arXiv:2304.11886v1 [math.NA] 24 Apr 2023

A BLOCK LANCZOS METHOD FOR LARGE-SCALE QUADRATIC MINIMIZATION PROBLEMS WITH ORTHOGONALITY CONSTRAINTS

BO FENG* AND GANG WU†

Abstract. Quadratic minimization problems with orthogonality constraints (QMPO) play an important role in many applications of science and engineering. However, some existing methods may suffer from low accuracy or heavy workload for large-scale QMPO. Krylov subspace methods are popular for large-scale optimization problems. In this work, we propose a block Lanczos method for solving the large-scale QMPO. In the proposed method, the original problem is projected into a small-sized one, and the Riemannian Trust-Region method is employed to solve the reduced QMPO. Convergence results on the optimal solution, the optimal objective function value, the multiplier and the KKT error are established. Moreover, we give the convergence speed of optimal solution, and show that if the block Lanczos process terminates, then an exact KKT solution is derived. Numerical experiments illustrate the numerical behavior of the proposed algorithm, and demonstrate that it is more powerful than many state-of-the-art algorithms for large-scale quadratic minimization problems with orthogonality constraints.

Key words. Quadratic minimization problems with orthogonality constraints (QMPO), Block Lanczos method, Block Krylov subspace, Large-scale optimization problem, Riemannian Trust-Region method.

AMS subject classifications. 65F15, 65F10, 90C20, 90C26.

1. Introduction.

We are interested in solving the following large-scale quadratic minimization problems with orthogonality constraints (QMPO)

$$\min_{U\in\mathbb{O}^{n\times\ell}} \left\{ f(U) := \mathrm{tr}(U^THU) + 2\ \mathrm{tr}(U^TG) \right\}, \quad n > \ell > 1, \tag{1.1}$$

where $H = H^T \in \mathbb{R}^{n\times n}$ is symmetric, $G \in \mathbb{R}^{n\times\ell}$, $\mathbb{O}^{n\times\ell} := \{U \in \mathbb{R}^{n\times\ell} \mid U^TU = I\}$. This problem of (1.1) arises from many practical problems such as orthogonal least squares regression (OLSR) [39], large graph clustering [37], multidimensional similarity structure analysis [5, Chapter 19], the Maxbet problem from canonical correlation analysis [15, 28], multi-view subspace clustering [47], and so on.

Indeed, the famous *unbalanced Procrustes problem* [7, 9, 10, 17, 35, 40, 41, 42]

$$\min_{U\in\mathbb{O}^{n\times\ell}} \|AU - B\|_{\mathrm{F}}, \tag{1.2}$$

is a special case of QMPO, where $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times\ell}$. By the first-order optimality conditions for unbalanced Procrustes problem (1.2) (cf. [7, Theorem 3.8] and [10, Theorem 3.1]), we have following first-order optimality necessary conditions on QMPO (1.1). There are some other types of optimality conditions on (1.1), for more details, refer to [7, 10, 41, 42].

*School of Mathematics, China University of Mining and Technology, 221116, Jiangsu, P.R. China. E-mail: bofeng@cumt.edu.cn.

†Corresponding author. School of Mathematics, China University of Mining and Technology, Xuzhou, 221116, Jiangsu, P.R. China. E-mail: gangwu@cumt.edu.cn. This author is supported by the National Natural Science Foundation of China under grant 12271518, the Key Research and Development Project of Xuzhou Natural Science Foundation under grant KC22288, and the Open Project of Key Laboratory of Data Science and Intelligence Education of the Ministry of Education under grant DSIE202203.

THEOREM 1.1. *[7, Theorem 3.8] and [10, Theorem 3.1] If $U \in \mathbb{O}^{n\times\ell}$ is a local minimizer of (1.1), then there is a symmetric matrix $\Lambda \in \mathbb{R}^{\ell\times\ell}$ such that*

$$HU + U\Lambda = -G. \tag{1.3}$$

If U is a global minimizer of (1.1), then

$$U^T HU + \Lambda = -U^T G = -G^T U \succcurlyeq \mathbf{O}. \tag{1.4}$$

Moreover, if $n = \ell$, then (1.1) reduces to the *balanced Procrustes problem* [16, 17, 25]

$$\min_{U\in\mathbb{O}^{\ell\times\ell}} \mathrm{tr}(U^T G). \tag{1.5}$$

In this case, we have a closed-form solution of (1.5) by using the SVD decomposition of G [16, 17]. And if $\ell = 1$, then (1.1) reduces to the classical *trust-region subproblem* [3, 6, 8, 12, 18, 20, 26, 29, 31, 45, 46].

Unfortunately, there is no closed-form solution for (1.1) generally. Some necessary or sufficient conditions for local and/or global minimizer of (1.1) were established in [7, 10, 41, 42]. Many iterative methods have been developed for the more general *optimization problems with orthogonal constraints*, which can be applied to solve (1.1) directly. For instance, Absil *et al.* proposed a Riemannian Trust-Region (RTR) algorithm [1, 2] for optimizing a smooth function on a Riemannian manifold. In [27], Jiang and Dai proposed a framework for a constraint preserving update scheme for optimization on Stiefel manifold. In [38], Wen and Yin applied the Cayley transform to preserve the orthogonal constraints and develop curvilinear search algorithms with lower flops compared to those based on projections and geodesics. In [24], structured quasi-Newton methods were studied for optimization problems with orthogonality constraints. In [14], Gao *et al.* proposed a proximal linearized augmented Lagrangian algorithm for solving optimization problems with orthogonality constraints. A first-order framework was proposed in [13] for optimization problems with orthogonal constraints.

The generalized power iteration (GPI) is one of the most popular methods for (1.1) [30]. However, this method often suffers from the difficulty of slow convergence, and more detailed analysis is desired for the convergence of GPI. Recently, a novel eigenvalue-based approach was proposed in [42] to solve the unbalanced Procrustes problem (1.2). This method also applies to the QMPO problem. It was proven that (1.1) can be equivalently transformed into an eigenvalue minimization whose solution can be computed by solving a related eigenvector-dependent nonlinear eigenvalue problem. However, one has to solve an n-by-n (possibly dense) symmetric eigenproblem in each iteration of this algorithm, and the algorithm may converge very slowly if there is no subspace speeding up.

To the best of our knowledge, there are few specialized methods for solving *large-scale* QMPO (1.1). Some existing methods may suffer from low accuracy or heavy workload for large-scale QMPO. Krylov subspace method is a powerful tool for solving large-scale optimization problems [11, 18, 26, 45, 43, 44, 46]. As far as we know, it seems that there is no (block) Krylov subspace method for the large-scale QMPO (1.1) till now. In this paper, we propose a block Krylov subspace method to solve (1.1), in which the large-scale QMPO (1.1) is reduced into a small-sized one by using projection techniques. Furthermore, we establish the convergence results on the optimal solution, the optimal objective function value, the multiplier, as well as the KKT error. We give the convergence speed of optimal solution, and show that if the block Lanczos process terminates, then an exact KKT solution is derived, which satisfies the first order optimality in Theorem 1.1 and also the necessary condition (1.4) for a global minimizer. Numerical experiments on both synthetic and real-world data sets demonstrate that the proposed algorithm is superior to many state-of-the-art approaches for solving the large-scale QMPO (1.1).

This paper is organized as follows. In Section 2, we propose a block Lanczos method for solving the large-scale QMPO. The convergence of the proposed method is established in Section 3. Numerical experiments are performed in Section 4 to show the numerical behavior of the new algorithm. Some concluding remarks are given in Section 5. Throughout this paper, we denote by $(\cdot)^T$ the transpose of a matrix or vector, by $\mathcal{R}(E)$ the range space of a matrix E, and by $E \otimes F$ the Kronecker product of E and F. In this paper, $E \succcurlyeq \mathbf{O}$ ($E \succ \mathbf{O}$) implies that E is symmetric semi-positive definite (positive definite). Let $F = [\boldsymbol{f}_1, \dots, \boldsymbol{f}_p] \in \mathbb{R}^{p\times q}$, then

$$\mathrm{vec}(F) = (\boldsymbol{f}_1^T, \dots, \boldsymbol{f}_p^T)^T \in \mathbb{R}^{pq}.$$

Let $\mathbf{0}$, $\mathbf{O}$ and I be the zero vector, zero matrix and identity matrix, respectively, whose orders are clear from the context.

2. A block Lanczos method for solving the large-scale QMPO. In this section, we propose a block Lanczos method to solve (1.1). Let $G = V_1 K$ be the economized QR decomposition of G, where $V_1 \in \mathbb{O}^{n\times\ell}$. As H is a symmetric matrix, we use the k-step block Lanczos process [16, 32, 33] to generate an orthonormal basis $\mathbf{V}_k$ for the block Krylov subspace

$$\mathcal{K}_k := \mathcal{K}_k(H, V_1) = span\{V_1, HV_1, \dots, H^{k-1}V_1\}.$$

Moreover, we have the following relation for this process [16, 32, 33]

$$H\mathbf{V}_k = \mathbf{V}_k T_k + V_{k+1} N_k (E_\ell^{(k\ell)})^T, \tag{2.1}$$

where $\mathbf{V}_k := [V_1,\ V_2, \dots,\ V_k] \in \mathbb{R}^{n\times k\ell}$, $\mathbf{V}_k^T\mathbf{V}_k = I_{k\ell}$, $\mathbf{V}_k^T V_{k+1} = \mathbf{O}$, $V_{k+1}^T V_{k+1} = I_\ell$, $N_k \in \mathbb{R}^{\ell\times\ell}$ is upper triangular, and $E_\ell^{k\ell}$ denotes the last ℓ columns of the identity matrix $I_{k\ell}$. Here

$$T_k = \mathbf{V}_k^T H \mathbf{V}_k = \begin{pmatrix} M_1 & N_1^T & & & \\ N_1 & M_2 & N_2^T & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & M_{k-1} & N_{k-1}^T \\ & & & N_{k-1} & M_k \end{pmatrix} \in \mathbb{R}^{k\ell\times k\ell}, \tag{2.2}$$

is block tridiagonal, with $M_j \in \mathbb{R}^{\ell\times\ell}$, and $N_j \in \mathbb{R}^{\ell\times\ell}$ being upper triangular, $j = 1, 2, \dots, k$.

In the proposed method, (1.1) reduces to the following small-sized *constrained* problem:

$$\min_{\substack{U\in\mathbb{O}^{n\times\ell} \\ \mathcal{R}(U)\subseteq\mathcal{K}_k}} \left\{ f(U) = \mathrm{tr}(U^T H U) + 2\ \mathrm{tr}(U^T G) \right\}. \tag{2.3}$$

Indeed, (2.3) can be equivalently rewritten as the following *reduced* QMPO:

$$\min_{P\in\mathbb{O}^{k\ell\times\ell}} \left\{ \widetilde{f}(P) = \mathrm{tr}(P^T T_k P) + 2\ \mathrm{tr}(P^T G_k) \right\}, \tag{2.4}$$

where

$$G_k = \mathbf{V}_k^T G = \mathbf{V}_k^T V_1 K = \begin{pmatrix} K \\ \mathbf{O} \end{pmatrix} \in \mathbb{R}^{k\ell \times \ell}.$$

Let P_k be a solution to (2.4), then we use

$$U_k = \mathbf{V}_k P_k = \arg \min_{\substack{U \in \mathbb{O}^{n\times \ell} \\ \mathcal{R}(U) \subseteq \mathcal{K}_k}} f(U) \tag{2.5}$$

as an approximation to the optimal solution U_*, and

$$f(U_k) = f(\mathbf{V}_k P_k) = \widetilde{f}(P_k) \tag{2.6}$$

is an approximation to the optimal value $f(U_*)$. By Theorem 1.1, there is a symmetric matrix $\Lambda_k \in \mathbb{R}^{k\times k}$ such that

$$T_k P_k + P_k \Lambda_k + G_k = 0. \tag{2.7}$$

Consequently, we reduce the large-scale QMPO (1.1) to a $k\ell$-by-$k\ell$ small-sized one. In practice, one can exploit the Riemannian Trust-Region (RTR) method [1, 2] to solve (2.4). The proposed algorithm is given in Algorithm 1. One refers to Section 4 for more details on practical implementations.

