SIAM J. MATRIX ANAL. APPL.
Vol. 41, No. 2, pp. 432–463

RIEMANNIAN MODIFIED POLAK–RIBIÈRE–POLYAK CONJUGATE GRADIENT ORDER REDUCED MODEL BY TENSOR TECHNIQUES*

YAO-LIN JIANG† AND KANG-LI XU‡

Abstract. This paper presents a new Riemannian modified Polak–Ribière–Polyak conjugate gradient algorithm to construct the reduced systems of quadratic-bilinear systems. We eliminate the orthogonality and homogeneity constraints of the truncated $\mathcal{H}_2$ optimal model order reduction problem and turn this constrained minimization problem into an unconstrained Riemannian optimization problem on the Grassmann manifold. Due to the compactness of this manifold, the existence of the minimum solution can be guaranteed. Applying tensor techniques, the Riemannian gradient of the cost function is derived. Additionally, we design a new Riemannian MPRP conjugate gradient scheme using the differentiated retraction and the scaled vector transport. The resulting search direction always provides a descent direction. The global convergence of the proposed algorithm is established. Moreover, our algorithm is also applicable to the minimization problems of linear and bilinear systems. Finally, two numerical tests are reported to illustrate the effectiveness of the proposed algorithm.

Key words. model order reduction, Riemannian optimization, tensors, Riemannian gradient, global convergence

AMS subject classifications. 53B21, 65D99, 93C15, 15A69, 58E50

DOI. 10.1137/19M1257147

1. Introduction.

More and more attention has been paid to model order reduction (MOR) because it is effective for approximating large-scale or complex systems by lower-order systems with high accuracy, while it can preserve some important dynamical properties of the original system such as stability, structure, and passivity. Replacing the large-scale system by a reduced model provides a quick and feasible way for us to achieve the design, optimization, simulation, and control processes. Since MOR was proposed, the techniques for linear systems have been well studied, e.g., see [1, 2, 3, 4, 5, 6]. Bilinear systems as a special class of nonlinear systems are analogous to linear systems in some ways. Thus, their linear counterparts can be used as reference for the development of MOR theory and numerical methods about bilinear systems. For further details, one can see [7, 8, 9, 10, 11].

Compared with linear and bilinear cases, it may be difficult to explore the MOR methods for nonlinear systems because of their complex structures. Since many models in engineering fields are nonlinear, increasing interest is devoted to studying MOR approaches for nonlinear systems. As we know, there are different ways to deal with them. One way is to approximate a weakly nonlinear system by a bilinear system using the Carleman bilinearisation procedure [7, 9]. Then, the well-known MOR methods for bilinear systems can be used to obtain the reduced systems. The trajec-

*Received by the editors April 18, 2019; accepted for publication (in revised form) by P. Benner January 22, 2020; published electronically April 13, 2020.
https://doi.org/10.1137/19M1257147
Funding: The work of the authors was supported by National Natural Science Foundation of China grants 11871393 and 61663043 and by International Science and Technology Cooperation Program of Shaanxi Key Research & Development Plan grant 2019KWZ-08.
†Corresponding author. School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049 China (yljiang@mail.xjtu.edu.cn).
‡School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049 China (klxuxa@xjtu.edu.cn).

tory piecewise-linear (TPWL) approach proposed in [12] is another method. The idea behind the TPWL method is to represent the original nonlinear system by a combination of piecewise-linear systems and then each piece is reduced by applying the existing MOR approaches for linear systems. For the other nonlinear MOR methods, one can refer to [13, 14].

In this paper, we consider a class of popular nonlinear systems, called quadratic-bilinear (QB) systems. This kind of system was first proposed in [15] to construct the reduced models of nonlinear systems. It shows that the nonlinear system, whose nonlinear function is a linear combination of the set of element functions in integrated circuit simulation, can be transformed into an exact equivalent QB system. It is obtained by adding some additional variables and introducing algebraic equations directly from these new variables or differential equations generated by the Lie derivative. A key strength of such a system in QB form is that it is an equivalent representation of the discussed nonlinear system, without involving Taylor expansion and approximation procedure. Thus, it is applicable to weakly and strongly nonlinear systems. Another comparable advantage of the QB system is that it keeps the nonlinear function in quadratic and bilinear terms instead of cubic or higher-order terms, which can reduce the computational burden effectively. However, the equivalent representation of a nonlinear system is not unique and it can be rewritten as differential algebraic equations or ordinary differential equations (ODEs). Moreover, some literature has been dedicated to exploring the MOR techniques of QB systems [16, 17, 18].

For linear systems and their bilinear counterparts, the $\mathcal{H}_2$ norm is commonly used for measuring the MOR error between the original system and its reduced system. In [19], based on the kernels of the Volterra series expression of the input-output relationship, the authors defined the $\mathcal{H}_2$ norm and its truncated version of QB systems. In addition, some important results and relations about the truncated $\mathcal{H}_2$ norm and the truncated Gramians were derived. One can apply the truncated $\mathcal{H}_2$ norm to minimize the mismatch between the original system and its reduced one.

Manifolds are important tools for solving scientific or engineering problems. The notion of a manifold is the generalization of the concept for a smooth surface in a Euclidean space. Riemannian optimization, i.e., optimization on a Riemannian manifold, provides feasible techniques to solve the minimization problem whose constraints can be characterized as a Riemannian manifold. For an overview, one can see [20, 21] and references therein. More and more researchers have realized the importance of Riemannian optimization. The considered applications of these differential-geometric techniques mainly include the approximation or decomposition of tensors [22, 23, 24, 25], the eigenvalue problem [26], the low-rank approximation of Lyapunov equations [27], and the matrix completion problem [28].

It is notable that some Riemannian optimization techniques have been applied to the $\mathcal{H}_2$ optimal MOR problems of linear and bilinear systems [29, 30, 31, 32, 33, 34, 35, 36]. Instead of solving the $\mathcal{H}_2$ optimal MOR problem by the oblique projection, the paper [29] minimized this optimal problem over a compact subset of stable reduced models parameterized by the column-orthogonal matrices. There are a number of valid reasons for this. First, the minimization solution is guaranteed to exist over this subset. Next, the proposed algorithm is rigorously convergent and it inherits an interesting advantage of Riemannian optimization, namely, all the iterates are feasible. Additionally, it can provide a satisfactory approximation to the $\mathcal{H}_2$ optimal MOR problem in the context of the oblique projection. Other advantages of the proposed algorithm are that it has rich convergence analysis and can yield good global convergence behavior. So far, it has been an attractive topic to minimize a smooth cost function on a Riemannian manifold. Interestingly, a structure-preserving MOR method for port-Hamiltonian systems was explored based on the Riemannian optimization on a product manifold [30]. Differently, a new Riemannian optimization algorithm was proposed in [31] to reduce the port-Hamiltonian system by using the projection. Its efficiency was illustrated by a numerical example with 8000 order. Motivated by these good properties, we investigate the truncated $\mathcal{H}_2$ optimal MOR problem of QB systems via Riemannian optimization.

1.1. Contributions and outline. We focus on the truncated $\mathcal{H}_2$ optimal MOR problem of QB systems and develop a globally convergent Riemannian modified Polak–Ribière–Polyak (MPRP) conjugate gradient algorithm. Our main contributions are as follows. First, we show that the truncated $\mathcal{H}_2$ optimal MOR problem has orthogonality and homogeneity constraints. Thus, we solve this minimization problem on the Grassmann manifold. In this way, the original constrained optimization problem is turned into an unconstrained Riemannian optimization problem. Due to the compactness of the Grassmann manifold, the existence of the minimum solution is guaranteed. Second, we derive the Riemannian gradient of the cost function based on tensor techniques.

It is well known that the conjugate gradient method has the advantage of algorithmic simplicity at a satisfactory convergence speed and is more suitable for large-scale problems than second-order methods, such as the Newton's method and the trust region method. Therefore, using the differential-geometric notions on the Grassmann manifold, we design a new Riemannian MPRP conjugate gradient scheme. This is another contribution of our paper. The resulting algorithm always provides a descent direction for the cost function, which depends neither on the convexity of the cost function nor on the line-search method. Besides, the global convergence of the proposed algorithm with the common line-search conditions is established, which is a strength over the state-of-the-art methods for QB systems. The QB system may degenerate to a bilinear system or linear system as the corresponding terms are eliminated. Thus, the above findings are also applicable to linear and bilinear cases. As with the other Riemannian optimization methods for MOR, our algorithm needs to solve Sylvester equations iteratively. We adopt some strategies to address these computational issues. Two numerical tests illustrate that the proposed algorithm performs effectively under an acceptable computational cost.

The remainder of this paper is organized as follows. Some useful properties for tensors are derived and some basic definitions for QB systems are given in section 2. In section 3, we turn the optimization problem for QB systems into a minimization problem on the Grassmann manifold and the existence of its minimum solution is discussed. In section 4, we derive all the essential geometric objects and the Riemannian MPRP conjugate gradient algorithm is proposed to construct the reduced systems. The global convergence is established in section 5. Section 6 is devoted to the computational issues of the proposed algorithm. Numerical experiments are presented in section 7, and our conclusions are made in section 8.

1.2. Notation conventions. Throughout the paper, we use the following notation, where the Riemannian notation follow [20, 37]. Assume that all the matrices in this paper have compatible dimensions. We use $\mathrm{tr}(\cdot)$ for the trace operator and use $\mathrm{vec}(\cdot)$ for the vectorization of a matrix. The matrix A is said to be stable if all its eigenvalues are located on the left half open plane. The Stiefel manifold, denoted by $\mathrm{St}(n, r)$, is the set of all $n \times r$ column-orthogonal real matrices. The set of all r-dimensional linear subspaces of $\mathbb{R}^n$ is the Grassmann manifold $\mathrm{Gr}(n, r)$. The symbol

O_r denotes the orthogonal group, i.e., the set of all $r \times r$ orthogonal matrices. Let $\otimes$ denote the Kronecker product and $\odot$ be the Khatri–Rao product of two matrices with the same number of columns. We define $E^{\otimes}(X,Y) = (X \otimes Y) + (Y \otimes X)$ for matrices X and Y. The symmetric part of the matrix X is defined by $\mathrm{sym}(X) = \frac{1}{2}(X + X^T)$.

Let g be a real-valued function defined on a Riemannian manifold $\mathcal{M}$ endowed with the Riemannian metric $\langle \cdot, \cdot \rangle_x$. The corresponding induced norm is denoted by $\| \cdot \|_x$. Let $T_x\mathcal{M}$ be the tangent space of $\mathcal{M}$ at x and all the tangent spaces form the tangent bundle $T\mathcal{M}$. The differential $\mathrm{D}g(x)$ of g at x is defined as a mapping from $T_x\mathcal{M}$ to $T_{g(x)}\mathbb{R}$ such that $\mathrm{D}g(x)[\zeta] \in T_{g(x)}\mathbb{R}$ for $\zeta \in T_x\mathcal{M}$. The Riemannian gradient of g at x is denoted by grad $g(x)$. Let ϕ be a linear operator on $T_x\mathcal{M}$, and let ϕ^* be the adjoint operator of ϕ, i.e., for all $\eta, \xi \in T_x\mathcal{M}$ it holds that $\langle \eta, \phi\xi \rangle_x = \langle \phi^*\eta, \xi \rangle_x$. Let $\eta^\flat$ represent the flat of η, i.e., $\eta^\flat : T_x\mathcal{M} \to \mathbb{R} : \xi \mapsto \langle \eta, \xi \rangle_x$. The retraction $\mathcal{R}$ is a C^1 mapping from $T\mathcal{M}$ onto $\mathcal{M}$ with the following properties: (1) $\mathcal{R}_x(0_x) = x$, where 0_x is the zero element of $T_x\mathcal{M}$; (2) $\mathrm{D}\mathcal{R}_x(0_x) = id_{T_x\mathcal{M}}$, where $id_{T_x\mathcal{M}}$ denotes the identity mapping on $T_x\mathcal{M}$. The symbol $\mathcal{R}_x$ denotes the restriction of $\mathcal{R}$ to $T_x\mathcal{M}$. The vector transport $\mathcal{T}$ with associated retraction $\mathcal{R}$ is a mapping such that $\mathcal{T}_\eta\xi \in T_{\mathcal{R}_x(\eta)}\mathcal{M}$ and $\mathcal{T}_\eta$ is linear for all $\eta, \xi \in T_x\mathcal{M}$. The Ring–Wirth nonexpansive condition is that $\mathcal{T}$ satisfies $\|\mathcal{T}_\eta\xi\|_{\mathcal{R}_x(\eta)} \le \|\xi\|_x$. The vector transport $\mathcal{T}_S$ is isometric if $\langle \mathcal{T}_{S_\eta}\xi, \mathcal{T}_{S_\eta}\zeta \rangle_{\mathcal{R}_x(\eta)} = \langle \xi, \zeta \rangle_x$. The vector transport $\mathcal{T}^{\mathcal{R}}$ by differentiated retraction is defined by $\mathcal{T}_\eta^{\mathcal{R}}\xi = \mathrm{D}\mathcal{R}_x(\eta)[\xi] = \frac{d}{dt}\mathcal{R}_x(\eta + t\xi)|_{t=0}$.

