well as the distance $\lambda_{\text{opt}} - \lambda_k$ between the Lagrange multiplier $\lambda_{\text{opt}}$ and its approximation $\lambda_k$. Moreover, we generalize these results to the convergence of Krylov subspace method for the cubic regularization problem, and improve some bounds due to X. Jia et al [18].

In this paper, we only consider the situation of easy case of the TRS (1). Further study includes the perturbation analysis on this problem, and proposing more efficient methods for judging and solving the TRS (1) in the (nearly) hard case. These are very interesting topics and are definitely a part of our future work.

References

[1] S. Adachi, S. Iwata, Y. Nakatsukasa, and A. Takeda, Solving the trust-region subproblem by a generalized eigenvalue problem, SIAM J. Optim., 27 (2017), pp. 269–291.

[2] T. Bianconcini, G. Liuzzi, B. Morini, and M. Sciandrone, On the use of iterative methods in cubic regularization for unconstrained optimization, Comput. Optim. Appl., 60 (2015), pp. 35–57.

[3] Y. Carmon and J. Duchi, First-order method for nonconvex quadratic minimization, SIAM Rev., 62 (2020), pp. 395–436.

[4] Y. Carmon and J. Duchi, Analysis of Krylov subspace solutions of regularized nonconvex quadratic problems, in Proceedings of the 32nd International Conference on Neural Information Processing Systems, 2018, pp. 10728–10738.

[5] C. Cartis, N. Gould, and P. Toint, Adaptive cubic regularisation methods for unconstrained optimization. Part I: Motivation, convergence and numerical results, Math. Programming, 127 (2011), pp. 245–295.
[6] C. Cartis, N. Gould, and M. Lange, "On monotonic estimates of the norm of the minimizers of regularized quadratic functions in Krylov spaces," BIT Numerical Mathematics, 60 (2020), pp. 583–589.

[7] A. Conn, N. Gould, and P. Toint, Trust-Region Methods, SIAM, Philadelphia, PA, 2000.

[8] J. Erway, P. Gill, and J. Griffin, "Iterative methods for finding a trust-region step," SIAM J. Optim., 20 (2009), pp. 1110–1131.

[9] W. Gander, G.H. Golub, and U. von Matt, "A constrained eigenvalue problem," Linear Algebra Appl., 114 (1989), pp. 815–839.

[10] G.H. Golub and U. von Matt, "Quadratically constrained least squares and quadratic problems," Numer. Math., 59 (1991), pp. 561–580.

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

[12] N. Gould, D. Robinson, and H. Thorne, "On solving trust-region and other regularized subproblems in optimization," Math. Program. Comput., 2 (2010), pp. 21–57.

[13] N. Gould and V. Simoncini, "Error estimates for iterative algorithms for minimizing regularized quadratic subproblems," Optimization Methods and Soft., 35 (2020), pp. 304–328.

[14] A. Greenbaum, Iterative Methods for Solving Linear Systems, Frontiers in Appl. Math. 17, SIAM, Philadelphia, 1997.

[15] W. Hager, "Minimizing a quadratic over a sphere," SIAM J. Optim., 12 (2001) pp. 188–208.

[16] W. Hager and Y. Krylyuk, "Graph partitioning and continuous quadratic programming," SIAM J. Discrete Math., 12 (1999), pp. 500–523.

[17] R. Horn and C. Johnson, Matrix Analysis, 2nd edition, Cambridge University Press, Cambridge, UK, 2013.

[18] X. Jia, X. Liang, C. Shen and L. Zhang, "Solving the cubic regularization model by a nested restarting Lanczos method," SIAM J. Matrix Anal. Appl., 43 (2022), pp. 812–839.

[19] Z. Jia, "The convergence of generalized Lanczos methods for large unsymmetric eigenproblems," SIAM J. Matrix Anal. Appl., 16 (1995), pp. 843–862.

[20] Z. Jia and F. Wang, "The convergence of the generalized Lanczos trust-region method for the trust-region subproblem," SIAM J. Optim., 31 (2021), pp. 887–914.

[21] F. Lieder, "Solving large-scale cubic regularization by a generalized eigenvalue problem," SIAM J. Optim., 30 (2020), pp. 3345–3358.

[22] L. Lukšan, C. Matonoha and J. Vlček, "On Lagrange multipliers of trust-region subproblems," BIT, 48 (2008), pp. 763–768.

[23] J. Moré and D. Sorensen, "Computing a trust region step," SIAM J. Sci. Statist. Comput., 4 (1983) pp. 553–572.

[24] Y. Nesterov and B. Polyak, "Cubic regularization of Newton method and its global performance," Math. Program. 108 (2006), pp. 177–205.

[25] J. Nocedal and S. Wright, Numerical Optimization, 2nd edition., Springer, New York, 2006.

[26] F. Rendl and H. Wolkowicz, "A semidefinite framework for trust region subproblems with applications to large scale minimization," Math. Program., 77 (1997), pp. 273–299.

[27] M. Rojas, S. Santos, and D. Sorensen, "A new matrix-free algorithm for the large-scale trust-region subproblem," SIAM J. Optim., 11 (2000), pp. 611–646.

[28] M. Rojas, S. Santos, and D. Sorensen, "Algorithm 873: LSTRS: MATLAB software for large-scale trust-region subproblems and regularization," ACM Trans. Math. Software, 34 (2008), pp. 1–28.

[29] M. Rojas and D. Sorensen, "A trust-region approach to the regularization of large-scale discrete forms of ill-posed problems," SIAM J. Sci. Comput., 23 (2002), pp. 1843–1861.

[30] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd edition., SIAM, Philadelphia, PA, 2003.

[31] Y. Saad, Numerical Method for Large Eigenvalue Problems, 2nd edition., SIAM, Philadelphia, PA, 2011.

[32] D. Sorensen, "Newton’s method with a model trust region modification," SIAM J. Numer. Anal., 19 (1982), pp. 409–426.

[33] D. Sorensen, "Minimization of a large-scale quadratic function subject to a spherical constraint," SIAM J. Optim., 7 (1997), pp. 141–161.

[34] T. Steihaug, "The conjugate gradient method and trust regions in large scale optimization," SIAM J. Numer. Anal., 20 (1983), pp. 626–637.
[35] G.W. Stewart, *Matrix Algorithms II, Eigensystems*, SIAM, Philadelphia, 2001.

[36] P. Tao and L. An, *A DC optimization algorithm for solving the trust-region subproblem*, SIAM J. Optim., 8 (1998), pp. 476–505.

[37] P. Toint, *Towards an efficient sparsity exploiting Newton method for minimization*, in Sparse Matrices and Their Uses, I. Duff, ed., Academic Press, London, 1981, pp. 57-88.

[38] L. Zhang, C. Shen, and R. Li, *On the generalized Lanczos trust-region method*, SIAM J. Optim., 27 (2017), pp. 2110–2142.

[39] L. Zhang and C. Shen, *A nested Lanczos method for the trust-region subproblem*, SIAM J. Sci. Comput., 40 (2018), pp. A2005–A2032.

China University of Mining and Technology, Xuzhou, 221116, P.R. China.

*Email address: bofeng@cumt.edu.cn*

China University of Mining and Technology, Xuzhou, 221116, P.R. China.

*Email address: gangwu@cumt.edu.cn*