Algorithm 1 A block Lanczos method for large-scale QMPO

Input: $H \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{n\times \ell}$, and $k_{\max}$.
Output: U_k.
Set $V_0 = \mathbf{O} \in \mathbb{R}^{n\times \ell}$, $N_0 = \mathbf{O} \in \mathbb{R}^{\ell\times \ell}$ and $k = 0$;
Compute the economized QR decomposition: $G = V_1 K$, where $V_1 \in \mathbb{R}^{n\times \ell}$;
Let $M_1 = V_1^T H V_1$;
whlie $k \le k_{\max}$
  $k = k + 1$;
  Let $L_k = HV_k - V_k M_k - V_{k-1} N_{k-1}^T$;
  Compute the economized QR decomposition: $L_k = V_{k+1} N_k$;
  Let $M_{k+1} = V_{k+1}^T H V_{k+1}$;
  Solve the reduced QMPO

$$P_k = \arg \min_{P \in \mathbb{O}^{k\ell\times \ell}} \left\{ \mathrm{tr}(P^T T_k P) + 2\ \mathrm{tr}(P^T G_k) \right\},$$

  where $T_k \in \mathbb{R}^{k\ell\times k\ell}$ is defined in (2.2) and $G_k = \begin{pmatrix} K \\ \mathbf{O} \end{pmatrix} \in \mathbb{R}^{k\ell\times \ell}$;
  if the convergence criterion is satisfied % `Refer to` (4.1)
    $U_k = \mathbf{V}_k P_k$;
  end if
end whlie

---

3. Convergence analysis. In this section, we show the convergence of Algorithm 1. We first need the following three lemmas. The first lemma follows from the definition of the Kronecker product and [22, Theorem 4.4.5].

LEMMA 3.1. *Let $X \in \mathbb{R}^{s\times s}$ and $Y \in \mathbb{R}^{m\times m}$, then $(I_m \otimes X)+(Y \otimes I_s)$ is nonsingular if and only if $X+\lambda_Y I$ is nonsingular, where λ_Y is an eigenvalue of Y. Moreover, if $X = X^T$ and $Y = Y^T$, then*

$$(I_m \otimes X) + (Y \otimes I_s) \text{ is symmetric,} \tag{3.1a}$$

$$\lambda_{\max}\big((I_m \otimes X) + (Y \otimes I_s)\big) = \lambda_{\max}(X) + \lambda_{\max}(Y), \tag{3.1b}$$

$$\lambda_{\min}\big((I_m \otimes X) + (Y \otimes I_s)\big) = \lambda_{\min}(X) + \lambda_{\min}(Y). \tag{3.1c}$$

The second lemma is from [22, Section 4.2, Problem 25].

LEMMA 3.2. *[22] Let $C \in \mathbb{R}^{t\times s}$, $E \in \mathbb{R}^{p\times q}$, $X \in \mathbb{R}^{q\times s}$ and $Y \in \mathbb{R}^{p\times t}$. Then*

$$\mathrm{tr}(C^T Y^T E X) = \underline{\boldsymbol{y}}^T (C \otimes E)\underline{\boldsymbol{x}} \tag{3.2}$$

where $\underline{\boldsymbol{y}} = \mathrm{vec}(Y)$ and $\underline{\boldsymbol{x}} = \mathrm{vec}(X)$.

The third lemma is the polar decomposition of a full column rank matrix.

LEMMA 3.3. *[23, Theorem 8.1] Let $Y \in \mathbb{R}^{p\times s}$ $(p \geq s)$ with $rank(Y) = s$. There exists a unique matrix $Q \in \mathbb{R}^{p\times s}$ with orthonormal columns and a unique symmetric positive definite matrix S such that $Y = QS$. The matrix S is given by $S = (Y^TY)^{\frac{1}{2}}$.*

We are ready to consider the convergence of the proposed method. Let U_* be a global minimizer of (1.1), then U_* is also a local solution. It follows from Theorem 1.1 that, there is a symmetric matrix $\Lambda_* \in \mathbb{R}^{\ell\times\ell}$, such that (1.3) holds. Let the eigendecompositions of H and Λ_* be

$$H = WDW^T = W\ diag(\mu_1, \mu_2, \ldots, \mu_n)\ W^T, \quad \text{with} \quad \mu_1 \geq \mu_2 \geq \cdots \geq \mu_n, \tag{3.3a}$$

$$\Lambda_* = Z\Gamma Z^T = Z\ diag(\gamma_1, \gamma_2, \ldots, \gamma_\ell)\ Z^T, \quad \text{with} \quad \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_\ell, \tag{3.3b}$$

where $W \in \mathbb{R}^{n\times n}$ and $Z = [\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_\ell] \in \mathbb{R}^{\ell\times\ell}$ are orthonormal matrices.

In this paper, we make the following assumption

$$\textit{Assumption}:\ \mathcal{H}_* := (I_\ell \otimes H) + (\Lambda_* \otimes I_n) \ \text{ is nonsingular.} \tag{3.4}$$

Hence, it follows from Lemma 3.1 that $H + \gamma_i I$ are nonsingular, $i = 1, 2, \ldots, \ell$. Consider the two index sets

$$\mathcal{I} = \{i \mid H + \gamma_i I \succ \mathbf{O},\ \text{i.e.,}\ \mu_n + \gamma_i > 0,\ \text{where}\ 1 \leq i \leq \ell\},$$
$$\mathcal{J} = \{i \mid H + \gamma_i I \ \text{ is nonsingular and}\ \mu_n + \gamma_i < 0,\ \text{where}\ 1 \leq i \leq \ell\}.$$

Thus, we have that $\mathcal{I} \cup \mathcal{J} = \{1, 2, \ldots, \ell\}$ and $\mathcal{I} \cap \mathcal{J} = \varnothing$. It was shown in [42, Theorem 2.4] that $\mu_1 + \gamma_n \geq 0$. In other words, $H + \gamma_i I$ will never be a negative definite matrix, $i = 1, 2, \ldots, \ell$. Therefore, if $\mathcal{J} \neq \varnothing$, there is an integer $1 \leq s_i \leq n$, such that

$$\mu_n + \gamma_i \leq \cdots \leq \mu_{s_i} + \gamma_i < 0 < \mu_{s_i+1} + \gamma_i \leq \cdots \leq \mu_1 + \gamma_i, \quad \text{for} \quad i \in \mathcal{J}.$$

Consider

$$\phi_i := \frac{a_i b_i}{|(\mu_{s_i} + \gamma_i)(\mu_{s_i+1} + \gamma_i)|} \geq 1 \quad \text{for} \quad i \in \mathcal{J},$$

where

$$a_i = \max\big\{ -(\mu_n + \gamma_i), \mu_1 - \mu_{s_i+1} - (\mu_{s_i} + \gamma_i)\big\},$$
$$b_i = \max\big\{\mu_1 + \gamma_i, \mu_{s_1+1} - \mu_n + \mu_{s_i} + \gamma_i\big\}.$$

The definitions of a_i and b_i are due to the embedding of $[\mu_n+\gamma_i,\mu_{s_i}+\gamma_i]\cup[\mu_{s_i+1}+\gamma_i,\mu_1+\gamma_i]\subseteq[-a_i,\mu_{s_i}+\gamma_i]\cup[\mu_{s_i+1}+\gamma_i,b_i]$ into intervals of equal lengths [21, section 3.1].

• First, we consider the distance between the optimal solution U_* and the search subspace $\mathcal{K}_k$. Indeed, a necessary condition for the convergence of the proposed method is that the distance tends to zero.

THEOREM 3.4. *Let*

$$\epsilon_k = \min_{X\in\mathbb{R}^{k\ell\times\ell}} \|U_* - \mathbf{V}_k X\|_{\mathrm{F}} = \|(I-\mathbf{V}_k\mathbf{V}_k^T)U_*\|_{\mathrm{F}}.$$

Under the previous notation and assumption, we have

$$\epsilon_k \le 2\cdot\sqrt{\sum_{i\in\mathcal{I}}\left(\frac{\sqrt{\kappa_i}-1}{\sqrt{\kappa_i}+1}\right)^{2(k+1)}+\sum_{i\in\mathcal{J}}\left(\frac{\sqrt{\phi_i}-1}{\sqrt{\phi_i}+1}\right)^{k-1}}, \tag{3.5}$$

where κ_i is the 2-condition number of $H+\gamma_i I$ for $i\in\mathcal{I}$.

Proof. From $HU_*+U_*\Lambda_* = -G$ and $\Lambda_* = Z\Gamma Z^T$, we have

$$(H+\gamma_i I)U_*\boldsymbol{z}_i = -G\boldsymbol{z}_i \quad\text{for}\quad i=1,2,\ldots,\ell.$$

Denote by $\mathcal{K}_k^{(i)} := \mathcal{K}_k(H+\gamma_i I, G\boldsymbol{z}_i)$ and let $\mathcal{P}_t$ be the set of polynomials with degree no higher than t. It holds that

$$\begin{aligned}
\min_{\boldsymbol{y}\in\mathcal{K}_k^{(i)}}\|U_*\boldsymbol{z}_i-\boldsymbol{y}\|_2 &= \min_{p\in\mathcal{P}_k}\|U_*\boldsymbol{z}_i-p(H+\gamma_i I)G\boldsymbol{z}_i\|_2\\
&= \min_{p\in\mathcal{P}_k}\|[I+p(H+\gamma_i I)(H+\gamma_i I)]U_*\boldsymbol{z}_i\|_2\\
&= \min_{\substack{h\in\mathcal{P}_{k+1}\\ h(0)=1}}\|h(H+\gamma_i I)U_*\boldsymbol{z}_i\|_2 \le \min_{\substack{h\in\mathcal{P}_{k+1}\\ h(0)=1}}\|h(H+\gamma_i I)\|_2\\
&\overset{(3.3a)}{=} \min_{\substack{h\in\mathcal{P}_{k+1}\\ h(0)=1}}\|h(D+\gamma_i I)\|_2.
\end{aligned}$$

From the Assumption (3.4), we have that $H+\gamma_i I$ are nonsingular, $i=1,2,\ldots,\ell$. Then it follows from [19, Section 3.1] that, on one hand, if $i\in\mathcal{I}$,

$$\begin{aligned}
\min_{\boldsymbol{y}\in\mathcal{K}_k^{(i)}}\|U_*\boldsymbol{z}_i-\boldsymbol{y}\|_2 &\le \min_{\substack{h\in\mathcal{P}_{k+1}\\ h(0)=1}}\|h(D+\gamma_i I)\|_2\\
&\le \min_{\substack{h\in\mathcal{P}_{k+1}\\ h(0)=1}}\max_{t\in[\mu_n+\gamma_i,\mu_1+\gamma_i]}|h(t)| \le 2\left(\frac{\sqrt{\kappa_i}-1}{\sqrt{\kappa_i}+1}\right)^{k+1}.
\end{aligned} \tag{3.6}$$

On the other hand, if $i\in\mathcal{J}$,

$$\begin{aligned}
\min_{\boldsymbol{y}\in\mathcal{K}_k^{(i)}}\|U_*\boldsymbol{z}_i-\boldsymbol{y}\|_2 &\le \min_{\substack{h\in\mathcal{P}_{k+1}\\ h(0)=1}}\|h(D+\gamma_i I)\| \le \min_{\substack{h\in\mathcal{P}_{k+1}\\ h(0)=1}}\max_{t\in\mathcal{L}}|h(t)|\\
&\le 2\left(\frac{\sqrt{|a_ib_i|}-\sqrt{|(\mu_{s_i}+\gamma_i)(\mu_{s_i+1}+\gamma_i)|}}{\sqrt{|a_ib_i|}+\sqrt{|(\mu_{s_i}+\gamma_i)(\mu_{s_i+1}+\gamma_i)|}}\right)^{\lfloor\frac{k+1}{2}\rfloor} = 2\left(\frac{\sqrt{\phi_i}-1}{\sqrt{\phi_i}+1}\right)^{\lfloor\frac{k+1}{2}\rfloor},
\end{aligned} \tag{3.7}$$

where $\mathcal{L} = [-a_i, \mu_{s_i} + \gamma_i] \cup [\mu_{s_i+1} + \gamma_i, b_i]$ and $\lfloor \cdot \rfloor$ stands for the integer part of a number.