2. Definitions and properties. In this section, we first give some properties of third-order tensors and then review some definitions about QB systems.

2.1. The properties of tensors. We first give the concept of matricization on tensors. Using the frontal slices $\mathcal{X}_i \in \mathbb{R}^{n_1 \times n_2}$ $(i = 1, 2, \ldots, n_3)$ of a third-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its mode-$\mu$ unfoldings are expressed as (see [38])

$$\begin{aligned}
\mathcal{X}^{(1)} &= \begin{bmatrix} \mathcal{X}_1 & \mathcal{X}_2 & \cdots & \mathcal{X}_{n_3} \end{bmatrix} \in \mathbb{R}^{n_1 \times n_2 n_3}, \\
\mathcal{X}^{(2)} &= \begin{bmatrix} \mathcal{X}_1^T & \mathcal{X}_2^T & \cdots & \mathcal{X}_{n_3}^T \end{bmatrix} \in \mathbb{R}^{n_2 \times n_3 n_1}, \\
\mathcal{X}^{(3)} &= \begin{bmatrix} \mathrm{vec}(\mathcal{X}_1) & \mathrm{vec}(\mathcal{X}_2) & \cdots & \mathrm{vec}(\mathcal{X}_{n_3}) \end{bmatrix}^T \in \mathbb{R}^{n_3 \times n_1 n_2}.
\end{aligned}$$

Another concept we want to state is the mode-μ products of the tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with matrices $U \in \mathbb{R}^{m \times n_\mu}$ $(\mu = 1, 2, 3)$, denoted by $\mathcal{X} \times_\mu U$. This generates a new tensor $\mathcal{Y}$ and its mode-μ matricization is $\mathcal{Y}^{(\mu)} = U\mathcal{X}^{(\mu)}$. For a tensor $\mathcal{G} := \mathcal{X} \times_1 U \times_2 M \times_3 N$, we further get

$$\mathcal{G}^{(1)} = U\mathcal{X}^{(1)}(N \otimes M)^T, \ \mathcal{G}^{(2)} = M\mathcal{X}^{(2)}(N \otimes U)^T, \ \mathcal{G}^{(3)} = N\mathcal{X}^{(3)}(M \otimes U)^T,$$

where $U \in \mathbb{R}^{m_1 \times n_1}$, $M \in \mathbb{R}^{m_2 \times n_2}$, and $N \in \mathbb{R}^{m_3 \times n_3}$. Obviously, the size of the tensor $\mathcal{G}$ is $m_1 \times m_2 \times m_3$. One can see [38, 39] for more basic knowledge of tensors.

A tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_1 \times n_1}$ is called symmetric if $\mathcal{X}^{(1)}(u \otimes v) = \mathcal{X}^{(1)}(v \otimes u)$ for vectors $u, v \in \mathbb{R}^{n_1}$ [16]. This yields an additional property, i.e., $\mathcal{X}^{(2)} = \mathcal{X}^{(3)}$. Denoting by $T_{(n_1,n_2)} = [I_{n_1}\ 0_{n_1 \times n_2}] \otimes [I_{n_1}\ 0_{n_1 \times n_2}]$ and $\hat{T}_{(n_1,n_2)} = [0_{n_2 \times n_1}\ I_{n_2}] \otimes [0_{n_2 \times n_1}\ I_{n_2}]$, and utilizing the symmetric tensors $\mathcal{X} \in \mathbb{R}^{n_1 \times n_1 \times n_1}$ and $\mathcal{Y} \in \mathbb{R}^{n_2 \times n_2 \times n_2}$, we construct $\mathcal{Z}^{(1)} \in \mathbb{R}^{(n_1+n_2) \times (n_1+n_2)^2}$ as

$$\mathcal{Z}^{(1)} = \begin{bmatrix} \mathcal{X}^{(1)} T_{(n_1,n_2)} \\ \mathcal{Y}^{(1)} \hat{T}_{(n_1,n_2)} \end{bmatrix}. \tag{2.1}$$

Treating $\mathcal{Z}^{(1)}$ as the mode-1 matricization of $\mathcal{Z} \in \mathbb{R}^{(n_1+n_2) \times (n_1+n_2) \times (n_1+n_2)}$, the expressions of its mode-2 and mode-3 matricizations can be obtained in Lemma 2.1.

The matrix $\mathcal{Z}^{(1)}$ is constructed to obtain the expression of the mode-2 matricization $\mathcal{H}_e^{(2)}$ in section 3.

LEMMA 2.1. *Consider a tensor $\mathcal{Z} \in \mathbb{R}^{(n_1+n_2)\times(n_1+n_2)\times(n_1+n_2)}$, whose mode-1 matricization is constructed as (2.1). Then, its mode-2 and mode-3 matricizations are equal, which are given by*

$$\mathcal{Z}^{(\mu)} = \begin{bmatrix} \mathcal{X}^{(\mu)} T_{(n_1,n_2)} \\ \mathcal{Y}^{(\mu)} \hat{T}_{(n_1,n_2)} \end{bmatrix}, \qquad \mu \in \{2,3\}.$$

Proof. Please see Appendix A for the proof of this lemma. ∎

We next introduce some properties about the trace operator, the Kronecker product, and the vectorization operator:

$$\operatorname{tr}(A_1^T A_2) = \operatorname{vec}(A_1)^T \operatorname{vec}(A_2),\ (A_1 \otimes A_2)^T = (A_1^T \otimes A_2^T),$$
$$(A_1A_2 \otimes B_1B_2) = (A_1 \otimes B_1)(A_2 \otimes B_2),\ \operatorname{vec}(A_1A_2A_3) = (A_3^T \otimes A_1)\operatorname{vec}(A_2).$$

Utilizing the above knowledge, we obtain Lemmas 2.2 and 2.3, which will be useful in deriving the Riemannian gradient in subsection 4.2.

LEMMA 2.2. *Consider two tensors $\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_2}$, $\mathcal{Y} \in \mathbb{R}^{n_2\times n_2\times n_2}$ and a matrix $U \in \mathbb{R}^{n_2\times n_1}$. Let $\mathcal{X}^{(\mu)}$ and $\mathcal{Y}^{(\mu)}$ be the mode-μ matricizations of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Then, it holds that*

$$\operatorname{tr}\big(\mathcal{X}^{(1)}(\mathcal{Y}^{(1)})^T U\big) = \operatorname{tr}\big(\mathcal{X}^{(2)}\big(\mathcal{Y}^{(2)}(I_{n_2}\otimes U)\big)^T\big) = \operatorname{tr}\big(\mathcal{X}^{(3)}\big(\mathcal{Y}^{(3)}(I_{n_2}\otimes U)\big)^T\big). \tag{2.2}$$

Particularly, when U is the identity matrix I_{n_2}, it holds that

$$\operatorname{tr}\big(\mathcal{X}^{(1)}(\mathcal{Y}^{(1)})^T\big) = \operatorname{tr}\big(\mathcal{X}^{(2)}(\mathcal{Y}^{(2)})^T\big) = \operatorname{tr}\big(\mathcal{X}^{(3)}(\mathcal{Y}^{(3)})^T\big). \tag{2.3}$$

Proof. Using the ith frontal slices of $\mathcal{X}$ and $\mathcal{Y}$, denoted by $\mathcal{X}_i$ and $\mathcal{Y}_i$, it follows that

$$\begin{aligned}
\operatorname{tr}\big(\mathcal{X}^{(1)}(\mathcal{Y}^{(1)})^T U\big) &= \operatorname{tr}\Big([\mathcal{X}_1\ \mathcal{X}_2\ \cdots\ \mathcal{X}_{n_2}]\ [\mathcal{Y}_1\ \mathcal{Y}_2\ \cdots\ \mathcal{Y}_{n_2}]^T U\Big) \\
&= \sum_{i=1}^{n_2} \operatorname{tr}\left(\mathcal{X}_i \mathcal{Y}_i^T U\right) = \sum_{i=1}^{n_2} \operatorname{tr}\left(\mathcal{X}_i^T U^T \mathcal{Y}_i\right) \\
&= \operatorname{tr}\Big([\mathcal{X}_1^T\ \mathcal{X}_2^T\ \cdots\ \mathcal{X}_{n_2}^T]\ \big([\mathcal{Y}_1^T\ \mathcal{Y}_2^T\ \cdots\ \mathcal{Y}_{n_2}^T](I_{n_2}\otimes U)\big)^T\Big) \\
&= \operatorname{tr}\Big(\mathcal{X}^{(2)}\big(\mathcal{Y}^{(2)}(I_{n_2}\otimes U)\big)^T\Big).
\end{aligned}$$

Since $\operatorname{vec}(U^T\mathcal{Y}_i) = (I_{n_2}\otimes U)^T\operatorname{vec}(\mathcal{Y}_i)$ and $\operatorname{tr}\left(\mathcal{X}_i^T U^T \mathcal{Y}_i\right) = \operatorname{vec}(\mathcal{X}_i)^T\operatorname{vec}(U^T\mathcal{Y}_i)$, one has

$$\begin{aligned}
&\operatorname{tr}\Big(\mathcal{X}^{(3)}\big(\mathcal{Y}^{(3)}(I_{n_2}\otimes U)\big)^T\Big) \\
&= \operatorname{tr}\Big(\mathcal{X}^{(3)}\big[(I_{n_2}\otimes U^T)\operatorname{vec}(\mathcal{Y}_1)\ \ \cdots\ \ (I_{n_2}\otimes U^T)\operatorname{vec}(\mathcal{Y}_{n_2})\big]\Big) \\
&= \operatorname{tr}\Big(\big[\operatorname{vec}(\mathcal{X}_1)\ \cdots\ \operatorname{vec}(\mathcal{X}_{n_2})\big]^T\ \big[\operatorname{vec}(U^T\mathcal{Y}_1)\ \cdots\ \operatorname{vec}(U^T\mathcal{Y}_{n_2})\big]\Big) \\
&= \sum_{i=1}^{n_2}\operatorname{vec}(\mathcal{X}_i)^T\operatorname{vec}(U^T\mathcal{Y}_i) = \sum_{i=1}^{n_2}\operatorname{tr}\left(\mathcal{X}_i^T U^T\mathcal{Y}_i\right) = \operatorname{tr}\Big(\mathcal{X}^{(1)}(\mathcal{Y}^{(1)})^T U\Big).
\end{aligned}$$

Thereby, the results in (2.2) are proved. Moreover, the results in (2.3) can be derived when U is the identity matrix I_{n_2}. ∎

The above lemma generalizes Lemma 2.2 in [19]. Furthermore, we obtain the following lemma.

LEMMA 2.3. *Consider tensors $\mathcal{X} \in \mathbb{R}^{n_1 \times n_1 \times n_1}$, $\mathcal{Y} \in \mathbb{R}^{n_2 \times n_2 \times n_2}$ and matrices $U \in \mathbb{R}^{n_2 \times n_1}$, $M \in \mathbb{R}^{n_1 \times n_2}$, and $N \in \mathbb{R}^{n_1 \times n_2}$. Assume that $\mathcal{X}$ and $\mathcal{Y}$ are symmetric; then, we have*

$$\operatorname{tr}\Big(\mathcal{X}^{(1)}(M \otimes N)(\mathcal{Y}^{(1)})^T U\Big) = \operatorname{tr}\Big(\mathcal{X}^{(1)}(N \otimes M)(\mathcal{Y}^{(1)})^T U\Big). \tag{2.4}$$

Proof. We view $\mathcal{X}^{(1)}(M \otimes N)$ as the mode-1 matricization of the tensor $\mathcal{X} \times_2 N^T \times_3 M^T$. Since $\mathcal{X}$ and $\mathcal{Y}$ are symmetric, one has $\mathcal{X}^{(2)} = \mathcal{X}^{(3)}$ and $\mathcal{Y}^{(2)} = \mathcal{Y}^{(3)}$. Using Lemma 2.2, it follows that

$$\begin{aligned}
\operatorname{tr}\Big(\mathcal{X}^{(1)}(M \otimes N)(\mathcal{Y}^{(1)})^T U\Big) &= \operatorname{tr}\Big(N^T \mathcal{X}^{(2)}(M \otimes I_{n_1})\big(\mathcal{Y}^{(2)}(I_{n_2} \otimes U)\big)^T\Big) \\
&= \operatorname{tr}\Big(N^T \mathcal{X}^{(2)}(M \otimes U^T)(\mathcal{Y}^{(2)})^T\Big) \\
&= \operatorname{tr}\Big(N^T \mathcal{X}^{(3)}(M \otimes U^T)(\mathcal{Y}^{(3)})^T\Big) \\
&= \operatorname{tr}\Big(\mathcal{X}^{(1)}(N \otimes M)(\mathcal{Y}^{(1)})^T U\Big),
\end{aligned}$$

which completes the proof. ∎

2.2. Truncated $\mathcal{H}_2$ norm of QB systems. We now briefly review some definitions about QB systems. In view of the fact that some practical examples can be represented by QB ODE systems, e.g., Burgers' equation and some nonlinear transmission line circuits, we consider the QB system with the form

$$\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + H\big(x(t) \otimes x(t)\big) + \displaystyle\sum_{k=1}^{m} N_k x(t) u_k(t) + Bu(t), \\ y(t) = Cx(t), \ x(0) = 0, \end{cases} \tag{2.5}$$

where $A \in \mathbb{R}^{n \times n}, N_k \in \mathbb{R}^{n \times n}$ for $k = 1, 2, \ldots, m$, the matrix $H \in \mathbb{R}^{n \times n^2}$ induced by Hessian, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{p \times n}$. The variables $x(t) \in \mathbb{R}^n$ and $y(t) \in \mathbb{R}^p$ denote the state and output of the system (2.5), respectively; the function $u_k(t)$ is the kth element of the input $u(t)$. Clearly, the system (2.5) is bilinear if $H = 0$, and it further degenerates to a linear system as $H = 0$ and $N_k = 0$.

In the following, the truncated $\mathcal{H}_2$ norm for the QB system (2.5) is defined.

DEFINITION 2.4 (see [19]). *Let Σ be the QB system (2.5) with the stable matrix A. Its truncated $\mathcal{H}_2$ norm is defined as*

$$\|\Sigma\|^2_{\mathcal{H}_2^{\mathcal{T}}} = \operatorname{tr}\left(\sum_{i=1}^{3} \int_0^{+\infty} \cdots \int_0^{+\infty} C g_i(t_1, \ldots, t_i) g_i^T(t_1, \ldots, t_i) C^T dt_1 \cdots dt_i\right),$$

where $g_1(t_1) = e^{At_1}B$, $g_2(t_1, t_2) = e^{At_2}[N_1 \ N_2 \ \cdots \ N_m](I_m \otimes e^{At_1}B)$, and $g_3(t_1, t_2, t_3) = e^{At_3}H(e^{At_2}B \otimes e^{At_1}B)$.

Note that the matrix H in (2.5) can be viewed as the mode-1 matricization of a third-order tensor $\mathcal{H} \in \mathbb{R}^{n \times n \times n}$, i.e., $H = \mathcal{H}^{(1)}$. Following the matricization process in [38], the mode-2 and mode-3 matricizations of $\mathcal{H}$ can be obtained, which are denoted by $\mathcal{H}^{(2)}$ and $\mathcal{H}^{(3)}$, respectively. As shown in [16], the tensor $\mathcal{H}$ can be modified such that the resulting tensor $\tilde{\mathcal{H}}$ is symmetric without changing the dynamics of $x(t)$ in (2.5). Without loss of generality, this paper also assumes that $\mathcal{H}$ is symmetric. It follows that $\mathcal{H}^{(2)} = \mathcal{H}^{(3)}$.

The truncated controllability Gramian $P_{\mathcal{T}}$ is defined as

$$P_{\mathcal{T}} = \sum_{i=1}^{3} P_i,\ P_i = \int_0^{+\infty} \cdots \int_0^{+\infty} g_i(t_1, \ldots, t_i) g_i^T(t_1, \ldots, t_i) dt_1 \cdots dt_i, \tag{2.6}$$

and the truncated observability Gramian $Q_{\mathcal{T}}$ is defined as

$$Q_{\mathcal{T}} = \sum_{i=1}^{3} Q_i,\ Q_i = \int_0^{+\infty} \cdots \int_0^{+\infty} h_i(t_1, \ldots, t_i) h_i^T(t_1, \ldots, t_i) dt_1 \cdots dt_i, \tag{2.7}$$

where $h_1(t_1) = e^{A^T t_1} C^T$, $h_2(t_1, t_2) = e^{A^T t_2}[N_1^T \ \cdots \ N_m^T](I_m \otimes e^{A^T t_1} C^T)$, and $h_3(t_1, t_2, t_3) = e^{A^T t_3} \mathcal{H}^{(2)}(e^{A t_1} B \otimes e^{A^T t_2} C^T)$. Then for the stable matrix A, we have

$$AP_{\mathcal{T}} + P_{\mathcal{T}} A^T + H(P_1 \otimes P_1) H^T + \sum_{k=1}^{m} N_k P_1 N_k^T + BB^T = 0, \tag{2.8a}$$

$$A^T Q_{\mathcal{T}} + Q_{\mathcal{T}} A + \mathcal{H}^{(2)}(P_1 \otimes Q_1)(\mathcal{H}^{(2)})^T + \sum_{k=1}^{m} N_k^T Q_1 N_k + C^T C = 0, \tag{2.8b}$$

where

$$AP_1 + P_1 A^T + BB^T = 0,\ A^T Q_1 + Q_1 A + C^T C = 0.$$

Additionally, the matrices P_i and Q_i are the solutions of the following equations:

$$AP_2 + P_2 A^T + \sum_{k=1}^{m} N_k P_1 N_k^T = 0,\ AP_3 + P_3 A^T + H(P_1 \otimes P_1) H^T = 0,$$
$$A^T Q_2 + Q_2 A + \sum_{k=1}^{m} N_k^T Q_1 N_k = 0,\ A^T Q_3 + Q_3 A + \mathcal{H}^{(2)}(P_1 \otimes Q_1)(\mathcal{H}^{(2)})^T = 0.$$

One can refer to [40] for more details.

Obviously, Definition 2.4 makes sense and the truncated Gramians exist when the matrix A in (2.5) is stable. Therefore, we assume that all the QB systems discussed in this paper have the stable matrix A. The truncated $\mathcal{H}_2$ norm can be computed with the help of the truncated Gramians $P_{\mathcal{T}}$ and $Q_{\mathcal{T}}$, i.e., it holds that

$$\|\Sigma\|_{\mathcal{H}_2^{\mathcal{T}}}^2 = \operatorname{tr}(C P_{\mathcal{T}} C^T) = \operatorname{tr}(B^T Q_{\mathcal{T}} B). \tag{2.9}$$

From Definition 2.4, we can obtain the truncated version of the $\mathcal{H}_2$ norm for bilinear systems [8] and the corresponding $\mathcal{H}_2$ norm of linear systems [1, 2]. Hence, the formula (2.9) also holds for bilinear and linear systems. For the bilinear case, $P_T = P_1 + P_2$ and $Q_T = Q_1 + Q_2$, which still satisfy (2.8) with $H = 0$. For the linear case, the formula (2.9) holds with $P_T = P_1$ and $Q_T = Q_1$.

3. Riemannian optimization problem on the Grassmann manifold and the existence of its minimum. We now discuss the truncated $\mathcal{H}_2$ optimal MOR problem. Our goal is to construct the reduced system of order r ($r \ll n$),

$$\hat{\Sigma}: \begin{cases} \dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{H}\big(\hat{x}(t)\otimes\hat{x}(t)\big) + \sum_{k=1}^{m}\hat{N}_k\hat{x}(t)u_k(t) + \hat{B}u(t), \\ \hat{y}(t) = \hat{C}\hat{x}(t), \end{cases} \tag{3.1}$$

such that it is the truncated $\mathcal{H}_2$ optimal in the context of the orthogonal projection, where $x(t) \approx V\hat{x}(t)$ with $V^TV = I_r$ and the matrices are

$$\begin{aligned} &\hat{A} = V^TAV \in \mathbb{R}^{r\times r},\ \hat{H} = V^TH(V\otimes V) \in \mathbb{R}^{r\times r^2}, \\ &\hat{N}_k = V^TN_kV \in \mathbb{R}^{r\times r},\ \hat{B} = V^TB \in \mathbb{R}^{r\times m},\ \hat{C} = CV \in \mathbb{R}^{p\times r}. \end{aligned} \tag{3.2}$$

For this, we consider the truncated $\mathcal{H}_2$ norm of the error system $\Sigma_e := \Sigma - \hat{\Sigma}$, which is defined as

$$A_e = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, N_{k,e} = \begin{bmatrix} N_k & 0 \\ 0 & \hat{N}_k \end{bmatrix}, H_e = \begin{bmatrix} HT_{(n,r)} \\ \hat{H}\hat{T}_{(n,r)} \end{bmatrix}, B_e = \begin{bmatrix} B \\ \hat{B} \end{bmatrix},\ C_e = [C \ \ -\hat{C}].$$

Since $\mathcal{H}$ is symmetric, $\hat{\mathcal{H}}$ is also symmetric and satisfies $\hat{\mathcal{H}}^{(2)} = \hat{\mathcal{H}}^{(3)}$. In the following, the matrix $H_e \in \mathbb{R}^{(n+r)\times(n+r)^2}$ is viewed as the mode-1 matricization of the tensor $\mathcal{H}_e$ with size $(n+r)\times(n+r)\times(n+r)$. From Lemma 2.1, we can get the explicit expression of its mode-2 matricization $\mathcal{H}_e^{(2)}$.

For the error system Σ_e, its truncated Gramians

$$P_{\mathcal{T}e} = \begin{bmatrix} P_{\mathcal{T}} & X_{\mathcal{T}} \\ X_{\mathcal{T}}^T & \hat{P}_{\mathcal{T}} \end{bmatrix},\ Q_{\mathcal{T}e} = \begin{bmatrix} Q_{\mathcal{T}} & Y_{\mathcal{T}} \\ Y_{\mathcal{T}}^T & \hat{Q}_{\mathcal{T}} \end{bmatrix} \tag{3.3}$$

are the solutions of the Lyapunov equations

$$\begin{aligned} &A_eP_{\mathcal{T}e} + P_{\mathcal{T}e}A_e^T + H_e(P_{1e}\otimes P_{1e})H_e^T + \sum_{k=1}^{m}N_{k,e}P_{1e}N_{k,e}^T + B_eB_e^T = 0, \\ &A_e^TQ_{\mathcal{T}e} + Q_{\mathcal{T}e}A_e + \mathcal{H}_e^{(2)}(P_{1e}\otimes Q_{1e})(\mathcal{H}_e^{(2)})^T + \sum_{k=1}^{m}N_{k,e}^TQ_{1e}N_{k,e} + C_e^TC_e = 0, \end{aligned} \tag{3.4}$$

with

$$P_{1e} = \begin{bmatrix} P_1 & X_1 \\ X_1^T & \hat{P}_1 \end{bmatrix},\ Q_{1e} = \begin{bmatrix} Q_1 & Y_1 \\ Y_1^T & \hat{Q}_1 \end{bmatrix}, \tag{3.5}$$

which satisfy the Lyapunov equations

$$A_eP_{1e} + P_{1e}A_e^T + B_eB_e^T = 0,\ A_e^TQ_{1e} + Q_{1e}A_e + C_e^TC_e = 0. \tag{3.6}$$

We denote the square of the truncated $\mathcal{H}_2$ error by $\tilde{f}(V) = \|\Sigma_e\|^2_{\mathcal{H}_2^{\mathcal{T}}}$. According to (2.9), we can find that

$$\tilde{f}(V) = \mathrm{tr}(C_eP_{\mathcal{T}e}C_e^T) = \mathrm{tr}(CP_{\mathcal{T}}C^T - 2CX_{\mathcal{T}}\hat{C}^T + \hat{C}\hat{P}_{\mathcal{T}}\hat{C}^T), \tag{3.7}$$

where $X_{\mathcal{T}}$ and $\hat{P}_{\mathcal{T}}$ respectively satisfy

$$AX_{\mathcal{T}} + X_{\mathcal{T}}\hat{A}^T + H(X_1\otimes X_1)\hat{H}^T + \sum_{k=1}^{m}N_kX_1\hat{N}_k^T + B\hat{B}^T = 0, \tag{3.8}$$

$$\hat{A}\hat{P}_{\mathcal{T}} + \hat{P}_{\mathcal{T}}\hat{A}^T + \hat{H}(\hat{P}_1 \otimes \hat{P}_1)\hat{H}^T + \sum_{k=1}^{m} \hat{N}_k\hat{P}_1\hat{N}_k^T + \hat{B}\hat{B}^T = 0, \tag{3.9}$$

in which X_1 and $\hat{P}_1$ are determined by the equations

$$AX_1 + X_1\hat{A}^T + B\hat{B}^T = 0,\ \hat{A}\hat{P}_1 + \hat{P}_1\hat{A}^T + \hat{B}\hat{B}^T = 0. \tag{3.10}$$

In order to find the reduced system (3.1) that can minimize the truncated $\mathcal{H}_2$ error over all the r order stable reduced models generated by the orthogonal projection, we solve the minimization problem

$$\text{minimize} \quad \tilde{f}(V) \quad \text{subject to} \quad V^TV = I_r. \tag{3.11}$$

It implies that the domain of $\tilde{f}$ is $\mathrm{St}(n, r)$.