From $G\boldsymbol{z}_i \in \mathcal{R}(V_1)$, we obtain

$$\mathcal{K}_k^{(i)} = \mathcal{K}_k(H, G\boldsymbol{z}_i) \subseteq \mathcal{K}_k(H, V_1) = \mathcal{K}_k \quad \text{for} \quad i = 1, 2, \ldots, \ell.$$

So we have

$$\begin{aligned}
\epsilon_k^2 &= \left\|(I - \mathbf{V}_k \mathbf{V}_k^T) U_*\right\|_{\mathrm{F}}^2 \overset{(3.3b)}{=} \left\|(I - \mathbf{V}_k \mathbf{V}_k^T) U_* Z\right\|_{\mathrm{F}}^2 \\
&= \sum_{i=1}^{\ell} \|(I - \mathbf{V}_k \mathbf{V}_k^T) U_* \boldsymbol{z}_i\|_2^2 = \sum_{i=1}^{\ell} \min_{\boldsymbol{y} \in \mathcal{K}_k} \|U_* \boldsymbol{z}_i - \boldsymbol{y}\|_2^2 \\
&= \sum_{i \in \mathcal{I}} \min_{\boldsymbol{y} \in \mathcal{K}_k} \|U_* \boldsymbol{z}_i - \boldsymbol{y}\|_2^2 + \sum_{i \in \mathcal{J}} \min_{\boldsymbol{y} \in \mathcal{K}_k} \|U_* \boldsymbol{z}_i - \boldsymbol{y}\|_2^2 \\
&\le \sum_{i \in \mathcal{I}} \min_{\boldsymbol{y} \in \mathcal{K}_k^{(i)}} \|U_* \boldsymbol{z}_i - \boldsymbol{y}\|_2^2 + \sum_{i \in \mathcal{J}} \min_{\boldsymbol{y} \in \mathcal{K}_k^{(i)}} \|U_* \boldsymbol{z}_i - \boldsymbol{y}\|_2^2,
\end{aligned}$$

which, together with (3.6) and (3.7), yields (3.5). ∎

REMARK 3.1. *Theorem 3.4 shows that the rate at which ϵ_k converges to 0 strictly relies on the distribution of the spectrum of $H + \gamma_i I, i = 1, 2, \ldots, \ell$. In particular, the convergence rate of ϵ_k is comparable to that of conjugate gradient method provided that $(I_\ell \otimes H) + (\Lambda_* \otimes I_n) \succ \mathbf{O}$.*

• Second, we show the convergence of $f(U_k)$. To this aim, we consider the upper bound of $f(U_k) - f(U_*)$.

THEOREM 3.5. *Suppose that $\|(I - \mathbf{V}_k \mathbf{V}_k^T) U_*\|_2 < 1$. Then*

$$0 \le f(U_k) - f(U_*) \le 2(\mu_1 + \gamma_1) \cdot \epsilon_k^2. \tag{3.8}$$

Proof. For any $U \in \mathbb{O}^{n \times \ell}$, we show that

$$0 \le f(U) - f(U_*) = \mathrm{tr}[(U_* - U)^T H (U_* - U)] + \mathrm{tr}[(U_* - U)\Lambda_*(U_* - U)^T]. \tag{3.9}$$

Indeed,

$$\begin{aligned}
0 \le f(U) - f(U_*) &= \mathrm{tr}(U^T H U) - \mathrm{tr}(U_*^T H U_*) + 2\ \mathrm{tr}[(U - U_*)^T G] \\
&\overset{(1.3)}{=} \mathrm{tr}(U^T H U) - \mathrm{tr}(U_*^T H U_*) + \mathrm{tr}[(U - U_*)^T G] - \mathrm{tr}[(U - U_*)^T (H U_* + U_* \Lambda_*)] \\
&= \mathrm{tr}(U^T H U) + \mathrm{tr}[(U - U_*)^T G] - \mathrm{tr}(U^T H U_*) + \mathrm{tr}[(U_* - U)^T U_* \Lambda_*] \\
&= \mathrm{tr}[(U - U_*)^T (H U + G)] + \mathrm{tr}[(U_* - U)^T U_* \Lambda_*] \\
&= \mathrm{tr}[(U - U_*)^T H (U - U_*)] + \mathrm{tr}[(U - U_*)^T (H U_* + G)] + \mathrm{tr}[(U_* - U)^T U_* \Lambda_*] \\
&\overset{(1.3)}{=} \mathrm{tr}[(U - U_*)^T H (U - U_*)] + 2\ \mathrm{tr}[(U_* - U)^T U_* \Lambda_*] \\
&= \mathrm{tr}[(U - U_*)^T H (U - U_*)] + \mathrm{tr}[(U_* - U)^T (U_* - U) \Lambda_*] + \mathrm{tr}[(U_* - U)^T (U_* + U) \Lambda_*] \\
&= \mathrm{tr}[(U - U_*)^T H (U - U_*)] + \mathrm{tr}[(U_* - U) \Lambda_* (U_* - U)^T],
\end{aligned}$$

where the last inequality is from the facts that $U \in \mathbb{O}^{n \times \ell}$, $\Lambda_* = \Lambda_*^T$, and

$$\begin{aligned}
\mathrm{tr}[(U_* - U)^T (U_* + U)\Lambda] &= \mathrm{tr}[(U_*^T U - U^T U_*)\Lambda_*] \\
&= \mathrm{tr}(U_*^T U \Lambda_*) - \mathrm{tr}(U^T U_* \Lambda_*) \\
&= \mathrm{tr}(U \Lambda_* U_*^T) - \mathrm{tr}(U_* \Lambda_* U^T) \\
&= \mathrm{tr}(U \Lambda_* U_*^T) - \mathrm{tr}(U \Lambda_* U_*^T) = 0,
\end{aligned}$$

so we have (3.9).

We are ready to prove (3.8). It follows from [22, Theorem 3.3.16 (c)] that

$$\begin{aligned}\left|1-\sigma_\ell(\mathbf{V}_k\mathbf{V}_k^TU_*)\right| =&\left|\sigma_\ell(U_*)-\sigma_\ell(\mathbf{V}_k\mathbf{V}_k^TU_*)\right|\\ &\leq \left\|U_*-\mathbf{V}_k\mathbf{V}_k^TU_*\right\|_2=\|(I-\mathbf{V}_k\mathbf{V}_k^T)U_*\|_2<1.\end{aligned}$$

Hence, $\sigma_\ell(\mathbf{V}_k\mathbf{V}_k^TU_*)>0$, i.e., $\mathbf{V}_k\mathbf{V}_k^TU_*\in\mathbb{R}^{n\times\ell}$ has full column rank. By Lemma 3.3, there is a unique orthonormal matrix $\widetilde{U}_k\in\mathbb{C}^{n\times\ell}$ and a symmetric positive definite matrix M, such that $\mathbf{V}_k\mathbf{V}_k^TU_*=\widetilde{U}_kM$. Thus, $\mathcal{R}(\widetilde{U}_k)\subseteq\mathcal{R}(\mathbf{V}_k)=\mathcal{K}_k$. By (2.5),

$$\begin{aligned}0&\leq f(U_k)-f(U_*)\leq f(\widetilde{U}_k)-f(U_*)\\ &\overset{(3.9)}{=}\operatorname{tr}[(\widetilde{U}_k-U_*)^TH(\widetilde{U}_k-U_*)]+\operatorname{tr}[(U_*-\widetilde{U}_k)\Lambda_*(U_*-\widetilde{U}_k)^T]\\ &=\operatorname{tr}[(\widetilde{U}_k-U_*)^TH(\widetilde{U}_k-U_*)]+\operatorname{tr}[\Lambda_*(U_*-\widetilde{U}_k)^T(U_*-\widetilde{U}_k)]\\ &\overset{(3.2)}{=}\left[\operatorname{vec}(U_*-\widetilde{U}_k)\right]^T\cdot\mathcal{H}_*\cdot\operatorname{vec}\big(U_*-\widetilde{U}_k\big)\\ &\overset{(3.1b)}{\leq}(\mu_1+\gamma_1)\cdot\left\|\operatorname{vec}(U_*-\widetilde{U}_k)\right\|_2^2=(\mu_1+\gamma_1)\cdot\left\|U_*-\widetilde{U}_k\right\|_{\mathrm{F}}^2.\end{aligned}\tag{3.10}$$

Notice that

$$\begin{aligned}\left\|\widetilde{U}_k-U_*\right\|_{\mathrm{F}}^2&=\left\|\widetilde{U}_k-\mathbf{V}_k\mathbf{V}_k^TU_*-(I-\mathbf{V}_k\mathbf{V}_k^T)U_*\right\|_{\mathrm{F}}^2\\ &=\left\|\widetilde{U}_k-\mathbf{V}_k\mathbf{V}_k^TU_*\right\|_{\mathrm{F}}^2+\left\|(I-\mathbf{V}_k\mathbf{V}_k^T)U_*\right\|_{\mathrm{F}}^2\\ &=\left\|\widetilde{U}_k-\mathbf{V}_k\mathbf{V}_k^TU_*\right\|_{\mathrm{F}}^2+\epsilon_k^2.\end{aligned}\tag{3.11}$$

Next, we consider $\left\|\widetilde{U}_k-\mathbf{V}_k\mathbf{V}_k^TU_*\right\|_{\mathrm{F}}^2$. We have that

$$\begin{aligned}\left\|\widetilde{U}_k-\mathbf{V}_k\mathbf{V}_k^TU_*\right\|_{\mathrm{F}}^2=&\|\widetilde{U}_k\|_{\mathrm{F}}^2+\|\mathbf{V}_k\mathbf{V}_k^TU_*\|_{\mathrm{F}}^2-2\operatorname{tr}(\widetilde{U}_k^T\mathbf{V}_k\mathbf{V}_k^TU_*)\\ =&\|U_*\|_{\mathrm{F}}^2+\|\mathbf{V}_k\mathbf{V}_k^TU_*\|_{\mathrm{F}}^2-2\operatorname{tr}(M)\\ =&\|U_*\|_{\mathrm{F}}^2-\|\mathbf{V}_k\mathbf{V}_k^TU_*\|_{\mathrm{F}}^2+2\left(\|\mathbf{V}_k\mathbf{V}_k^TU_*\|_{\mathrm{F}}^2-\operatorname{tr}(M)\right)\\ =&\left\|(I-\mathbf{V}_k\mathbf{V}_k^T)U_*\right\|_{\mathrm{F}}^2+2\left[\operatorname{tr}(M^TM)-\operatorname{tr}(M)\right]\\ =&\epsilon_k^2+2\left[\operatorname{tr}(M^TM)-\operatorname{tr}(M)\right].\end{aligned}$$

From $M=M^T\succ\mathbf{O}$ and $\sigma_i(M)=\sigma_i(\mathbf{V}_k\mathbf{V}_k^TU_*)\leq 1,\ i=1,2,\ldots,\ell$, we get

$$\operatorname{tr}(M^TM)-\operatorname{tr}(M)=\sum_{i=1}^{\ell}\sigma_i^2(M)-\sum_{i=1}^{\ell}\sigma_i(M)=\sum_{i=1}^{\ell}\Big(\sigma_i^2(M)-\sigma_i(M)\Big)\leq 0,$$

and $\left\|\widetilde{U}_k-\mathbf{V}_k\mathbf{V}_k^TU_*\right\|_{\mathrm{F}}\leq\epsilon_k$. A combination of (3.10) and (3.11) yields (3.8). □

REMARK 3.2. *We note that*

$$\mu_1+\gamma_1\leq\|H\|_2+\|\Lambda_*\|_2\overset{(1.3)}{=}\|H\|_2+\|HU_*+G\|_2\leq 2\|H\|_2+\|G\|_2.$$

That is, $(\mu_1+\gamma_1)$ is uniformly bounded, and $\epsilon_k\to 0$ implies $f(U_k)\to f(U_*)$.