Reusing (2.9), we find that $\tilde{f}(V)$ can also be expressed as

$$\tilde{f}(V) = \mathrm{tr}(B^TQ_{\mathcal{T}}B + 2B^TY_{\mathcal{T}}\hat{B} + \hat{B}^T\hat{Q}_{\mathcal{T}}\hat{B}),$$

where $Y_{\mathcal{T}}$ and $\hat{Q}_{\mathcal{T}}$ are the solutions of the equations

$$A^TY_{\mathcal{T}} + Y_{\mathcal{T}}\hat{A} + \mathcal{H}^{(2)}(X_1 \otimes Y_1)(\hat{\mathcal{H}}^{(2)})^T + \sum_{k=1}^{m} N_k^TY_1\hat{N}_k - C^T\hat{C} = 0, \tag{3.12}$$

$$\hat{A}^T\hat{Q}_{\mathcal{T}} + \hat{Q}_{\mathcal{T}}\hat{A} + \hat{\mathcal{H}}^{(2)}(\hat{P}_1 \otimes \hat{Q}_1)(\hat{\mathcal{H}}^{(2)})^T + \sum_{k=1}^{m} \hat{N}_k^T\hat{Q}_1\hat{N}_k + \hat{C}^T\hat{C} = 0, \tag{3.13}$$

in which

$$A^TY_1 + Y_1\hat{A} - C^T\hat{C} = 0,\ \hat{A}^T\hat{Q}_1 + \hat{Q}_1\hat{A} + \hat{C}^T\hat{C} = 0. \tag{3.14}$$

Equations (3.12)–(3.14) are derived based on Lemma 2.1 and for brevity we show the derivation in the first part of Appendix B. Since the above two expressions of $\tilde{f}(V)$ are equivalent, we adopt the expression in (3.11) to explore the minimization problem.

Clearly, the minimization problem (3.11) has an orthogonality constraint, since V is column-orthogonal. Further, we demonstrate that it characterizes the homogeneity constraint.

THEOREM 3.1. *Let Σ be the QB system* (2.5) *and $\hat{\Sigma}$ be the reduced QB system* (3.1) *constructed by the column-orthogonal matrix V, where the matrices A and $\hat{A}$ are stable. Then, it has $\tilde{f}(V) = \tilde{f}(VO)$ for arbitrary matrix O belonging to O_r.*

Proof. Let $\tilde{\Sigma}$ be the reduced QB system generated by VO, where its state space is constructed by the system matrices $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{H}$, and $\tilde{N}_k$ for $k = 1, 2, \ldots, m$. According to the constructions of $\hat{\Sigma}$ and $\tilde{\Sigma}$, one gets $\tilde{A} = O^T\hat{A}O$, $\tilde{B} = O^T\hat{B}$, $\tilde{C} = \hat{C}O$, $\tilde{H} = O^T\hat{H}(O \otimes O)$, and $\tilde{N}_k = O^T\hat{N}_kO$. We denote the error system of Σ and $\tilde{\Sigma}$ by $\tilde{\Sigma}_e$, where its system matrices are

$$\tilde{A}_e = \begin{bmatrix} A & 0 \\ 0 & \tilde{A} \end{bmatrix}, \tilde{N}_{k,e} = \begin{bmatrix} N_k & 0 \\ 0 & \tilde{N}_k \end{bmatrix}, \tilde{H}_e = \begin{bmatrix} HT_{(n,r)} \\ \tilde{H}\hat{T}_{(n,r)} \end{bmatrix}, \tilde{B}_e = \begin{bmatrix} B \\ \tilde{B} \end{bmatrix},\ \tilde{C}_e = [C \ \ -\tilde{C}].$$

Subsequently, it holds that $\tilde{A}_e = O_e^TA_eO_e$, $\tilde{N}_{k,e} = O_e^TN_{k,e}O_e$, $\tilde{B}_e = O_e^TB_e$, and $\tilde{C}_e = C_eO_e$, where $O_e = \mathrm{diag}\{I_n, O\}$. Besides, one has $T_{(n,r)}(O_e \otimes O_e) = T_{(n,r)}$ and $(O \otimes O)\hat{T}_{(n,r)} = \tilde{T}_{(n,r)}(O_e \otimes O_e)$, which follows that $\tilde{H}_e = O_e^TH_e(O_e \otimes O_e)$.

Notice that the equations about P_{ie} $(i \in \{1, \mathcal{T}\})$ in (3.4) and (3.6) can be represented as

$$A_e P_{ie} + P_{ie} A_e^T + F_{ie} = 0 \tag{3.15}$$

for the corresponding F_{ie}. Similarly, it is observed that the Lyapunov equations of the matrices $\tilde{P}_{ie}$ $(i \in \{1, \mathcal{T}\})$ for $\tilde{\Sigma}_e$ take the form

$$\tilde{A}_e \tilde{P}_{ie} + \tilde{P}_{ie} \tilde{A}_e^T + \tilde{F}_{ie} = 0, \tag{3.16}$$

where $\tilde{F}_{1e} = \tilde{B}_e \tilde{B}_e^T$ and $\tilde{F}_{\mathcal{T}e} = \tilde{H}_e(\tilde{P}_{1e} \otimes \tilde{P}_{1e})\tilde{H}_e^T + \sum_{k=1}^m \tilde{N}_{k,e} \tilde{P}_{1e} \tilde{N}_{k,e}^T + \tilde{B}_e \tilde{B}_e^T$. Using the relationships of the system matrices for Σ_e and $\tilde{\Sigma}_e$, we further find that $\tilde{F}_{ie} = O_e^T F_{ie} O_e$. Then, from (3.16), one has

$$A_e O_e \tilde{P}_{ie} O_e^T + O_e \tilde{P}_{ie} O_e^T A_e^T + F_{ie} = 0,$$

which implies that $P_{ie} = O_e \tilde{P}_{ie} O_e^T$ for $i \in \{1, \mathcal{T}\}$ due to the unique solution of (3.15). Thus, the cost function for $\tilde{\Sigma}$ satisfies $\tilde{f}(VO) = \operatorname{tr}(\tilde{C}_e \tilde{P}_{\mathcal{T}e} \tilde{C}_e^T) = \tilde{f}(V)$. This concludes the proof. ∎

Theorem 3.1 shows that a feasible point of the minimization problem (3.11) is an equivalence class $[V]$ represented by $V \in \operatorname{St}(n, r)$, i.e.,

$$[V] = \{VO | V \in \operatorname{St}(n, r), O \in O_r\}.$$

In this paper, the Grassmann manifold $\operatorname{Gr}(n, r)$ is represented as a quotient manifold of the Stiefel manifold $\operatorname{St}(n, r)$ with the canonical projection $\pi : V \in \operatorname{St}(n, r) \mapsto [V] \in \operatorname{Gr}(n, r)$, i.e., $\operatorname{Gr}(n, r) = \operatorname{St}(n, r)/O_r$. So we conclude that the minimization problem (3.11) can be considered on $\operatorname{Gr}(n, r)$. Using the function $\tilde{f}$, we have one and only one function $f : [V] \in \operatorname{Gr}(n, r) \mapsto \tilde{f}(V) \in \mathbb{R}$. Thus, we solve the Riemannian optimization problem

$$\text{minimize} \quad f([V]) \quad \text{subject to} \quad [V] \in \operatorname{Gr}(n, r). \tag{3.17}$$

It is natural to consider *the existence of the minimum* of the Riemannian optimization problem (3.17). Since f is continuous and $\operatorname{Gr}(n, r)$ is compact, according to the extreme value theorem, there exists a point $[V^{opt}] \in \operatorname{Gr}(n, r)$ such that

$$f([V^{opt}]) = \inf_{[V] \in \operatorname{Gr}(n,r)} f([V]) < +\infty,$$

which implies that a minimum solution of the optimization problem (3.17) exists on $\operatorname{Gr}(n, r)$.

Remark 1. Notice that the stability of $\hat{A}$ is a basic requirement for the $\mathcal{H}_2$ optimal MOR problem. Generally, it cannot be ensured automatically by the Petrov–Galerkin approach even if $A + A^T$ is negative definite. However, this requirement can be achieved by the Galerkin approach since $V(A + A^T)V^T$ is negative definite, which implies that $\hat{A}$ is stable. Moreover, the above analysis shows that constructing the reduced QB system (3.1) by an orthogonal projection allows us to pose the $\mathcal{H}_2$ optimal MOR problem on a Grassmann manifold. Meanwhile, combining with its compactness, we can guarantee the existence of the minimum. Hence, we focus on generating the reduced QB system (3.1) by the Galerkin approach.

4. The Riemannian MPRP conjugate gradient algorithm on the Grassmann manifold. In order to solve the minimization problem (3.17), we need to give all the necessary Riemannian geometry concepts on Grassmann manifolds and all the closed-form expressions that will be used in our algorithm. It is worth mentioning that the Riemannian gradient of the cost function f is derived and a new Riemannian MPRP conjugate gradient scheme is designed. We refer to [20] for the theory of Riemannian manifolds.

4.1. The tangent space. The Euclidean space $\mathbb{R}^{n\times r}$ is a vector space and it thus is a linear Riemannian manifold endowed with the inner product

$$\langle \xi_1, \xi_2\rangle_V = \mathrm{tr}(\xi_2^T\xi_1) \quad \forall\ \xi_1, \xi_2 \in \mathbb{R}^{n\times r},$$

whereas $\mathrm{St}(n, r)$ and $\mathrm{Gr}(n, r)$ are not flat. Their tangent spaces, which are vector spaces with the elements in $\mathbb{R}^{n\times r}$, can be viewed as the linear approximations of manifolds.

Let $T_V\mathrm{St}(n, r)$ be the tangent space of $\mathrm{St}(n, r)$ at V. The tangent space $T_V\mathrm{St}(n, r)$ is decomposed into the direct sum of the vertical space $\mathcal{V}_V$ and the horizontal space $\mathcal{H}_V$. The former contains directions for movement within the equivalence class $[V]$, while the latter contains directions for movement into a new equivalence class. The elements in $\mathcal{V}_V$ take the form of VS with $S^T = -S \in \mathbb{R}^{r\times r}$, and $\mathcal{H}_V$ is given by

$$\mathcal{H}_V = \{V_\perp K|\ K \in \mathbb{R}^{(n-r)\times r}\},$$

where $V_\perp$ denotes the matrix such that $[V\ V_\perp] \in O_n$. Intuitively, the element in $\mathcal{H}_V$ is specified by the orthogonal projection (see [20, 23])

$$\Pi_V = I_n - VV^T = V_\perp V_\perp^T. \tag{4.1}$$

For Riemannian optimization on the quotient manifold, the vertical and horizontal spaces play important roles in computation. We denote by $T_{[V]}\mathrm{Gr}(n, r)$ the tangent space of $\mathrm{Gr}(n, r)$ at $[V]$. For any tangent vector $\zeta \in T_{[V]}\mathrm{Gr}(n, r)$, there exists one and only one vector $\bar{\zeta} \in \mathcal{H}_V$ satisfying $\mathrm{D}\pi(V)[\bar{\zeta}] = \zeta$. The element $\bar{\zeta}$ is called the horizontal lift of ζ at V, which is used as a matrix representation of ζ in numerical computation. Obviously, it holds that $V^T\bar{\zeta} = 0$. Equipped with the Riemannian metric

$$\langle\cdot,\cdot\rangle_{[V]} : T_{[V]}\mathrm{Gr}(n, r) \times T_{[V]}\mathrm{Gr}(n, r) \to \mathbb{R}, \quad \langle\zeta_1, \zeta_2\rangle_{[V]} = \mathrm{tr}(\bar{\zeta}_2^T\bar{\zeta}_1), \tag{4.2}$$

the manifold $\mathrm{Gr}(n, r)$ is turned into a Riemannian manifold, where $\bar{\zeta}_1$ and $\bar{\zeta}_2$ are the horizontal lifts of ζ_1 and ζ_2, respectively. Here, the induced norm is defined as $\|\zeta\|_{[V]} = \sqrt{\langle\zeta, \zeta\rangle_{[V]}}$.

4.2. The Riemannian gradient. Once the Riemannian metric is fixed, the definition of the Riemannian gradient for a smooth scalar field g on a Riemannian manifold $\mathcal{M}$ can be introduced. The Riemannian gradient of g at x on $\mathcal{M}$ is defined as the unique tangent vector belonging to $T_x\mathcal{M}$ and satisfying

$$\langle \mathrm{grad}\ g(x),\ \zeta\rangle_x = \mathrm{D}g(x)[\zeta] \quad \forall\ \zeta \in T_x\mathcal{M}.$$

Let $\hat{f}$ be an extension of $\tilde{f}$ from $\mathrm{St}(n, r)$ to $\mathbb{R}^{n\times r}$, defined as $\hat{f} : V \in \mathbb{R}^{n\times r} \mapsto \mathrm{tr}(CP_{\mathcal{T}}C^T - 2CX_{\mathcal{T}}\hat{C}^T + \hat{C}\hat{P}_{\mathcal{T}}\hat{C}^T) \in \mathbb{R}$. Let grad $f([V])$ denote the Riemannian gradient of the cost function f at $[V]$. Next, we use the Euclidean gradient grad $\hat{f}(V)$ of $\tilde{f}$ to derive grad $f([V])$.

Before going on, we need to introduce the following Sylvester or Lyapunov equations:

$$A^TY_2 + Y_2\hat{A} + \sum_{k=1}^{m} N_k^TY_1\hat{N}_k = 0, \quad \hat{A}^T\hat{Q}_2 + \hat{Q}_2\hat{A} + \sum_{k=1}^{m} \hat{N}_k^T\hat{Q}_1\hat{N}_k = 0, \tag{4.3}$$

$$A^TY_3 + Y_3\hat{A} + \mathcal{H}^{(2)}(X_1 \otimes Y_1)(\hat{\mathcal{H}}^{(2)})^T = 0, \tag{4.4}$$

$$\hat{A}^T\hat{Q}_3 + \hat{Q}_3\hat{A} + \hat{\mathcal{H}}^{(2)}(\hat{P}_1 \otimes \hat{Q}_1)(\hat{\mathcal{H}}^{(2)})^T = 0. \tag{4.5}$$

We also need to show the following relations based on the definitions of truncated Gramians:

$$Y_{\mathcal{T}} = \sum_{i=1}^{3} Y_i, \ \hat{Q}_{\mathcal{T}} = \sum_{i=1}^{3} \hat{Q}_i, \tag{4.6}$$

which will be used to simplify the expression of grad $f([V])$. Specific derivation processes for (4.3)–(4.6) are given in the second part of Appendix B.

Denote by $\mathcal{I}^{(1)} = H(V \otimes V)$ and $\mathcal{P}^{(2)} = \mathcal{H}^{(2)}(V \otimes V)$. Let $\mathcal{K}^{(\mu)}$ and $\mathcal{S}^{(\mu)}$ be the mode-μ matricizations of the tensors $\mathcal{K} := \mathcal{H} \times_1 Y_1^T \times_2 X_1^T \times_3 X_1^T \in \mathbb{R}^{r\times r\times r}$ and $\mathcal{S} := \hat{\mathcal{H}} \times_1 \hat{Q}_1 \times_2 \hat{P}_1 \times_3 \hat{P}_1 \in \mathbb{R}^{r\times r\times r}$, respectively. The following theorem is given.

Theorem 4.1. *Consider the QB system* (2.5) *and its reduced system* (3.1) *constructed by V, where the matrices A and $\hat{A}$ are stable. Then, the horizontal lift of* grad $f([V])$ *at V is computed by*

$$\overline{\text{grad } f}([V]) = \Pi_V(\text{grad } \hat{f}(V)), \tag{4.7}$$

where grad $\hat{f}(V)$ *is given by*

$$\begin{aligned}\text{grad } \hat{f}(V) = 2\Big(&A^TVM_1 + AVM_1^T + C^T\hat{C}\hat{P}_{\mathcal{T}} - C^TCX_{\mathcal{T}} + BB^T(Y_{\mathcal{T}} + Y_3)\\ &+ B\hat{B}^T(\hat{Q}_{\mathcal{T}} + \hat{Q}_3) + \mathcal{I}^{(1)}(\mathcal{K}^{(1)} + \mathcal{S}^{(1)})^T + 2\big(\mathcal{P}^{(2)}(\mathcal{K}^{(2)} + \mathcal{S}^{(2)})^T\big)\\ &+ \sum_{k=1}^{m}\big(N_k^T(VY_1^TN_kX_1 + V\hat{Q}_1\hat{N}_k\hat{P}_1) + N_k(VX_1^TN_k^TY_1 + V\hat{P}_1\hat{N}_k^T\hat{Q}_1)\big)\Big)\end{aligned} \tag{4.8}$$

with $M_1 = Y_1^TX_{\mathcal{T}} + \hat{Q}_1\hat{P}_{\mathcal{T}} + (Y_2 + 2Y_3)^TX_1 + (\hat{Q}_2 + 2\hat{Q}_3)\hat{P}_1$.

Proof. We show the proof in three parts. In the first part, it begins by giving the property of the trace operator. For the matrices M, N, L_1, L_2, U_1, and U_2 with corresponding dimensions, if M and N satisfy

$$ML_1 + L_1N + U_1 = 0, \ L_2M + NL_2 + U_2 = 0, \tag{4.9}$$

then it holds that tr(U_1L_2)=tr(U_2L_1). One can see [29, Lemma A1] for the proof.

For any $\xi \in \mathbb{R}^{n\times r}$, differentiating both sides of (3.8), we find that $\mathrm{D}X_{\mathcal{T}}[\xi]$ satisfies

$$A\mathrm{D}X_{\mathcal{T}}[\xi] + \mathrm{D}X_{\mathcal{T}}[\xi]\hat{A}^T + X_{\mathcal{T}}\xi^TA^TV + X_{\mathcal{T}}V^TA^T\xi + BB^T\xi + D_x = 0, \tag{4.10}$$

where D_x is given by

$$D_x = HE^{\otimes}(\mathrm{D}X_1[\xi], X_1)\hat{H}^T + HE^{\otimes}(X_1\xi^T, X_1V^T)H^TV + H(X_1V^T \otimes X_1V^T)H^T\xi + \sum_{k=1}^{m} \big(N_k\mathrm{D}X_1[\xi]\hat{N}_k^T + N_kX_1\xi^TN_k^TV + N_kX_1V^TN_k^T\xi\big)$$

in which $\mathrm{D}X_1[\xi]$ is obtained by differentiating the first equation in (3.10) and satisfies

$$A\mathrm{D}X_1[\xi] + \mathrm{D}X_1[\xi]\hat{A}^T + X_1\xi^TA^TV + X_1V^TA^T\xi + BB^T\xi = 0. \tag{4.11}$$

Similarly, we can get the equations related to $\mathrm{D}\hat{P}_1[\xi]$ and $\mathrm{D}\hat{P}_{\mathcal{T}}[\xi]$, which satisfy

$$\hat{A}\mathrm{D}\hat{P}_1[\xi] + \mathrm{D}\hat{P}_1[\xi]\hat{A}^T + 2\mathrm{sym}(\xi^TAV\hat{P}_1 + V^TA\xi\hat{P}_1 + \xi^TB\hat{B}^T) = 0, \tag{4.12}$$

$$\hat{A}\mathrm{D}\hat{P}_{\mathcal{T}}[\xi] + \mathrm{D}\hat{P}_{\mathcal{T}}[\xi]\hat{A}^T + 2\mathrm{sym}(\xi^TAV\hat{P}_{\mathcal{T}} + V^TA\xi\hat{P}_{\mathcal{T}} + \xi^TB\hat{B}^T) + D_p = 0, \tag{4.13}$$

where

$$\begin{aligned} D_p = {} & \hat{H}E^{\otimes}(\mathrm{D}\hat{P}_1[\xi], \hat{P}_1)\hat{H}^T + 2\mathrm{sym}(\xi^TH(V\hat{P}_1 \otimes V\hat{P}_1)\hat{H}^T) \\ & + V^THE^{\otimes}(\xi\hat{P}_1, V\hat{P}_1)\hat{H}^T + \hat{H}E^{\otimes}(\hat{P}_1\xi^T, \hat{P}_1V^T)H^TV \\ & + \sum_{k=1}^{m} \big(\hat{N}_k\mathrm{D}\hat{P}_1[\xi]\hat{N}_k^T + 2\mathrm{sym}(\xi^TN_kV\hat{P}_1\hat{N}_k^T + N_k\xi\hat{P}_1\hat{N}_k^T)\big). \end{aligned}$$

We compute the directional derivative $\mathrm{D}\hat{f}(V)[\xi]$ by

$$\mathrm{D}\hat{f}(V)[\xi] = \mathrm{tr}\big(-2\hat{C}^TC\mathrm{D}X_{\mathcal{T}}[\xi] - 2\xi^TC^TCX_{\mathcal{T}} + 2\xi^TC^T\hat{C}\hat{P}_{\mathcal{T}} + \hat{C}^T\hat{C}\mathrm{D}\hat{P}_{\mathcal{T}}[\xi]\big). \tag{4.14}$$

Note that the expressions of the first and last terms in the above formula should be investigated in detail. Thus, in the second part, our task is to drive these expressions.

First, let us study the term $\mathrm{tr}(-2\hat{C}^TC\mathrm{D}X_{\mathcal{T}}[\xi])$. Applying the property (4.9) to the first equation in (3.14) and (4.10) leads to

$$\begin{aligned} \mathrm{tr}\left(-\hat{C}^TC\mathrm{D}X_{\mathcal{T}}[\xi]\right) &= \mathrm{tr}\Big((X_{\mathcal{T}}\xi^TA^TV + X_{\mathcal{T}}V^TA^T\xi + BB^T\xi + D_x)Y_1^T\Big) \\ &= \mathrm{tr}\Big(A^TVY_1^TX_{\mathcal{T}}\xi^T + AVX_{\mathcal{T}}^TY_1\xi^T + BB^TY_1\xi^T \\ &\quad + HE^{\otimes}(\mathrm{D}X_1[\xi], X_1)\hat{H}^TY_1^T + HE^{\otimes}(X_1\xi^T, X_1V^T)H^TVY_1^T \\ &\quad + H(X_1V^T \otimes X_1V^T)H^T\xi Y_1^T + \sum_{k=1}^{m} \big(N_k^TVY_1^TN_kX_1\xi^T \\ &\quad + N_kVX_1^TN_k^TY_1\xi^T + N_k\mathrm{D}X_1[\xi]\hat{N}_k^TY_1^T\big)\Big). \end{aligned} \tag{4.15}$$

Now, we aim to study some terms appearing in (4.15). Using Lemma 2.3, due to the symmetry of $\mathcal{H}$ and $\hat{\mathcal{H}}$, it holds that

$$\mathrm{tr}\big(HE^{\otimes}(\mathrm{D}X_1[\xi], X_1)\hat{H}^TY_1^T\big) = 2\mathrm{tr}\big(H\left(\mathrm{D}X_1[\xi] \otimes X_1\right)\hat{H}^TY_1^T\big).$$

Since $H(\mathrm{D}X_1[\xi] \otimes X_1)$ is a matrix of size $n \times r^2$, it can be viewed as the mode-1 matricization of the tensor $\mathcal{H} \times_2 X_1^T \times_3 (\mathrm{D}X_1[\xi])^T \in \mathbb{R}^{n\times r\times r}$. Obviously, its mode-3 matricization is $(\mathrm{D}X_1[\xi])^T\mathcal{H}^{(3)}(X_1 \otimes I_n)$. Thus, based on Lemma 2.2, one gets

$$
\begin{aligned}
\operatorname{tr}\Big(HE^{\otimes}(\mathrm{D}X_1[\xi], X_1)\hat{H}^T Y_1^T\Big) &= 2\operatorname{tr}\Big((\mathrm{D}X_1[\xi])^T \mathcal{H}^{(3)}(X_1 \otimes I_n)(I_r \otimes Y_1)(\hat{\mathcal{H}}^{(3)})^T\Big) \\
&= 2\operatorname{tr}\Big((\mathrm{D}X_1[\xi])^T \mathcal{H}^{(2)}(X_1 \otimes Y_1)(\hat{\mathcal{H}}^{(2)})^T\Big) \\
&= 2\operatorname{tr}\Big((X_1\xi^T A^T V + X_1 V^T A^T \xi + BB^T\xi)^T Y_3\Big) \\
&= 2\operatorname{tr}\Big((A^T V Y_3^T X_1 + AVX_1^T Y_3 + BB^T Y_3)\xi^T\Big),
\end{aligned} \tag{4.16}
$$

where the last second equality is obtained by applying the property of the trace operator (4.9) to (4.4) and (4.11).