• Third, we show the convergence of U_k. To do this, we pay special attention to the distance between the global optimal solution U_* and the approximate solution U_k.

Notice that U_* may be non-unique. In this case, the convergence of U_k is difficult to define. We first establish a sufficient condition for the uniqueness of U_*.

Theorem 3.6. *Let U_* be a global optimal solution of (1.1), and define*

$$\delta(U_*) := \inf_{\substack{X\in U_*+\mathbb{O}^{n\times\ell}\\ X\neq \mathbf{O}}} \frac{\operatorname{tr}(X^T H X)+\operatorname{tr}(X\Lambda_* X^T)}{\|X\|_{\mathrm{F}}^2}, \tag{3.12}$$

where

$$U_*+\mathbb{O}^{n\times\ell} = \left\{X \mid X = U_*+Q \quad \textit{where} \quad Q\in\mathbb{O}^{n\times\ell}\right\}.$$

Then we have that

(i) $\delta(U_*)\geq 0$. Moreover, if $\delta(U_*)>0$, then U_* is the unique global optimal solution to (1.1).

(ii) If the infimum in (3.12) is attainable, then $\delta(U_*)>0$ if and only if U_* is a unique global optimal solution to (1.1).

(iii) We have $\delta(U_*)\geq\lambda_{\min}\big((I_\ell\otimes H)+(\Lambda_*\otimes I_n)\big)=\mu_n+\gamma_\ell$. Specifically, if $\ell=1$, then $\Lambda_*\in\mathbb{R}$ is a scalar, and $\delta(U_*)=\lambda_{\min}(H)+\Lambda_*=\mu_n+\Lambda_*\geq 0$.

Proof. (i) We prove it by contradiction. Suppose that $\delta(U_*)<0$, there is a matrix $\widetilde{X}\in(U_*+\mathbb{O}^{n\times\ell})\backslash\{\mathbf{O}\}$, such that

$$\frac{\operatorname{tr}(\widetilde{X}^T H\widetilde{X})+\operatorname{tr}(\widetilde{X}\Lambda_*\widetilde{X}^T)}{\|\widetilde{X}\|_{\mathrm{F}}^2}<0.$$

Hence, there is a matrix $L\in\mathbb{O}^{n\times\ell}$, such that $\widetilde{X}=U_*+L$. By (3.9),

$$0\leq\frac{f(-L)-f(U_*)}{\|U_*+L\|_{\mathrm{F}}^2}=\frac{\operatorname{tr}(\widetilde{X}^T H\widetilde{X})+\operatorname{tr}(\widetilde{X}\Lambda_*\widetilde{X}^T)}{\|\widetilde{X}\|_{\mathrm{F}}^2}<0,$$

which is a contradiction. As a result, we have $\delta(U_*)\geq 0$.

Moreover, if $\delta(U_*)>0$ while U_* is non-unique, then there is a matrix $U_{**}\in\mathbb{O}^{n\times\ell}$ such that $U_*\neq U_{**}$ and $f(U_*)=f(U_{**})$. As $U_*-U_{**}\in(U_*+\mathbb{O}^{n\times\ell})\backslash\{\mathbf{O}\}$, it follows that

$$\begin{aligned}0<\delta(U_*)&\leq\frac{\operatorname{tr}[(U_*-U_{**})^T H(U_*-U_{**})]+\operatorname{tr}[(U_*-U_{**})\Lambda_*(U_*-U_{**})^T]}{\|U_*-U_{**}\|_{\mathrm{F}}^2}\\ &\overset{(3.9)}{=}\frac{f(U_{**})-f(U_*)}{\|U_*-U_{**}\|_{\mathrm{F}}^2}=0,\end{aligned}$$

a contradiction. Thus, U_* is a unique global optimal solution to (1.1).

(ii) If the infimum in (3.12) is attainable, then there is a matrix $X_*\in(U_*+\mathbb{O}^{n\times\ell})\backslash\{\mathbf{O}\}$, such that

$$X_*=\arg\min_{\substack{X\in U_*+\mathbb{O}^{n\times\ell}\\ X\neq \boldsymbol{O}}}\frac{\operatorname{tr}(X^T H X)+\operatorname{tr}(X\Lambda_* X^T)}{\|X\|_{\mathrm{F}}^2}.$$

Thus, there is a matrix $\widehat{X}_*\in\mathbb{O}^{n\times\ell}$ such that $X_*=U_*-\widehat{X}_*$. So we obtain from (3.9) that

$$f(\widehat{X}_*)-f(U_*)=\operatorname{tr}(X_*^T H X_*)+\operatorname{tr}(X_*\Lambda_* X_*^T)=\delta(U_*)\cdot\|X_*\|_{\mathrm{F}}^2.$$

If U_* is the unique global optimal solution to (1.1), then $f(\widehat{X}_*) - f(U_*) > 0$. That is, $\delta(U_*) > 0$, and we have (ii) from (i).

(iii) We have that

$$\begin{aligned}\delta(U_*) &\geq \inf_{\substack{X\in\mathbb{R}^{n\times\ell}\\ X\neq \boldsymbol{O}}} \frac{\operatorname{tr}(X^T H X)+\operatorname{tr}(X\Lambda_* X^T)}{\|X\|_{\mathrm{F}}^2} = \inf_{\substack{X\in\mathbb{R}^{n\times\ell}\\ X\neq \boldsymbol{O}}} \frac{\operatorname{tr}(X^T H X)+\operatorname{tr}(\Lambda_* X^T X)}{\|X\|_{\mathrm{F}}^2} \\ &\overset{(3.2)}{=} \inf_{\substack{X\in\mathbb{R}^{n\times\ell}\\ X\neq \boldsymbol{O}}} \frac{(\operatorname{vec}(X))^T \cdot \Big((I_\ell \otimes H)+(\Lambda_* \otimes I_n)\Big)\cdot \operatorname{vec}(X)}{\|\operatorname{vec}(X)\|_2^2} \\ &\overset{(3.1a)}{=} \lambda_{\min}(\mathcal{H}_*) \overset{(3.1c)}{=} \mu_n + \gamma_\ell.\end{aligned}$$

In particular, if $\ell=1$, then (1.1) reduces to a trust-region subproblem. In this case, $\Lambda_*=\lambda_*\in\mathbb{R}$ is a scalar, $H_* := H+\lambda_* I \succcurlyeq \mathbf{O}$ [20, Lemma 2.1], $\mathbb{O}^{n\times\ell}= \mathcal{B} := \{\boldsymbol{x} \mid \|\boldsymbol{x}\|_2=1\}$, and

$$\delta(U_*)=\inf_{\boldsymbol{x}\in\mathcal{S}} \boldsymbol{x}^T H_* \boldsymbol{x}, \quad \text{with} \quad \mathcal{S} = \left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{x} = \frac{U_*-\boldsymbol{p}}{\|U_*-\boldsymbol{p}\|}, \quad \text{where} \;\; U_* \neq \boldsymbol{p} \in \mathcal{B} \right\} \subseteq \mathcal{B}.$$

Denote by $\widetilde{\mathcal{S}} = \left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{x} = \frac{\boldsymbol{p}-U_*}{\|\boldsymbol{p}-U_*\|}, \quad \text{where} \;\; U_* \neq \boldsymbol{p} \in \mathcal{B} \right\}$, we see that

$$\inf_{\boldsymbol{x}\in\mathcal{S}\cup\widetilde{\mathcal{S}}} \boldsymbol{x}^T H_* \boldsymbol{x} = \min\left\{ \inf_{\boldsymbol{x}\in\mathcal{S}} \boldsymbol{x}^T H_* \boldsymbol{x}, \;\; \inf_{\boldsymbol{x}\in\widetilde{\mathcal{S}}} \boldsymbol{x}^T H_* \boldsymbol{x} \right\},$$

and $\inf_{\boldsymbol{x}\in\mathcal{S}} \boldsymbol{x}^T H_* \boldsymbol{x} = \inf_{\boldsymbol{x}\in\widetilde{\mathcal{S}}} \boldsymbol{x}^T H_* \boldsymbol{x}$. Therefore, $\delta(U_*) = \inf_{\boldsymbol{x}\in\mathcal{S}\cup\widetilde{\mathcal{S}}} \boldsymbol{x}^T H_* \boldsymbol{x}$. As

$$\mathcal{S}\cup\widetilde{\mathcal{S}} = \left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{x} = \pm\frac{\boldsymbol{p}-U_*}{\|\boldsymbol{p}-U_*\|}, \quad \text{where} \;\; U_* \neq \boldsymbol{p} \in \mathcal{B} \right\}$$

is dense on $\mathcal{B}$ [31, p. 91], it follows that

$$\delta(U_*) = \inf_{\boldsymbol{x}\in\mathcal{B}} \boldsymbol{x}^T (H+\lambda_* I)\boldsymbol{x} = \min_{\boldsymbol{x}\in\mathcal{B}} \boldsymbol{x}^T (H+\lambda_* I)\boldsymbol{x} = \mu_n + \lambda_* \geq 0,$$

which completes the proof. ∎

We are ready to consider the convergence of U_k.

THEOREM 3.7. *Let U_* be the global optimal solution of* (1.1). *If $\delta(U_*) > 0$ and $\|(I - \mathbf{V}_k\mathbf{V}_k^T)U_*\|_2 < 1$, then we have*

$$\|U_k - U_*\|_{\mathrm{F}} \leq \sqrt{\frac{2(\mu_1+\gamma_1)}{\delta(U_*)}} \cdot \epsilon_k. \tag{3.13}$$

Specifically, if $\mathcal{H}_* \succ \mathbf{O}$, then

$$\|U_k - U_*\|_{\mathrm{F}} \leq \sqrt{2\varkappa_*} \cdot \epsilon_k, \tag{3.14}$$

where $\varkappa_*$ is the 2-condition number of $\mathcal{H}_*$.

Proof. We notice that $U_* - U_k \in (U_* + \mathbb{O}^{n\times\ell}) \cup (-U_* + \mathbb{O}^{n\times\ell})$. It follows from (3.9) and (3.12) that

$$\begin{aligned} f(U_k)-f(U_*) &= \frac{\operatorname{tr}[(U_*-U_k)^T H (U_*-U_k)] + \operatorname{tr}[(U_*-U_k)\Lambda_*(U_*-U_k)^T]}{\|U_*-U_k\|_{\mathrm{F}}^2} \cdot \|U_*-U_k\|_{\mathrm{F}}^2 \\ &\geq \delta(U_*) \cdot \|U_* - U_k\|_{\mathrm{F}}^2, \end{aligned} \tag{3.15}$$

so we obtain (3.13) from combining (3.15) and (3.8).

If $\mathcal{H}_* \succ \mathbf{O}$, it is seen from Theorem 3.6 (iii) and (3.15) that

$$\|U_* - U_k\|_{\mathrm{F}}^2 \le \frac{\delta(U_*)\|U_* - U_k\|_{\mathrm{F}}^2}{\mu_n + \gamma_\ell} \le \frac{f(U_k) - f(U_*)}{\mu_n + \gamma_\ell} \overset{(3.8)}{\le} \frac{2(\mu_1 + \gamma_1)}{\mu_n + \gamma_\ell} \cdot \epsilon_k^2 \overset{(3.1a)}{=} 2\varkappa \cdot \epsilon_k^2,$$

which yields (3.14). ∎

Remark 3.3. *Theorem 3.7 indicates that $\delta(U_*)$ plays an important role in the convergence of U_k. More precisely, U_k may converge slowly as $\delta(U_*)$ is close to zero. Specifically, U_k is difficult to define the convergence as $\delta(U_*) = 0$, which coincides with the results established in Theorem 3.6.*

• Fourth, we consider the KKT error $\|HU_k + U_k\Lambda_k + G\|_{\mathrm{F}}$ and the upper bound on $\|\Lambda_* - \Lambda_k\|_{\mathrm{F}}$, where Λ_k is defined in (2.7).

Theorem 3.8. *Denote by*

$$R_k = HU_k + U_k\Lambda_k + G.$$

Then

(i) We have that

$$\max\left\{\|R_k\|_{\mathrm{F}}, \|\Lambda_* - \Lambda_k\|_{\mathrm{F}}\right\} \le \|\mathcal{H}_*\|_2 \cdot \|U_k - U_*\|_{\mathrm{F}}. \tag{3.16}$$

(ii) If $\mathcal{H}_* \succ \mathbf{O}$ and $\|(I - \mathbf{V}_k\mathbf{V}_k^T)U_*\|_2 < 1$, then

$$\max\left\{\|R_k\|_{\mathrm{F}}, \|\Lambda_* - \Lambda_k\|_{\mathrm{F}}\right\} \le \sqrt{2}\|\mathcal{H}_*\|_2 \cdot \epsilon_k. \tag{3.17}$$

(iii) If $\mathcal{K}_q$ is an invariant subspace of H, then

$$\|R_q\|_{\mathrm{F}} = \|HU_q + U_q\Lambda_q + G\|_{\mathrm{F}} = 0 \quad \text{and} \quad -U_q^T G \succcurlyeq \mathbf{O}.$$

That is, U_q satisfies the first order optimality in Theoem 1.1 and also the necessary condition (1.4) for a global minimizer.

Proof. (i) We notice that

$$P_k = \arg\min_{P \in \mathbb{O}^{k\ell\times\ell}} \left\{\mathrm{tr}(P^T T_k P) + 2\,\mathrm{tr}(P^T G_k)\right\},$$

and $G = \mathbf{V}_k\mathbf{V}_k^T G = \mathbf{V}_k G_k$. From (2.1),

$$\begin{aligned} R_k =& HU_k + U_k\Lambda_k + G \\ =& H\mathbf{V}_k P_k + \mathbf{V}_k P_k \Lambda_k + G \\ =& \left(\mathbf{V}_k T_k + W_{k+1}N_k(E_\ell^{(k\ell)})^T\right) P_k + \mathbf{V}_k P_k\Lambda_k + \mathbf{V}_k G_k \\ =& \mathbf{V}_k\big(T_k P_k + P_k\Lambda_k + G_k\big) + W_{k+1}N_k(E_r^{(k\ell)})^T P_k \\ \overset{(2.7)}{=}& W_{k+1}N_k(E_\ell^{(k\ell)})^T P_k, \end{aligned} \tag{3.18}$$

and $\mathbf{V}_k^T R_k = \mathbf{O}$. Thus, we have from $U_k = \mathbf{V}_k P_k$ that

$$\|HU_k + U_k\Lambda_* + G\|_{\mathrm{F}}^2 = \|R_k + U_k(\Lambda_* - \Lambda_k)\|_{\mathrm{F}}^2 = \|R_k\|_{\mathrm{F}}^2 + \|\Lambda_* - \Lambda_k\|_{\mathrm{F}}^2.$$

As a result,

$$\max\left\{\|R_k\|_{\rm F}, \|\Lambda_* - \Lambda_k\|_{\rm F}\right\} \leq \|HU_k + U_k\Lambda_* + G\|_{\rm F}. \tag{3.19}$$

Notice that

$$\begin{aligned}\|HU_k + U_k\Lambda_* + G\|_{\rm F} &\overset{(1.3)}{=} \|HU_k + U_k\Lambda_* - (HU_* + U_*\Lambda_*)\|_{\rm F}\\ &= \|H(U_k - U_*) + (U_k - U_*)\Lambda_*\|_{\rm F}\\ &= \left\|\operatorname{vec}\big(H(U_k - U_*) + (U_k - U_*)\Lambda_*\big)\right\|_2\\ &= \|\left[(I_\ell \otimes H) + (\Lambda_* \otimes I_n)\right]\operatorname{vec}(U_k - U_*)\|_2 \qquad (3.20)\\ &\leq \|\mathcal{H}_*\|_2 \cdot \|U_k - U_*\|_{\rm F}.\end{aligned}$$

So we have (3.16) from (3.19) and (3.20).

(ii) If $\mathcal{H}_* = (I_\ell \otimes H) + (\Lambda_* \otimes I_n) \succ \mathbf{O}$, then

$$\begin{aligned}&\left\|\left[(I \otimes H) + (\Lambda_* \otimes I)\right]\operatorname{vec}(U_k - U_*)\right\|_2^2\\ \leq&\left\|\mathcal{H}_*^{\frac{1}{2}}\right\|_2^2 \cdot \left\|\mathcal{H}_*^{\frac{1}{2}}\operatorname{vec}(U_k - U_*)\right\|_2^2\\ =&\|\mathcal{H}_*\|_2 \cdot \left[\operatorname{vec}(U_k - U_*)\right]^T(I \otimes H + \Lambda_* \otimes I)\ \operatorname{vec}(U_k - U_*)\\ =&\|\mathcal{H}_*\|_2 \cdot \left(\operatorname{tr}[(U_k - U_*)^T H(U_k - U_*)] + \operatorname{tr}[\Lambda_*(U_k - U_*)^T(U_k - U_*)]\right)\\ \overset{(3.9)}{=}&\|\mathcal{H}_*\|_2 \cdot \big(f(U_k) - f(U_*)\big) \overset{(3.8)}{\leq} 2\|\mathcal{H}_*\|_2(\mu_1 + \gamma_1) \cdot \epsilon_k^2\\ \leq&2\|\mathcal{H}_*\|_2^2 \cdot \epsilon_k^2.\end{aligned}$$

A combination of (3.19) and (3.20) gives (3.17).

(iii) In this case, we have from (2.1) that $H\mathbf{V}_q = \mathbf{V}_q T_q$. Recall that $G = \mathbf{V}_q\mathbf{V}_q^T G = \mathbf{V}_q G_q$. Hence,

$$R_q = HU_q + U_q\Lambda_q + G = H\mathbf{V}_q P_q + \mathbf{V}_q(P_q\Lambda_q + G_q) \overset{(2.7)}{=} (H\mathbf{V}_q - \mathbf{V}_q T_q)P_q = \mathbf{O}.$$

By (1.4), $-P_q^T G_q \succcurlyeq \mathbf{O}$, and $-U_q^T G = -P_q^T\mathbf{V}_q^T\mathbf{V}_q G_q = -P_q^T G_q \succcurlyeq \mathbf{O}$. ∎

REMARK 3.4. *It is known that if the block Lanczos process terminates at the q-th step, then $\mathcal{K}_q$ is an invariant subspace of H, and q is no larger than the number of distinct eigenvalues of H [32, 33, 34]. This can happen in some applications such as the orthogonal least squares regression (OLSR) model [39] for supervised learning, where H is often a low-rank matrix.*

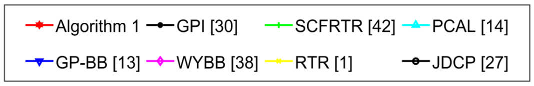
4. Numerical experiments. In this section, we perform numerical experiments to illustrate the numerical behavior of the proposed method. All the numerical experiments were run on a AMD R7 5800H CPU 3.20 GHz with 16GB RAM under Windows 11 operation system. The experimental results are obtained from using MATLAB R2022a implementation with machine precision $u_{\rm machine} \approx 2.22 \times 10^{-16}$. To show the efficiency of Algorithm 1, we compare it with seven state-of-the-art approaches for solving (1.1), including the RTR method [1], the SCFRTR method [42], the PCAL method [14], the GP-BB method [13], the WYBB method [38], the JDCP method [27], as well as the GPI method [30].

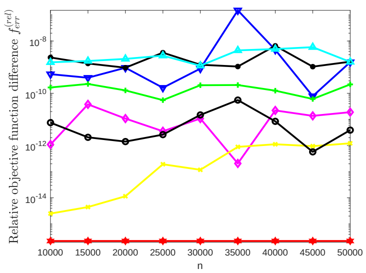
In all the experiments, we first normalize H and G by using $\|G\|_{\rm F}$, and use the following stopping criterion [27, 38, 42]

$$\frac{|f(U_k) - f(U_{k+1})|}{|f(U_k)| + 1} \leq \varepsilon_f,\ \frac{\|U_k - U_{k+1}\|_{\rm F}}{\sqrt{n}} \leq \varepsilon_U,\ \|R_k\|_{\rm F} \leq \varepsilon_g,\ \text{and } k \leq k_{\max}, \tag{4.1}$$

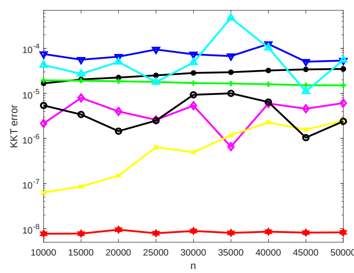
with $\varepsilon_f = 10^{-10}$, $\varepsilon_U = 10^{-6}$, $\varepsilon_g = 10^{-5}$ and $k_{\max} = 1000$.

Fig. 4.1. *Example 4.1: Numerical results on the synthetic data.*

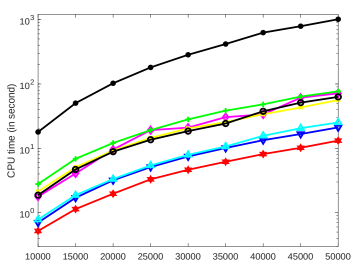




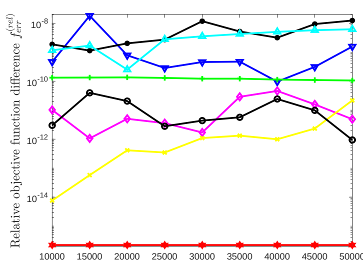
(a) $\ell = 10$: The relative objective function difference $f_{\text{err}}^{(\text{rel})}$.



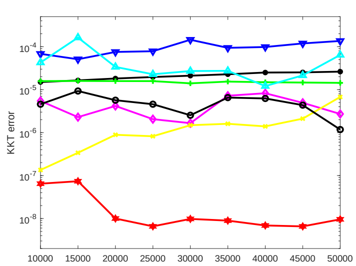
(b) $\ell = 10$: KKT error



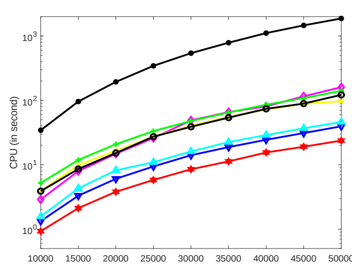
(c) $\ell = 10$: CPU time



(d) $\ell = 20$: The relative objective function difference $f_{\text{err}}^{(\text{rel})}$.



(e) $\ell = 20$: KKT error



(f) $\ell = 20$: CPU time

In the block Lanczos process, we make use of full reorthogonalization process when necessary. We stress that an advantage of the proposed method is that one can compute the KKT error cheaply. More precisely, we have from (3.18) that

$$
\begin{aligned}
\|R_k\|_{\mathrm{F}} &= \|HU_k + U_k\Lambda_k + G\|_{\mathrm{F}} \\
&= \sqrt{\|T_kP_k + P_k\Lambda_k + G_k\|_{\mathrm{F}}^2 + \left\|N_k(E_\ell^{(k\ell)})^TP_k\right\|_{\mathrm{F}}^2} \\
&= \sqrt{\|T_kP_k + P_k\Lambda_k + G_k\|_{\mathrm{F}}^2 + \left\|N_kP_k\big((k-1)\ell+1:k\ell,:\big)\right\|_{\mathrm{F}}^2}. \qquad (4.2)
\end{aligned}
$$

Moreover, $f(U_k) = \widetilde{f}(P_k)$ and $f(U_{k+1}) = \widetilde{f}(P_{k+1})$; refer to (2.6). As k increases, the main overhead in each iteration of our method lies in solving (2.4) by using the RTR method. Thus, we solve (2.4) every 5 steps in practical calculations.