For the $n\times n\times n$ tensor $\mathcal{H}\times_2 VX_1^T \times_3 \xi X_1^T$, its mode-1 and mode-3 matricizations are $H(X_1\xi^T \otimes X_1V^T)$ and $\xi X_1^T\mathcal{H}^{(3)}(X_1V^T \otimes I_n)$, respectively. Making full use of Lemmas 2.2 and 2.3, one has

$$
\begin{aligned}
\operatorname{tr}\Big(HE^{\otimes}(X_1\xi^T, X_1V^T)H^T V Y_1^T\Big) &= 2\operatorname{tr}\Big(H(X_1\xi^T \otimes X_1V^T)H^T V Y_1^T\Big) \\
&= 2\operatorname{tr}\Big(\xi X_1^T \mathcal{H}^{(3)}(X_1V^T \otimes I_n)\big(\mathcal{H}^{(3)}(I_n \otimes VY_1^T)\big)^T\Big) \\
&= 2\operatorname{tr}\Big(\mathcal{H}^{(2)}(VX_1^T \otimes VY_1^T)(\mathcal{H}^{(2)})^T X_1\xi^T\Big) \\
&= 2\operatorname{tr}\Big(\mathcal{P}^{(2)}(\mathcal{K}^{(2)})^T\xi^T\Big).
\end{aligned} \tag{4.17}
$$

Consider the first equation in (4.3) and the equation (4.11). Utilizing the trace property (4.9) again, we rewrite the last term in (4.15) as

$$
\begin{aligned}
\operatorname{tr}\left(\sum_{k=1}^{m} N_k \mathrm{D}X_1[\xi]\hat{N}_k^T Y_1^T\right) &= \operatorname{tr}\left(\sum_{k=1}^{m}(\mathrm{D}X_1[\xi])^T N_k^T Y_1 \hat{N}_k\right) \\
&= \operatorname{tr}\Big((X_1\xi^T A^T V + X_1 V^T A^T\xi + BB^T\xi)^T Y_2\Big) \\
&= \operatorname{tr}\Big(\big(A^T V Y_2^T X_1 + AVX_1^T Y_2 + BB^T Y_2\big)\,\xi^T\Big).
\end{aligned} \tag{4.18}
$$

Thus, based on these results derived in (4.16)–(4.18) and the relation in (4.6), the formula (4.15) is finally rewritten as

$$
\begin{aligned}
\operatorname{tr}\Big(-\hat{C}^T C \mathrm{D}X_{\mathcal{T}}[\xi]\Big) = \operatorname{tr}\Big(\Big(&A^T V Y_1^T X_{\mathcal{T}} + AVX_{\mathcal{T}}^T Y_1 + AVX_1^T(Y_2 + 2Y_3) + A^T V(Y_2 \\
&+ 2Y_3)^T X_1 + BB^T(Y_{\mathcal{T}} + Y_3) + \mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T + 2\mathcal{P}^{(2)}(\mathcal{K}^{(2)})^T \\
&+ \sum_{k=1}^{m}\big(N_k^T V Y_1^T N_k X_1 + N_k V X_1^T N_k^T Y_1\big)\Big)\xi^T\Big).
\end{aligned} \tag{4.19}
$$

Reusing Lemmas 2.2 and 2.3 and the trace property (4.9), the specific expression of the last term in (4.14) can be derived in a similar way, which takes the form

$$
\begin{aligned}
\operatorname{tr}\Big(\hat{C}^T\hat{C}\mathrm{D}\hat{P}_{\mathcal{T}}[\xi]\Big) = 2\operatorname{tr}\Big(\Big(&A^T V\hat{Q}_1\hat{P}_{\mathcal{T}} + AV\hat{P}_{\mathcal{T}}\hat{Q}_1 + A^T V(\hat{Q}_2 + 2\hat{Q}_3)\hat{P}_1 + AV\hat{P}_1(\hat{Q}_2 \\
&+ 2\hat{Q}_3) + B\hat{B}^T(\hat{Q}_{\mathcal{T}} + \hat{Q}_3) + \mathcal{I}^{(1)}(\mathcal{S}^{(1)})^T + 2\mathcal{P}^{(2)}(\mathcal{S}^{(2)})^T \\
&+ \sum_{k=1}^{m}\big(N_k V\hat{P}_1\hat{N}_k^T\hat{Q}_1 + N_k^T V\hat{Q}_1\hat{N}_k\hat{P}_1\big)\Big)\xi^T\Big).
\end{aligned} \tag{4.20}
$$

In the following part, we are ready to derive (4.7). Substituting (4.19) and (4.20) into (4.14) and using $\mathrm{D}\hat{f}(V)[\xi] = \mathrm{tr}(\xi^T \mathrm{grad}\ \hat{f}(V))$, it is easy to verify that the expression (4.8) for grad $\hat{f}(V)$ holds. Moreover, the Riemannian gradient in (4.7) can be obtained following from [20, equation (3.39)]. This concludes the proof. $\square$

Remark 2. Obviously, the minimization problem (3.17) implies the corresponding minimization problems for bilinear and linear systems when the corresponding terms are removed. Then, the Riemannian gradient grad $f(V)$ in (4.7) gives the gradients of the minimization problems for bilinear and linear systems. Specifically, by removing the term related to H, one can obtain grad $f(V)$ for the bilinear case, where

$$\begin{aligned}\mathrm{grad}\ \hat{f}(V) = 2\Big(&A^T V M_1 + A V M_1^T + C^T \hat{C} \hat{P}_{\mathcal{T}} - C^T C X_{\mathcal{T}} + B B^T Y_{\mathcal{T}} + B \hat{B}^T \hat{Q}_{\mathcal{T}} \\ &+ \sum_{k=1}^{m} \left(N_k^T (V Y_1^T N_k X_1 + V \hat{Q}_1 \hat{N}_k \hat{P}_1) + N_k (V X_1^T N_k^T Y_1 + V \hat{P}_1 \hat{N}_k^T \hat{Q}_1) \right)\Big)\end{aligned}$$

with $M_1 = Y_1^T X_{\mathcal{T}} + \hat{Q}_1 \hat{P}_{\mathcal{T}} + Y_2^T X_1 + \hat{Q}_2 \hat{P}_1$. Let $H = 0$ and $N_k = 0$, and then (4.7) gives the expression of grad $f(V)$ for linear case, which has also been derived in [29].

4.3. Riemannian MPRP conjugate gradient scheme. In [41], the MPRP conjugate gradient method in Euclidean spaces is presented based on the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method. Motivated by it and using the Riemannian BFGS (RBFGS) method [37], we next design a new Riemannian MPRP conjugate gradient scheme for solving the minimization problem (3.17).

Retractions are important concepts in Riemmannian geometry. On $\mathrm{Gr}(n, r)$, a frequently used retraction is

$$\mathcal{R}_{[V]}(\eta) = [q(V + \overline{\eta})], \tag{4.21}$$

where $\overline{\eta}$ is the horizontal lift of $\eta \in T_{[V]}\mathrm{Gr}(n, r)$ and $q(Z)$ denotes the Q factor of the known thin QR decomposition for $Z \in \mathbb{R}^{n \times r}$, i.e., $Z = QR$ and $q(Z) = Q$ with $Q \in \mathrm{St}(n, r)$ [20, 23]. Let $[V_i]$ be a current point on $\mathrm{Gr}(n, r)$. The above retraction is utilized to go back from the updates in $T_{[V_i]}\mathrm{Gr}(n, r)$ to $\mathrm{Gr}(n, r)$, which yields the next iterate

$$[V_{i+1}] = \mathcal{R}_{[V_i]}(\alpha_i \eta_i),$$

where η_i is the search direction at $[V_i]$ and $\alpha_i \in \mathbb{R}$ is the step-size produced by the line-search procedure.

The choice of an efficient search direction is important in the design of high-performance numerical algorithms. According to the RBFGS method, the search direction is given by $\eta_i = -\mathfrak{B}_i^{-1} \mathrm{grad}\ f([V_i])$, where $\mathfrak{B}_i^{-1}$ is a linear operator on $T_{[V_i]}\mathrm{Gr}(n, r)$ and is used to approximate the action of the inverse Hessian (see [37]) along grad $f([V_i])$. The inverse Hessian approximation update formula of $\mathfrak{B}_i^{-1}$ is written as

$$\mathfrak{H}_i = \tilde{\mathfrak{H}}_{i-1} - \frac{\tilde{\mathfrak{H}}_{i-1} y_{i-1} (\tilde{\mathfrak{H}}_{i-1}^* y_{i-1})^\flat}{(\tilde{\mathfrak{H}}_{i-1}^* y_{i-1})^\flat y_{i-1}} + \frac{s_{i-1} s_{i-1}^\flat}{s_{i-1}^\flat y_{i-1}},\ \tilde{\mathfrak{H}}_{i-1} = \mathcal{T}_{S_{\alpha_{i-1}\eta_{i-1}}} \circ \mathfrak{H}_{i-1} \circ \mathcal{T}_{S_{\alpha_{i-1}\eta_{i-1}}}^{-1}, \tag{4.22}$$

where $y_{i-1} = \rho_{i-1}^{-1} \mathrm{grad}\ f([V_i]) - \mathcal{T}_{S_{\alpha_{i-1}\eta_{i-1}}} \mathrm{grad}\ f([V_{i-1}])$ and $s_{i-1} = \mathcal{T}_{S_{\alpha_{i-1}\eta_i}} \alpha_{i-1} \eta_{i-1}$. Here, ρ_{i-1} is an arbitrary number satisfying that $|\rho_{i-1} - 1| \le L_\rho \|\alpha_{i-1} \eta_{i-1}\|_{[V_{i-1}]}$ and

$|\rho_{i-1}^{-1} - 1| \leq L_\rho \|\alpha_{i-1}\eta_{i-1}\|_{[V_{i-1}]}$ for $L_\rho > 0$. It can also be set to 1 for all i. Thereby, the direction can be rewritten as $\eta_i = -\mathfrak{H}_i \mathrm{grad}\ f([V_i])$. Taking $\mathfrak{H}_{i-1}$ in (4.22) as the identity mapping, we find that η_i can be formalized as

$$\eta_i = -\mathrm{grad}\ f([V_i]) + \beta_i \mathcal{T}_{S_{\alpha_{i-1}\eta_{i-1}}} \eta_{i-1} - \theta_i y_{i-1}, \tag{4.23}$$

where β_i and θ_i are parameters to be determined.

Based on (4.23) and following the idea in [41], we next design a new Riemannian MPRP conjugate gradient scheme on the Grassmann manifold and then construct a new update direction, which can generalize the MPRP conjugate gradient method to the Riemannian setting. However, the direct generalization will destroy some good properties of the MPRP conjugate gradient method, e.g., its descent direction and global convergence. Besides, as far as we know, it is difficult to obtain the expression of $\mathcal{T}_S$ in (4.23). To tackle these issues, we next take the following measures to (4.23). First, we replace the isometric vector transport $\mathcal{T}_S$ by $\mathcal{T}$ that satisfies the Ring–Wirth nonexpansive condition. Second, we modify y_i using the differentiated retraction operator. Finally, we concentrate on designing the parameter θ_i.