To measure the accuracy of the approximations from the algorithms, we make use of the relative objective function difference defined as [3]

$$
f_{\mathrm{err}}^{(\mathrm{rel})} := \frac{f(\breve{U}_*) - f(U_{\mathrm{best}})}{|f(U_{\mathrm{best}})|}, \qquad (4.3)
$$

where $\breve{U}_*$ is the computed solution of each method and U_{best} denotes the solution with the smallest objective value among all the solvers. Thus, $f_{\mathrm{err}}^{(\mathrm{rel})} = 0$ means that $\breve{U}_* = U_{\mathrm{best}}$.

4.1. Test on synthetic data. In this subsection, we make experiments on some synthetic data generated by using the MATLAB built-in function `sprand`:

$$
H = B + B^T, \quad \text{where} \quad B = \mathtt{sprand}(\mathtt{n},\mathtt{n},\mathtt{density}), \quad G = \mathtt{randn}(\mathtt{n},\ell),
$$

where $density = 0.05$, $n = 10000, 15000, 20000, \ldots, 50000$, and $\ell = 10, 20$, respectively.

The numerical results of the eight algorithms are ploted in Figure 4.1. It is seen from the figure that both the relative objective function difference $f_{\mathrm{err}}^{(\mathrm{rel})}$ and the KKT errors of of Algorithm 1 are the smallest, and our algorithm is the fastest one among the eight algorithms.

4.2. Test on the orthogonal least squares regression for feature extraction. Orthogonal least squares regression (OLSR) is a popular supervised learning method for linear discriminant analysis (LDA) [39]. Let $A = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m] \in \mathbb{R}^{n\times m}$ be the whole database with ℓ classes, where m is the number of samples and n is the number of features. Let $\widehat{A} = [\hat{\boldsymbol{a}}_1, \ldots, \hat{\boldsymbol{a}}_{\tilde{m}}] \in \mathbb{R}^{n\times\tilde{m}}$ be training data set, and $B = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{\tilde{m}}] \in \mathbb{R}^{\ell\times\tilde{m}}$ be the corresponding class indicator matrix, where $\tilde{m}$ is the number of training samples, and $\boldsymbol{b}_i = \boldsymbol{e}_j \in \mathbb{R}^\ell$ if the sample $\hat{\boldsymbol{a}}_i$ is in the j-th class, $1 \le i \le \tilde{m}, 1 \le j \le \ell$, where $\boldsymbol{e}_j$ is the j-th column of the identity matrix. In the experiment, we randomly choose 30% of the total samples as the training set. The details of of the fourteen data sets are listed in Table 4.1.

Let $\widetilde{A}, \widetilde{B}$ be the centered matrix of $\widehat{A}, \widehat{B}$, respectively. In the OLSR method, one aims to seek $U \in \mathbb{O}^{n\times\ell}$ such that [39]

$$
\min_{U\in\mathbb{O}^{n\times\ell}} \left\{\mathrm{tr}(U^THU) + 2\mathrm{tr}(U^TG)\right\}, \qquad (4.4)
$$

where $H = \widetilde{A}^T\widetilde{A} \in \mathbb{R}^{n\times n}$ and $G = \widetilde{A}^T\widetilde{B} \in \mathbb{R}^{n\times\ell}$. We run the eight algorithms on the fourteen databases, and the numerical results are reported in Table 4.2 and Table 4.3. Specifically, if the CPU time of an algorithm exceeds 3600 seconds or the KKT error

$\|H\breve{U}_* + \breve{U}_*\breve{\Lambda}_* + G\|_{\mathrm{F}} \geq 1$, we declare that the algorithm fails to converge and denote it by "–"".

We observe from Table 4.2 and 4.3 that Algorithm 1 is more powerful than the other seven popular algorithms for solving the OLSR model (4.4). More precisely, Algorithm 1 is the best in terms of the values of $f_{\mathrm{err}}^{(\mathrm{rel})}$, and the KKT errors from Algorithm 1 is the smallest except for the `ORL` and `Text-1` databases. Indeed, the KKT errors of our algorithm is about two to five orders lower than those of the others. Moreover, our algorithm is the fastest one except for the `YouTubeFace` database, which ours is the second fastest one.

TABLE 4.1
Summary of test datasets in Example 4.2.

Datasets	Feature (n)	Number of samples (m)	Number of classes (ℓ)	Background
`ORL`[1]	10304	400	40	Image
`Yale`[2]	10000	165	15	
`YouTubeFace`[3]	16384	56653	17	
`CLL_SUB_111`[4]	11340	111	3	Biological
`SMK_CAN_187`	19993	187	2	
`GLI_85`	22283	85	2	
`leukemia`	7070	72	2	
`nci9`	9712	60	9	
`20Newsgroups`[5]	26214	18846	20	Text
`RCV1_4Class`	29992	9625	4	
`Text-1`[6]	7511	1946	2	
`Cora-HA`	3989	400	7	
`Cora-OS`	6737	1246	4	
`Core-PL`	7949	1575	9	

[1] http://featureselection.asu.edu/datasets.php.
[2] https://www.face-rec.org/databases/.
[3] https://www.cs.tau.ac.il/~wolf/ytfaces/.
[4] The databases `CLL_SUB_111, SMK_CAN_187, GLI_85, leukemia` and `nci9` are available at https://jundongl.github.io/scikit-feature/datasets.html.
[5] The databases `20Newsgroups` and `RCV1_4Class.` are available at http://www.cad.zju.edu.cn/home/dengcai/Data/TextData.html
[6] The databases `Text-1, Cora-HA, Cora-OS` and `Core-PL` are available at http://www.escience.cn/people/fpnie/papers.html.

Table 4.2

Example 4.2: Numerical experiments on the OLSR model (4.4) for some data sets in Table 4.1, where the best results are in bold.

Datasets	CPU(s)	KKT error	$f_{\mathrm{err}}^{(\mathrm{rel})}$
ORL	**Alg.1: 1.01**	8.08e-06	4.37e-13
	SCFRTR: 10.62	8.86e-06	1.07e-11
	WYBB: 31.11	4.58e-05	6.31e-08
	JDCP: 32.05	2.55e-05	8.21e-09
	RTR: 24.25	**1.49e-07**	**0**
	GP-BB: 6.85	1.74e-03	1.19e-05
	PCAL: 8.34	1.32e-03	4.76e-05
	GPI: 1.22e+02	5.30e-03	9.70e-05
Yale	**Alg.1: 1.34**	**8.46e-08**	**0**
	SCFRTR: 18.16	8.72e-06	3.44e-09
	WYBB: 50.29	4.93e-05	1.92e-07
	JDCP: 45.30	5.07e-05	1.83e-07
	RTR: 75.44	9.96e-06	6.66e-10
	GP-BB: 14.83	1.74e-03	8.03e-05
	PCAL: 14.00	8.73e-04	6.20e-05
	GPI: 1.80e+02	4.60e-03	6.76e-05
YouTubeFace	**Alg.1**: 1.60e+02	**8.46e-06**	**0**
	SCFRTR: 8.15e+02	8.72e-06	1.98e-06
	WYBB: 7.81e+02	1.10e-03	1.24e-02
	JDCP: 1.00e+03	1.70e-03	5.42e-03
	RTR: –	–	–
	GP-BB: 41.07	4.85e-03	5.81e-02
	PCAL: **29.08**	6.43e-03	7.80e-02
	GPI: 3.31e+02	9.76e-03	1.94e-01
CLL_SUB_111	**Alg.1: 2.73**	**5.19e-09**	**0**
	SCFRTR: 5.34	6.76e-08	1.70e-07
	WYBB: 30.42	1.54e-03	1.20e-03
	JDCP: 26.06	2.80e-01	26.56
	RTR: 25.58	5.96e-07	1.83e-07
	GP-BB: –	–	–
	PCAL: –	–	–
	GPI: –	–	–
SMK_CAN_187	**Alg.1**: 8.87	**3.68e-09**	**0**
	SCFRTR: 22.13	1.90e-07	5.24e-13
	WYBB: 31.81	1.26e-05	3.33e-09
	JDCP: 33.76	2.05e-05	5.50e-09
	RTR: 25.58	2.79e-08	**0**
	GP-BB: 6.69	1.52e-02	3.80e-03
	PCAL: **6.65**	1.98e-02	6.10e-03
	GPI: 2.96e+02	2.71e-02	8.25e-05
GLI_85	**Alg.1: 5.69**	**1.80e-08**	**0**
	SCFRTR: 15.84	1.00e-07	4.78e-07
	WYBB: 27.99	1.00e-02	1.30e-03
	JDCP: 22.95	9.00e-02	1.20e-01
	RTR: 45.51	4.72e-08	2.17e-07
	GP-BB: –	–	–
	PCAL: –	–	–
	GPI: –	–	–
leukemia	**Alg.1: 0.55**	**1.48e-08**	**0**
	SCFRTR: 0.79	1.34e-06	3.39e-12
	WYBB: 0.84	2.47e-04	3.42e-08
	JDCP: 0.92	1.28e-05	3.37e-10
	RTR: 1.51	6.89e-06	1.10e-10
	GP-BB: 0.58	5.39e-04	3.25e-07
	PCAL: 0.90	4.32e-03	3.55e-05
	GPI: 6.25	5.23e-04	5.50e-07
nci9	**Alg.1: 0.51**	**1.79e-08**	**0**
	SCFRTR: 2.54	5.89e-08	**0**
	WYBB: 3.70	6.80e-05	1.17e-08
	JDCP: 3.69	2.70e-05	1.50e-09
	RTR: 5.44	1.28e-07	**0**
	GP-BB: 1.71	4.47e-03	2.42e-06
	PCAL: 3.35	2.44e-03	1.19e-05
	GPI: 15.04	5.92e-04	1.59e-07
Text-1	**Alg.1: 1.20**	4.72e-06	5.92e-10
	SCFRTR: 3.01	1.07e-06	6.60e-13
	WYBB: 3.79	9.51e-06	3.54e-08
	JDCP: 3.49	9.29e-06	1.66e-08
	RTR: 28.59	**7.44e-08**	**0**
	GP-BB: 1.50	2.87e-05	3.26e-07
	PCAL: 2.20	8.90e-05	8.14e-07
	GPI: 41.36	3.73e-05	6.50e-07

TABLE 4.3
Example 4.2: Numerical experiments on the OLSR model (4.4) for some data sets in Table 4.1, where the best results are in bold.

Datasets	CPU(s)	KKT error	$f_{\mathrm{err}}^{(\mathrm{rel})}$
20Newsgroups	**Alg.1: 56.03**	**8.93e-06**	**0**
	SCFRTR: 6.25e+02	2.03e-05	6.93e-08
	WYBB: 1.87e+03	3.37e-03	3.26e-05
	JDCP: 2.36e+03	2.70e-04	5.41e-06
	RTR: –	–	–
	GPBB: 1.16e+02	1.47e-02	3.29e-02
	PCAL: 1.00e+02	1.31e-02	2.65e-02
	GPI: 4.40e+02	5.49e-02	5.73e-01
Cora_HA	**Alg.1: 0.67**	**1.06e-08**	**0**
	SCFRTR: 4.89	7.66e-06	1.59e-10
	WYBB: 3.85	4.01e-05	1.56e-07
	JDCP: 4.07	1.99e-05	4.07e-09
	RTR: 50.61	3.24e-07	5.19e-13
	GP-BB: 2.64	1.74e-04	2.65e-06
	PCAL: 2.60	7.12e-05	2.65e-07
	GPI: 30.50	3.12e-03	8.64e-04
Cora_OS	**Alg.1: 1.04**	**8.37e-07**	**0**
	SCFRTR: 10.47	7.77e-06	9.89e-10
	WYBB: 10.13	4.16e-05	5.67e-07
	JDCP: 10.59	3.31e-05	3.20e-08
	RTR: 14.49	4.52e-06	1.05e-10
	GP-BB: 4.37	1.48e-04	4.02e-06
	PCAL: 4.23	2.45e-04	1.85e-05
	GPI: 27.02	3.40e-03	8.10e-05
Cora_PL	**Alg.1: 4.64**	**2.60e-08**	**0**
	SCFRTR: 17.98	5.13e-06	2.19e-10
	WYBB: 17.35	2.08e-05	6.44e-08
	JDCP: 19.18	3.73e-05	2.04e-07
	RTR: 1.73e+02	2.95e-07	5.90e-13
	GP-BB: 7.47	1.95e-04	4.39e-06
	PCAL: 7.92	2.73e-04	1.07e-05
	GPI: 64.01	4.20e-03	1.81e-05
RCV1_4Class	**Alg.1: 17.87**	**8.05e-06**	1.23e-08
	SCFRTR: 1.37e+02	9.75e-06	**0**
	WYBB: 87.703	9.71e-06	6.73e-08
	JDCP: 1.01e+02	9.72e-06	4.45e-08
	RTR: 2.97e+03	2.72e-06	2.32e-08
	GP-BB: 28.58	1.15e-04	2.42e-06
	PCAL: 20.75	4.04e-04	1.70e-05
	GPI: 4.48e+02	6.84e-05	1.29e-06

4.3. Test on large graph clustering. Spectral clustering is a very popular unsupervised machine learning methods [36]. There are two important problems in spectral clustering [37]. First, spectral clustering consists of two successive optimization stages, i.e., spectral embedding and spectral rotation, which may not lead to globally optimal solutions. Second, for large-scale problems, it is well known that spectral clustering methods are time-consuming with very high computational complexities. In order to deal with these two challenging problems, a new framework is proposed recently to perform spectral embedding and spectral rotation simultaneously (GCSED) [37]. Unlike the OLSR model, GCSED deals with an m-dimensional problem, where m is the number of samples.

Given the database $X = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m] \in \mathbb{R}^{n\times m}$, with m samples and n features drawn from ℓ classes. Let W be a similarity matrix, and D be a diagonal matrix with the diagonal elements being the row sum of W. Denote by $\widehat{W} = D^{-\frac{1}{2}} W D^{-\frac{1}{2}} \in \mathbb{R}^{m\times m}$, and $C = D^{\frac{1}{2}} Y (Y^T D Y)^{-\frac{1}{2}} \in \mathbb{R}^{m\times \ell}$, in each iteration of the GCSED algorithm, one needs to solve the following QMPO problem [37]

$$\min_{U\in\mathbb{O}^{m\times\ell}} \left\{ \mathrm{tr}\left(U^T(-\widehat{W})U\right) + 2\gamma\ \mathrm{tr}\left(U^T(-C)\right)\right\}, \tag{4.5}$$

where $Y \in \mathbb{R}^{m\times\ell}$ is the cluster indicator matrix updated in each iteration, and γ is a constant trade-off parameter. Similar to [37], we make use of the heat kernel weighting to construct graphs, and calculate edge weights between nodes as

$$W_{i,j} = \exp\left\{-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2t^2}\right\} \quad \text{for} \quad i,j = 1,2,\ldots,m.$$

In the experiments, we choose $t = 0.1$ and $\gamma = 0.1, 1$, respectively. The data sets used in this example are summarized in Table 4.4.

The numerical results are reported in Table 4.5, where we run eight algorithms on the eleven data sets. Some remarks are given. First, we see that all the algorithms are very fast in this example. Second, Algorithm 1 is the best one in terms of $f_{\rm err}^{\rm (rel)}$ in most of the situations. Third, the proposed algorithm is the best one in terms of KKT error in most of the situations. Indeed, the accuracy of our method can be about five to eight order higher than those of the other methods. Therefore, the proposed block Lanczos method is very promising to large-scale quadratic minimization problems with orthogonality constraints.

TABLE 4.4
Summary of test data sets in Example 4.3.

Dataset	Number of samples (m)	Feature (n)	Number of classes (ℓ)	Background
`AR`[7]	1680	1200	120	Image
`YaleB`[8]	2432	4069	38	Image
`Statlog`[9]	2310	19	7	Image segmentation
`madelon`[10]	2600	500	2	Artificial
`TDT2`[11]	9394	36771	30	Audio
`MNIST`[12]	70000	784	10	Handwritten text
`USPS`[13]	9298	256	10	Handwritten text
`PenDigits`[14]	10992	16	10	Handwritten text
`Reuters`	8293	18933	65	Text
`20Newsgroups`	18846	26214	20	Text
`Letters`[15]	20000	16	4	Text

[7] http://www.cad.zju.edu.cn/home/dengcai/Data/TextData.html.
[8] http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html.
[9] https://archive.ics.uci.edu/ml/machine-learning-databases/statlog/segment/.
[10] https://jundongl.github.io/scikit-feature/datasets.html.
[11] The databases `TDT2`, `Reuters` and `20Newsgroups` are available at http://www.cad.zju.edu.cn/home/dengcai/Data/TextData.html.
[12] http://yann.lecun.com/exdb/mnist/.
[13] https://archive.ics.uci.edu/ml/index.php.
[14] http://archive.ics.uci.edu/ml.
[15] https://archive.ics.uci.edu/ml/datasets/Letter+Recognition.

Table 4.5

Numerical experiments on the QMPO problem (4.5) for the data sets in Table 4.4, where the best results are in bold.

	$\gamma = 0.1$			$\gamma = 1$		
Dataset	CPU(s)	KKT error	$f_{\text{err}}^{(\text{rel})}$	CPU(s)	KKT error	$f_{\text{err}}^{(\text{rel})}$
AR	**Alg.1**: 0.56	**3.11e-09**	**0**	0.39	**2.71e-13**	**0**
	SCFRTR: 1.57	1.29e-06	1.35e-13	0.87	5.50e-06	9.72e-12
	GBB: 0.31	5.60e-06	1.43e-12	0.32	1.38e-06	2.33e-13
	AFBB: 0.29	1.00e-05	7.18e-12	0.30	6.14e-07	1.06e-13
	RTR: 0.49	1.54e-06	1.19e-13	0.52	1.20e-08	**0**
	GP-BB: 0.17	1.75e-05	8.11e-12	0.14	5.39e-06	2.76e-12
	PCAL: 0.46	2.93e-05	2.55e-11	0.35	1.53e-05	5.79e-11
	GPI: 1.71	1.39e-06	2.68e-12	0.34	7.18e-07	3.16e-12
YaleB	**Alg.1**: 0.13	**5.82e-09**	**0**	0.09	**2.06e-11**	**0**
	SCFRTR: 0.47	7.14e-07	3.95e-14	0.22	7.14e-06	1.42e-11
	GBB: 0.14	3.77e-05	1.19e-10	0.07	1.46e-06	6.34e-13
	AFBB: 0.10	2.06e-05	1.80e-11	0.06	4.58e-06	3.15e-12
	RTR: 0.31	4.59e-08	1.15e-15	0.10	2.46e-08	1.00e-15
	GP-BB: 0.06	4.79e-05	5.31e-11	0.03	3.73e-06	2.00e-12
	PCAL: 0.10	6.64e-05	2.82e-10	0.06	3.71e-06	3.15e-12
	GPI: 0.48	2.21e-06	3.97e-12	0.08	9.27e-07	1.11e-12
Statlog	**Alg.1**: 0.01	**3.14e-09**	**0**	0.01	**6.62e-08**	**0**
	SCFRTR: 0.02	3.51e-07	5.13e-15	0.05	1.22e-05	3.91e-11
	GBB: 0.01	2.10e-05	1.64e-10	0.01	1.09e-06	2.02e-13
	AFBB: 0.01	3.08e-06	1.24e-11	0.01	7.18e-06	1.87e-11
	RTR: 0.01	1.45e-06	8.30e-13	0.01	1.90e-07	7.45e-15
	GP-BB: 0.01	9.85e-06	2.46e-11	0.01	9.85e-06	4.82e-12
	PCAL: 0.01	5.22e-05	2.33e-09	0.01	2.60e-06	1.23e-12
	GPI: 0.05	1.55e-06	1.17e-11	0.01	5.73e-07	1.61e-13
madelon	**Alg.1**: 0.002	**1.72e-13**	**0**	0.002	**1.58e-14**	8.37e-16
	SCFRTR: 0.006	**2.09e-13**	4.60e-15	0.006	3.72e-14	4.45e-16
	GBB: 0.007	8.10e-06	5.48e-12	0.003	1.43e-08	3.76e-15
	AFBB: 0.015	4.10e-08	1.29e-14	0.005	1.38e-09	4.60e-15
	RTR: 0.007	9.85e-10	1.59e-14	0.004	4.68e-09	**0**
	GP-BB: 0.004	4.20e-08	2.51e-15	0.002	5.33e-09	5.23e-15
	PCAL: 0.011	9.60e-05	7.68e-10	0.002	4.57e-05	6.96e-10
	GPI: 0.027	9.23e-06	1.29e-14	0.005	1.07e-06	1.92e-13
TDT2	**Alg.1**: 0.29	**3.43e-14**	**0**	0.28	**2.19e-13**	**0**
	SCFRTR: 0.44	4.79e-06	1.97e-12	0.35	2.81e-08	3.69e-15
	GBB: 0.28	7.82e-06	1.61e-12	0.20	4.31e-06	3.87e-12
	AFBB: 0.23	8.64e-05	5.12e-13	0.19	6.03e-07	4.28e-14
	RTR: 1.79	1.04e-08	6.50e-15	0.36	5.64e-11	2.39e-15
	GP-BB: 0.16	1.15e-04	1.09e-10	0.10	2.29e-07	1.11e-14
	PCAL: 0.34	1.23e-05	1.29e-11	0.20	1.59e-05	4.29e-11
	GPI: 1.29	2.44e-06	8.40e-12	0.20	6.87e-07	9.80e-13
MNIST	**Alg.1**: 0.15	**1.71e-12**	2.43e-14	0.11	**1.37e-13**	9.73e-15
	SCFRTR: 0.64	**1.60e-12**	7.33e-14	0.39	**2.96e-13**	**0**
	GBB: 1.05	8.72e-06	6.42e-12	0.44	3.98e-06	5.32e-12
	AFBB: 1.56	1.11e-05	1.03e-11	0.52	7.71e-06	1.90e-11
	RTR: 0.59	9.57e-12	**0**	0.38	9.67e-12	9.36e-15
	GP-BB: 0.25	1.82e-06	4.41e-14	0.15	2.05e-06	3.85e-13
	PCAL: 0.38	8.58e-05	6.13e-10	0.23	1.89e-05	1.18e-10
	GPI: 4.15	4.15e-06	1.16e-11	0.36	9.55e-07	8.54e-13
USPS	**Alg.1**: 0.13	**6.55e-07**	**0**	0.08	**5.96e-08**	**0**
	SCFRTR: 0.40	5.57e-06	1.40e-12	0.38	2.81e-05	2.45e-10
	GBB: 0.08	3.61e-05	1.11e-10	0.07	2.15e-06	1.52e-12
	AFBB: 0.07	4.02e-05	1.37e-10	0.05	7.19e-06	2.11e-11
	RTR: 0.11	5.88e-06	9.16e-13	0.05	2.03e-06	1.03e-12
	GP-BB: 0.05	1.48e-05	2.86e-12	0.03	1.73e-05	1.24e-10
	PCAL: 0.07	1.77e-04	2.66e-09	0.05	6.10e-05	6.55e-10
	GPI: 0.28	4.50e-06	1.24e-11	0.06	9.65e-07	8.14e-13
PenDigits	**Alg.1**: 0.01	**2.23e-13**	**0**	0.01	1.06e-10	2.62e-15
	SCFRTR: 0.04	6.18e-13	1.68e-15	0.04	**8.62e-12**	2.24e-15
	GBB: 0.05	1.26e-06	1.56e-13	0.08	3.21e-06	8.68e-13
	AFBB: 0.06	4.78e-08	2.43e-15	0.07	2.54e-07	1.31e-14
	RTR: 0.04	2.14e-10	6.55e-15	0.06	2.27e-10	**0**
	GP-BB: 0.01	4.14e-06	3.89e-13	0.02	2.79e-06	7.17e-13
	PCAL: 0.02	4.51e-04	1.69e-08	0.03	1.55e-04	7.99e-09
	GPI: 0.24	4.19e-06	1.14e-11	0.05	9.71e-07	2.62e-15
Reuters	**Alg.1**: 1.10	**1.32e-08**	**0**	0.40	2.28e-07	1.50e-14
	SCFRTR: 5.02	1.10e-06	6.12e-14	2.28	4.16e-06	4.62e-12
	GBB: 2.78	6.62e-05	3.69e-10	0.47	5.32e-06	3.30e-12
	AFBB: 1.68	1.97e-05	9.54e-12	0.47	1.73e-06	3.44e-13
	RTR: 2.36	2.87e-07	7.45e-12	0.95	**3.15e-09**	**0**
	GP-BB: 0.66	1.21e-05	3.51e-12	0.20	5.00e-05	2.61e-10
	PCAL: 1.29	7.50e-05	4.93e-11	0.53	5.28e-06	6.44e-12
	GPI: 6.35	1.88e-06	7.38e-12	0.48	5.30e-07	9.72e-13
20Newsgroups	**Alg.1**: 0.14	**9.89e-13**	**0**	0.18	**4.70e-13**	**0**
	SCFRTR: 0.37	2.46e-07	3.44e-15	0.26	1.68e-03	4.71e-07
	GBB: 0.82	3.98e-08	1.58e-15	0.28	1.08e-05	1.18e-11
	AFBB: 0.73	8.64e-05	5.66e-11	0.31	4.60e-06	2.37e-12
	RTR: 1.36	1.22e-06	1.66e-14	0.34	1.59e-08	4.10e-15
	GP-BB: 0.22	4.94e-06	2.03e-12	0.10	7.73e-06	1.80e-11
	PCAL: 0.58	2.32e-07	5.42e-15	0.19	2.64e-05	2.31e-10
	GPI: 1.63	3.01e-06	1.15e-11	0.28	6.87e-07	7.69e-13
Letters	**Alg.1**: 0.34	**1.59e-09**	**0**	0.28	**2.85e-10**	**0**
	SCFRTR: 2.86	3.84e-06	1.32e-12	1.32	3.26e-06	1.60e-12
	GBB: 0.82	1.41e-05	1.63e-12	0.36	2.41e-05	7.58e-11
	AFBB: 0.73	2.63e-05	5.50e-11	0.38	2.28e-06	1.32e-12
	RTR: 1.05	8.74e-09	4.47e-15	0.52	8.87e-06	2.30e-11
	GP-BB: 0.74	1.49e-04	2.11e-09	0.29	2.07e-05	5.37e-11
	PCAL: 0.58	1.15e-04	1.24e-09	0.42	5.50e-05	4.17e-10
	GPI: 2.80	3.03e-06	8.07e-12	0.38	1.22e-06	3.02e-12