From the definition of the vector transport $\mathcal{T}^{\mathcal{R}}$ by differentiated retraction on $\mathrm{Gr}(n,r)$, one gets $\overline{\mathcal{T}^{\mathcal{R}}_{\eta_i}\xi_i} = \overline{\mathrm{D}\mathcal{R}_{[V_i]}(\eta_i)[\xi_i]}$ for $\eta_i, \xi_i \in T_{[V_i]}\mathrm{Gr}(n,r)$. It is specified as

$$\overline{\mathcal{T}^{\mathcal{R}}_{\eta_i}\xi_i} = \big(I - q(V_i + \overline{\eta}_i)q^T(V_i + \overline{\eta}_i)\big)\overline{\xi}_i\big(q^T(V_i + \overline{\eta}_i)(V_i + \overline{\eta}_i)\big)^{-1}, \tag{4.24}$$

where $\overline{\eta}_i$ and $\overline{\xi}_i$ are the horizontal lifts of η_i and ξ_i, respectively. The derivation of (4.24) is similar to [20, p. 173], and one can refer to [42] for more details.

However, $\mathcal{T}^{\mathcal{R}}$ in (4.24) does not always satisfy the Ring–Wirth nonexpansive condition. Note that the scaled vector transport $\mathcal{T}^S$ in [43] associated with $\mathcal{T}^{\mathcal{R}}$ is already defined. The horizontal lift of $\mathcal{T}^S$ on $\mathrm{Gr}(n,r)$ is expressed as

$$\overline{\mathcal{T}^S_{\eta_i}\xi_i} = \frac{\|\xi_i\|_{[V_i]}}{\|\mathcal{T}^{\mathcal{R}}_{\eta_i}\xi_i\|_{\mathcal{R}_{[V_i]}(\eta_i)}} \overline{\mathcal{T}^{\mathcal{R}}_{\eta_i}\xi_i}.$$

According to (4.2), we have $\|\mathcal{T}^S_{\eta_i}\xi_i\|_{\mathcal{R}_{[V_i]}(\eta_i)} = \|\xi_i\|_{[V_i]}$. Further, we give $\mathcal{T}$ by

$$\overline{\mathcal{T}_{\eta_i}\xi_i} = \begin{cases} \overline{\mathcal{T}^{\mathcal{R}}_{\eta_i}\xi_i} & \text{if } \|\mathcal{T}^{\mathcal{R}}_{\eta_i}\xi_i\|_{\mathcal{R}_{[V_i]}(\eta_i)} \leq \|\xi_i\|_{[V_i]}, \\ \overline{\mathcal{T}^S_{\eta_i}\xi_i} & \text{ortherwise.} \end{cases} \tag{4.25}$$

Clearly, it satisfies the Ring–Wirth nonexpansive condition. Thus, we adopt $\mathcal{T}$ in (4.25) to replace the isometric vector transport $\mathcal{T}_S$ in (4.23).

Further, we reformulate the horizontal lift of y_{i-1} as

$$\overline{y}_{i-1} = \overline{\mathrm{grad}\ f}([V_i]) - \overline{\mathcal{T}^{\mathcal{R}}_{\alpha_{i-1}\eta_{i-1}}\mathrm{grad}\ f([V_{i-1}])}.$$

To guarantee the global convergence, the parameters are designed as

$$\beta_i^{\mathrm{PRP}} = \frac{\langle \mathrm{grad}\ f([V_i]), y_{i-1}\rangle_{[V_i]}}{\|\mathrm{grad}\ f([V_{i-1}])\|^2_{[V_{i-1}]}},\ \theta_i = \frac{\langle \mathrm{grad}\ f([V_i]), \mathcal{T}_{\alpha_{i-1}\eta_{i-1}}\eta_{i-1}\rangle_{[V_i]}}{\|\mathrm{grad}\ f([V_{i-1}])\|^2_{[V_{i-1}]}}. \tag{4.26}$$

Based on the above formulae, we rewrite the update direction η_i in (4.23). Its horizontal lift is then constructed by

$$\bar{\eta}_i = \begin{cases} -\overline{\mathrm{grad}\ f}([V_0]) & \text{if } i = 0, \\ -\overline{\mathrm{grad}\ f}([V_i]) + \beta_i^{\mathrm{PRP}}\overline{\mathcal{T}_{\alpha_{i-1}\eta_{i-1}}\eta_{i-1}} - \theta_i \bar{y}_{i-1} & \text{if } i \geq 1. \end{cases} \tag{4.27}$$

In fact, various approaches are possible to determine the step-size α_i in the iterate formula. In general, the chosen α_i has to fulfill a certain quality requirement. As we know, there are two frequently used rules to compute α_i. One is the Armijo-type rule. Like in [41], one can choose $\alpha_i = \max\{\rho^l, l = 0, 1, \ldots\}$ for $0 < \rho < 1$ and $\varrho > 0$ such that

$$f(\mathcal{R}_{[V_i]}(\alpha_i \eta_i)) \leq f([V_i]) - \varrho \alpha_i^2 \|\eta_i\|^2_{[V_i]}. \tag{4.28}$$

The other is the Wolfe condition [42], namely, α_i should satisfy

$$f(\mathcal{R}_{[V_i]}(\alpha_i \eta_i)) \leq f([V_i]) + \delta_1 \alpha_i \langle \text{grad } f([V_i]), \eta_i \rangle_{[V_i]}, \tag{4.29a}$$

$$\frac{d}{dt} f(\mathcal{R}_{[V_i]}(t\eta_i))|_{t=\alpha_i} \geq \delta_2 \frac{d}{dt} f(\mathcal{R}_{[V_i]}(t\eta_i))|_{t=0}, \tag{4.29b}$$

where $0 < \delta_1 < \delta_2 < 1$. Based on the definition of $\mathcal{T}^{\mathcal{R}}$, the inequality (4.29b) can be rewritten as

$$\langle \text{grad} f([V_{i+1}]), \mathcal{T}^{\mathcal{R}}_{\alpha_i \eta_i} \eta_i \rangle_{[V_{i+1}]} \geq \delta_2 \langle \text{grad} f([V_i]), \eta_i \rangle_{[V_i]}.$$

4.4. The implementation algorithm. Since all the essential components of the Riemannian MPRP conjugate gradient method are provided, we may give a summary of the whole MOR process in Algorithm 1. In brief, the Riemannian MPRP conjugate gradient method is now simply denoted the RMPRP method.

Algorithm 1: RMPRP for QB systems

Input: The QB system (2.5) with the stable matrix A, the reduced order r, the tolerance $\varepsilon > 0$, and the initial matrix V_0 with $V_0^T V_0 = I_r$.

Output: The reduced QB system (3.1) with the stable matrix $\hat{A}$.

1. Compute $\overline{\text{grad } f}([V_0])$ according to (4.7) and (4.8). Set $i = 0$.
2. While $\|\text{grad } f([V_i])\|_{[V_i]} \geq \varepsilon$ do
3. Compute the search direction $\bar{\eta}_i$ by (4.27) where $\overline{\mathcal{T}_{\alpha_{i-1}\eta_{i-1}} \eta_{i-1}}$ and $\overline{\mathcal{T}^{\mathcal{R}}_{\alpha_{i-1}\eta_{i-1}} \text{grad } f([V_{i-1}])}$ are obtained via (4.24).
4. Obtain the step-size α_i and compute $[V_{i+1}] = \mathcal{R}_{[V_i]}(\alpha_i \eta_i)$ using (4.21).
5. Set $i = i + 1$.
6. End while
7. Let $[V^{opt}] = \lim_{i \to +\infty} [V_i]$. Obtain $\hat{A} = (V^{opt})^T A V^{opt}$, $\hat{H} = (V^{opt})^T H (V^{opt} \otimes V^{opt})$, $\hat{N}_k = (V^{opt})^T N_k V^{opt}$, $\hat{B} = (V^{opt})^T B$ and $\hat{C} = C V^{opt}$.

For Algorithm 1, some properties can be concluded as follows:

- According to (4.27), if grad $f([V_i]) \neq 0$, then one gets

$$\begin{aligned} \langle \eta_i, \text{grad } f([V_i]) \rangle_{[V_i]} &= -\|\text{grad } f([V_i])\|^2_{[V_i]} + \beta_i^{\text{PRP}} \langle \mathcal{T}_{\alpha_{i-1}\eta_{i-1}} \eta_{i-1}, \\ &\qquad \text{grad } f([V_i]) \rangle_{[V_i]} - \theta_i \langle y_{i-1}, \text{grad } f([V_i]) \rangle_{[V_i]} \\ &= -\|\text{grad } f([V_i])\|^2_{[V_i]} < 0. \end{aligned} \tag{4.30}$$

  Thus, the direction η_i for any i always provides a descent direction of f, independent of the line-search process and the convexity of f.
- From the Cauchy–Schwarz inequality, one further has

$$\|\text{grad } f([V_i])\|^2_{[V_i]} = |\langle \eta_i, \text{ grad } f([V_i]) \rangle_{[V_i]}| \leq \|\text{grad } f([V_i])\|_{[V_i]} \|\eta_i\|_{[V_i]}.$$

  This implies that $\|\text{grad } f([V_i])\|_{[V_i]} \leq \|\eta_i\|_{[V_i]}$ for grad $f([V_i]) \neq 0$.

- Exact line-search minimization for any i results in

$$0=\frac{d}{dt}f(\mathcal{R}_{[V_i]}(t\eta_i))|_{t=\alpha_i}=\mathrm{D}f([V_{i+1}])\left[\frac{d}{dt}\mathcal{R}_{[V_i]}(t\eta_i)|_{t=\alpha_i}\right]=\mathrm{D}f([V_{i+1}])[\mathcal{T}^{\mathcal{R}}_{\alpha_i\eta_i}\eta_i].$$

It means

$$\langle \text{grad } f([V_{i+1}]), \mathcal{T}^{\mathcal{R}}_{\alpha_i\eta_i}\eta_i\rangle_{[V_{i+1}]}=0,\ \langle \text{grad } f([V_{i+1}]), \mathcal{T}_{\alpha_i\eta_i}\eta_i\rangle_{[V_{i+1}]}=0.$$

Thus, we observe that $\theta_i=0$ for any i if the exact line-search is performed.

5. Global convergence analysis of the proposed algorithm. To show the global convergence of Algorithm 1, i.e., the generated sequence converges toward a local minimum, we first start this task by giving the definition of the radially Lipschitz continuous function.

DEFINITION 5.1 (the radially L-C[1] function [37]). *For $g: x\in\mathcal{M}\mapsto g(x)\in\mathbb{R}$ with retraction $\mathcal{R}$, the function $\breve{g}=g\circ\mathcal{R}$ is radially Lipschitz continuous if there exists a constant $l_1>0$ such that for all $x\in\mathcal{U}\subseteq\mathcal{M}$ and all $\eta\in T_x\mathcal{M}$, it holds that*

$$\left|\frac{d}{dt}\breve{g}(t\eta)|_{t=\alpha}-\frac{d}{dt}\breve{g}(t\eta)|_{t=0}\right|\le l_1\alpha\|\eta\|_x^2,$$

where α and η satisfy $\mathcal{R}_x(\alpha\eta)\in\mathcal{U}$.

Definition 5.2 gives the Riemannian version of Lipschitz continuous differentiable functions [37].

DEFINITION 5.2. *Let $\mathcal{T}^{\mathcal{R}}$ be a vector transport associated with the retraction $\mathcal{R}$. The cost function g is said to be Lipschitz continuously differentiable with respect to $\mathcal{T}^{\mathcal{R}}$ on $\mathcal{U}\subseteq\mathcal{M}$ if there exists a constant $l_2>0$ such that*

$$\|\mathcal{T}^{\mathcal{R}}_{\alpha\eta}\text{grad } g(x)-\text{grad } g(\mathcal{R}_x(\alpha\eta))\|_{\mathcal{R}_x(\alpha\eta)}\le l_2\alpha\|\eta\|_x \tag{5.1}$$

for all $x\in\mathcal{U}$ and η satisfying $\mathcal{R}_x(\alpha\eta)\in\mathcal{U}$.

In order to establish the global convergence of Algorithm 1, the following results are needed. First, Lemma 5.3 shows that the defined level set Ω is compact due to the compactness of $\mathrm{Gr}(n,r)$.

LEMMA 5.3. *For any given initial point $[V_0]$, the level set $\Omega=\{[V_i]\in\mathrm{Gr}(n,r)|\ f([V_i])\le f([V_0])\}$ is compact and f is bounded on Ω.*

As shown in [37], since f is twice continuously differentiable on the compact manifold $\mathrm{Gr}(n,r)$, we can find that the cost function f satisfies Definitions 5.1 and 5.2. The following lemmas are then obtained.