5. Concluding remarks. In this paper, we propose a block Lanczos method for the large-scale quadratic minimization problems with orthogonality constraints. Convergence analysis on the optimal value, the optimal solution, the multipliers and the KKT error is given. Theoretical results show that the convergence speed of the new method strictly depends on the distribution of the spectrum of $(I_\ell \otimes H)+(\Lambda_* \otimes I_n)$. Specifically, if $(I_\ell \otimes H) + (\Lambda_* \otimes I_n) \succ \mathbf{O}$, the convergence rate of the solution from the proposed method is comparable to that of conjugate gradient method. Numerical experiments demonstrate that the new algorithm is superior to many state-of-the-art methods for large-scale QMPO in terms of accuracy, KKT error and running time, especially when $\ell \ll n$.

There are still something deserve further investigation. For instance, as the step k increases, the main workload of Algorithm 1 is to solve (2.4). The computational overhead will be prohibitive if k is large, and we have to restrict the value of k, and efficient restarting techniques [32, 33, 34] are required for our block Krylov subspace method. On the other hand, we assume that the matrix $(I_\ell \otimes H)+(\Lambda_* \otimes I_n)$ is nonsingular in the convergence analysis. An interesting topic is to weaken this assumption for the analysis.

Acknowledgement. We are grateful to Prof. Leihong Zhang and Prof. Chungen Shen for providing us some codes and databases used in numerical experiments. Meanwhile, we would like to thank Dr. Yongyan Guo for helpful discussions.

REFERENCES

[1] P. Absil, C. Baker, and K. Gallivan, *Trust-region methods on Riemannian manifolds*, Found. Comput. Math., 7 (2007), pp. 303–330.
[2] P. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
[3] 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.
[4] A. Bojanczyk and A. Lutoborski, *The Procrustes problem for orthogonal Stiefel matrices*, SIAM J. Sci. Comput., 21 (1999), pp. 1291–1304.
[5] I. Borg and J. Lingoes, *Multidimensional Similarity Structure Analysis*, Springer-Verlag, New York, 1987.
[6] Y. Carmon and J. Duchi, *First-order method for nonconvex quadratic minimization*, SIAM Rev., 62 (2020), pp. 395–436.
[7] M. Chu and N. Trendafilov, *The orthogonally constrained regression revisited*, J. Comput. Graph. Stat., 10 (2001), pp. 746–771.
[8] A. Conn, N. Gould, and P. Toint, *Trust-Region Methods*, SIAM, Philadelphia, PA, 2000.
[9] A. Edelman, T. Arias, and S. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 303–353.
[10] L. Eldén and H. Park, *A Procrustes problem on the Stiefel manifold*, Numer. Math., 82 (1999), pp. 599–619.
[11] B. Feng and G. Wu, *On convergence of the generalized Lanczos trust-region method for trust-region subproblems*, preprint, arXiv:2207.12674v1 (2022).
[12] B. Feng and G. Wu, *First-order perturbation theory of trust-region subproblems*, preprint, arXiv:2212.02744 (2022).
[13] B. Gao, X. Liu, X. Chen, and Y. Yuan, *A new first-order algorithmic framework for optimization problems with orthogonality constraints*, SIAM J. Optim., 28 (2018), pp. 302–332.
[14] B. Gao, X. Liu, and Y. Yuan, *Parallelizable algorithms for optimization problems with orthogonality constraints*, SIAM J. Sci. Comput., 41 (2019), pp. A1949–A1983.
[15] J. Van de Geer, *Linear relations among k sets of variables*, Psychometrika, 49 (1984), pp. 70–94.
[16] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, MD, 2013.
[17] J. Gower and G. Dijksterhuis, *Procrustes Problems*, Oxford University Press, New York, 2004.

[18] 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.
[19] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers in Appl. Math. 17, SIAM, Philadephia, 1997.
[20] W. HAGER, *Minimizing a quadratic over a sphere*, SIAM J. Optim., 12 (2001) pp. 188–208.
[21] R. HERZOG AND E. SACHS, *Superlinear convergence of Krylov subspace methods for self-adjoint problems in Hilbert space*, SIAM J. Numer. Anal., 53 (2015) pp. 1304–1324.
[22] R. HORN AND C. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
[23] N. HIGHAM, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.
[24] J. HU, B. JIANG, L. LIN, Z. WEN, AND Y. YUAN, *Structured quasi-Newton methods for optimization with orthogonality constraints*, SIAM J. Sci. Comput., 41 (2019), pp. A2239–A2269.
[25] J. HURLEY AND R. CATTELL, *The Procrustes program: Producing direct rotation to test a hypothesized factor structure*, Beh. Sci., 7 (1962), pp. 258–262.
[26] 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.
[27] B. JIANG AND Y. DAI, *A framework of constraint preserving update schemes for optimization on Stiefel manifold*, Math. Program., 153 (2015), pp. 535–575.
[28] X. LIU, X. WANG, AND W. WANG, *Maximization of matrix trace function of product Stiefel manifolds*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 1489–1506.
[29] J. MORÉ AND D. SORENSEN, *Computing a trust region step*, SIAM J. Sci. Statist. Comput., 4 (1983) pp. 553–572.
[30] F. NIE, R. ZHANG, AND X. LI, *A generalized power iteration method for solving quadratic problem on the Stiefel manifold*, Sci. China Inf. Sci., 60 (2017), 112101.
[31] J. NOCEDAL AND S. WRITHT, *Numerical Optimization*, 2nd edition., Springer, New York, 2006.
[32] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 2nd edition., SIAM, Philadelphia, PA, 2003.
[33] Y. SAAD, *Numerical Method for Large Eigenvalue Problems*, 2nd edition., SIAM, Philadelphia, PA, 2011.
[34] G.W. STEWART, *Matrix algorithm: Volume II: Eigensystems*, SIAM, Philadelphia, 2001.
[35] T. VIKLANDS, *Algorithms for the Weighted Orthogonal Procrustes Problem and Other Least Squares Problems*, Ph.D. thesis, Umeå University, Umeå, Sweden, 2006.
[36] U. von Luxburg, *A tutorial on spectral clustering*, Stat. Comput., 17 (2007), pp. 395–416.
[37] Z. WANG, Z. LI, R. WANG, F. NIE, AND X. LI, *Large graph clustering with simultaneous spectral embedding and discretization*, IEEE Trans. Pattern Anal. Mach. Intell., 43 (2021), pp. 4426–4440.
[38] Z. WEN AND W. YIN, *A feasible method for optimization with orthogonality constraints*, Math. Program., 142 (2013), pp. 397–434.
[39] H. ZHAO, Z. WANG, AND F. NIE, *Orthogonal least squares regression for feature extraction*, Neurocomputing, 216 (2016), pp. 200–207.
[40] Z. ZHANG AND K. DU, *Successive projection method for solving the unbalanced Procrustes problem*, Sci. China Math., 49 (2006), pp. 971–986.
[41] Z. ZHANG, Y. QIU, AND K. DU, *Conditions for optimal solutions of unbalanced Procrustes problem on Stiefel manifold*, J. Comput. Math., 25 (2007), pp. 661–671.
[42] L. ZHANG, W. YANG, C. SHEN AND J. YING, *An eigenvalue-based method for the unbalanced Procrustes problem*, SIAM J. Matrix Anal. Appl., 41 (2020), pp. 957–983.
[43] L. ZHANG,W. YANG, C. SHEN AND R. LI, *A Krylov subspace method for the large-scale second-order cone complementarity problem*, SIAM J. Sci. Comput. 37 (2015) pp. A2046–A2075.
[44] L. ZHANG, C. SHEN, W. YANG, AND J. JÚDICE, *A Lanczos method for large-scale extreme Lorentz eigenvalue problems*, SIAM J. Matrix Anal. Appl. 39 (2018) pp. 611–631.
[45] L. ZHANG, C. SHEN, AND R. LI, *On the generalized Lanczos trust-region method*, SIAM J. Optim., 27 (2017), pp. 2110–2142.
[46] L. ZHANG AND C. SHEN, *A nested Lanczos method for the trust-region subproblem*, SIAM J. Sci. Comput., 40 (2018), pp. A2005–A2032.
[47] P. ZHANG, X. LIU, S. MEMBER, J. XIONG, S. ZHOU, W. ZHAO, E. ZHU, AND Z. CAI, *Consensus one-step multi-view subspace clustering*, IEEE Transactions on Knowledge and Data Engineering, 34 (2022), pp. 4676–4689.