LEMMA 5.4. *The function $\breve{f}=f\circ\mathcal{R}$ is radially Lipschitz continuous on Ω.*

LEMMA 5.5. *The cost function f is Lipschitz continuously differentiable with respect to $\mathcal{T}^{\mathcal{R}}$ on Ω.*

Obviously, these lemmas imply that there exists $\gamma_1>0$ such that $\|\text{grad } f([V])\|_{[V]}\le\gamma_1$ for all $[V]\in\Omega$.

LEMMA 5.6. *Considering Algorithm 1 with the Armijo-type rule (4.28), it holds that $\liminf_{i\to+\infty}\alpha_i\|\eta_i\|_{[V_i]}=0$.*

Proof. From (4.28) and (4.30), one finds that the sequence $\{f([V_i])\}$ is decreasing. Then, its boundedness on Ω is obtained. Subsequently, one has $\sum_{i=0}^{+\infty}\alpha_i\|\eta_i\|_{[V_i]}<+\infty$ by (4.28), which implies that this lemma holds. ∎

Lemmas 5.3 and 5.5 are used to obtain Lemma 5.7.

LEMMA 5.7. *Consider Algorithm* 1 *with the Armijo-type rule* (4.28). *If there exists a constant* $\varepsilon_1 > 0$ *satisfying* $\|\text{grad } f([V_i])\|_{[V_i]} \geq \varepsilon_1$ *for any* i, *then there exists a constant* $\gamma_2 > 0$ *such that* $\|\eta_i\|_{[V_i]} \leq \gamma_2$.

Proof. For $i = 0$, let $\gamma_2 = \gamma_1$, and then it is obvious that $\|\eta_i\|_{[V_i]} \leq \gamma_2$. Lemma 5.6 shows that there exists $i_0 > 0$ such that $2\gamma_1 l_2 \|\alpha_i \eta_i\|_{[V_i]}/\varepsilon_1^2 \leq \gamma_3$ for $0 < \gamma_3 < 1$ and all $i > i_0$. Notice that f satisfies Lemma 5.5 and $\mathcal{T}$ satisfies the Ring–Wirth nonexpansive condition. By utilizing the Cauchy–Schwarz inequality, one gets

$$\begin{aligned}
\|\eta_i\|_{[V_i]} &= \| - \text{grad } f([V_i]) + \beta_i^{\text{PRP}} \mathcal{T}_{\alpha_{i-1}\eta_{i-1}}\eta_{i-1} - \theta_i y_{i-1}\|_{[V_i]} \\
&\leq \|\text{grad } f([V_i])\|_{[V_i]} + 2\frac{\|\text{grad } f([V_i])\|_{[V_i]}\|y_{i-1}\|_{[V_i]}\|\mathcal{T}_{\alpha_{i-1}\eta_{i-1}}\eta_{i-1}\|_{[V_i]}}{\|\text{grad } f([V_{i-1}])\|^2_{[V_{i-1}]}} \\
&\leq \|\text{grad } f([V_i])\|_{[V_i]} + 2\frac{l_2\alpha_{i-1}\|\text{grad } f([V_i])\|_{[V_i]}\|\eta_{i-1}\|_{[V_{i-1}]}}{\|\text{grad } f([V_{i-1}])\|^2_{[V_{i-1}]}}\|\eta_{i-1}\|_{[V_{i-1}]} \\
&\leq \gamma_1 + \frac{2l_2\gamma_1\|\alpha_{i-1}\eta_{i-1}\|_{[V_{i-1}]}}{\varepsilon_1^2}\|\eta_{i-1}\|_{[V_{i-1}]} \\
&\leq \gamma_1 + \gamma_3\|\eta_{i-1}\|_{[V_{i-1}]} \leq \gamma_1 \sum_{j=0}^{i-i_0-1}\gamma_3^j + \gamma_3^{i-i_0}\|\eta_{i_0}\|_{[V_{i_0}]} \leq \frac{\gamma_1}{1-\gamma_3} + \|\eta_{i_0}\|_{[V_{i_0}]}.
\end{aligned}$$

Let $\gamma_2 = \max\{\frac{\gamma_1}{1-\gamma_3} + \|\eta_{i_0}\|_{[V_{i_0}]}, \|\eta_1\|_{[V_1]}, \ldots, \|\eta_{i_0}\|_{[V_{i_0}]}\}$, and this lemma is then proved. ∎

The global convergence of Algorithm 1 with the Armijo-type rule (4.28) is shown in Theorem 5.8, which can be regarded as a generalization of [41, Theorem 3.2].

THEOREM 5.8. *Let* $\{[V_i]\}$ *denote the sequence generated by Algorithm* 1 *with the Armijo-type rule* (4.28). *Then it holds that*

$$\liminf_{i\to+\infty} \|\text{grad } f([V_i])\|_{[V_i]} = 0. \tag{5.2}$$

Proof. A contradiction is used to prove this theorem. By the contrary, then there exists a constant $\varepsilon_2 > 0$ such that $\|\text{grad } f([V_i])\|_{[V_i]} \geq \varepsilon_2$ for any i. Due to Lemma 5.6, it has $\liminf_{i\to+\infty} \alpha_i\|\eta_i\|_{[V_i]} = 0$. If $\liminf_{i\to+\infty} \alpha_i > 0$, one gets $\liminf_{i\to+\infty} \|\text{grad } f([V_i])\|_{[V_i]} = 0$ since $\|\text{grad } f([V_i])\| \leq \|\eta_i\|_{[V_i]}$. This contradicts $\|\text{grad } f([V_i])\|_{[V_i]} \geq \varepsilon_2$.

When $\liminf_{i\to+\infty} \alpha_i = 0$, there exists a positive integer set Λ such that $\alpha_i \to 0$ for $i \in \Lambda \to +\infty$. From (4.28), one has

$$f(\mathcal{R}_{[V_i]}(\rho^{-1}\alpha_i\eta_i)) > f([V_i]) - \varrho\rho^{-2}\alpha_i^2\|\eta_i\|^2_{[V_i]}. \tag{5.3}$$

It follows

$$\breve{f}(\rho^{-1}\alpha_i\eta_i) - \breve{f}(0_{[V_i]}) = f(\mathcal{R}_{[V_i]}(\rho^{-1}\alpha_i\eta_i)) - f([V_i]) > -\varrho\rho^{-2}\alpha_i^2\|\eta_i\|^2_{[V_i]}.$$

Utilizing the definition of $\mathcal{T}^{\mathcal{R}}$ and Lemma 5.4, one obtains

$$\begin{aligned}
\breve{f}(\rho^{-1}\alpha_i\eta_i) - \breve{f}(0_{[V_i]}) &= \frac{d}{dt}\breve{f}(t\eta_i)|_{t=\rho^{-1}\alpha_i\sigma_i} \cdot \rho^{-1}\alpha_i \\
&= \left(\frac{d}{dt}\breve{f}(t\eta_i)|_{t=\rho^{-1}\alpha_i\sigma_i} - \frac{d}{dt}\breve{f}(t\eta_i)|_{t=0}\right) \cdot \rho^{-1}\alpha_i \\
&\quad + \frac{d}{dt}\breve{f}(t\eta_i)|_{t=0} \cdot \rho^{-1}\alpha_i \\
&\leq l_1\rho^{-2}\sigma_i\alpha_i^2\|\eta_i\|_{[V_i]}^2 + \rho^{-1}\alpha_i\langle\eta_i, \mathrm{grad}\ f([V_i])\rangle_{[V_i]},
\end{aligned} \tag{5.4}$$

where $\sigma_i \in (0,1)$. Using (5.3) and (5.4) yields

$$-\varrho\rho^{-2}\alpha_i^2\|\eta_i\|_{[V_i]}^2 < l_1\rho^{-2}\sigma_i\alpha_i^2\|\eta_i\|_{[V_i]}^2 + \rho^{-1}\alpha_i\langle\eta_i, \mathrm{grad}\ f([V_i])\rangle_{[V_i]}.$$

We then have $\|\mathrm{grad}\ f([V_i])\|_{[V_i]}^2 < (l_1\sigma_i + \varrho)\rho^{-1}\alpha_i\|\eta_i\|_{[V_i]}^2$. Since $\|\eta_i\|_{[V_i]}$ is bounded, it holds that $\|\mathrm{grad}\ f([V_i])\|_{[V_i]} \to 0$ for $i \in \Lambda \to +\infty$. This contradicts $\|\mathrm{grad}\ f([V_i])\|_{[V_i]} \geq \varepsilon_2$. Thus, the proof of this theorem is completed. ∎

Based on Lemmas 5.4 and 5.5, we can also establish the global convergence of Algorithm 1 with the Wolfe step-size control (4.29).

THEOREM 5.9. *Let $\{[V_i]\}$ be the sequence of iterates generated by Algorithm* 1 *with the Wolfe conditions* (4.29). *If $2l_2\alpha_i\|\eta_i\|_{[V_i]} \leq \|\mathrm{grad}\ f([V_{i+1}])\|_{[V_{i+1}]}$ holds for any i, then the iterate sequence in Algorithm* 1 *is globally convergent, i.e.,* (5.2) *holds.*

Proof. The detailed derivation is shown in Appendix C. ∎

Remark 3. TQB-IRKA with an oblique projection proposed in [19] performs well in simulating QB systems. The Riemannian Fletcher–Reeves-type conjugate gradient method in [43] provides a global convergence analysis with the strong Wolfe condition (see [43, p. 1015]). In this paper, Algorithm 1 is constructed by orthogonal projection, which has attractive properties that the search direction is always descent from any initial orthogonal point and that it generates globally convergent sequences only requiring the Armijo-type rule or the Wolfe condition. Besides, the convergence analysis given in this section can be straightforwardly generalized to any smooth cost function on a compact manifold, such as the Grassmann manifold and the Stiefel manifold.

6. Computational issues. The computationally heaviest operations in our algorithm are the computations of the reduced matrix $\hat{H} = V^T H(V \otimes V)$ and grad $f([V])$. The computation on grad $f([V])$ involves computing $\mathcal{I}^{(1)}(\mathcal{K}^{(1)} + \mathcal{S}^{(1)})^T$, $\mathcal{P}^{(2)}(\mathcal{K}^{(2)} + \mathcal{S}^{(2)})^T$ in (4.8) and solving the Sylvester equations, e.g., (3.8), (3.12), and (3.10). In addition, (3.8) and (3.12) also involve the computation related to the matrix H, i.e., $H(X_1 \otimes X_1)\hat{H}^T$ and $\mathcal{H}^{(2)}(X_1 \otimes Y_1)(\hat{\mathcal{H}}^{(2)})^T$. Hence, we can conclude that the computational issues in Algorithm 1 mainly contain two parts: the computations related to H and the solutions of the Sylvester equations. In the following, we will give the corresponding strategies to address these issues.

6.1. Computation related to the matrix $\boldsymbol{H}$ induced by Hessian. Let us focus on the computational issues of the terms related to H, such as $\mathcal{I}^{(1)}(\mathcal{K}^{(1)} + \mathcal{S}^{(1)})^T$ and $\mathcal{P}^{(2)}(\mathcal{K}^{(2)} + \mathcal{S}^{(2)})^T$ in (4.8), where the matrix multiplications are required when computing the components $\mathcal{I}^{(1)}$, $\mathcal{P}^{(2)}$, $\mathcal{K}^{(\mu)}$, and $\mathcal{S}^{(\mu)}$ ($\mu \in \{1,2\}$). These may cause unacceptable computational burden. Using the method in [16], which is presented based on the tensor multiplication properties, their corresponding computational costs can be reduced. In consideration of clear expression, we take $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T$ as the discussion target to describe the concrete procedure, shown in Algorithm 2. Other terms related to the matrix H can also be computed in a similar way.

Algorithm 2: Computation of $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T$ based on tensor multiplication

---

Input: The matrices H, V, Y_1, and X_1.
Output: $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T = H(V \otimes V)(Y_1^T H(X_1 \otimes X_1))^T$.

1. Restructure $\mathcal{X} \in \mathbb{R}^{n\times r\times n}$ such that $\mathcal{X}^{(2)} = V^T \mathcal{H}^{(2)}$.
2. Restructure $\mathcal{I} \in \mathbb{R}^{n\times r\times r}$ such that $\mathcal{I}^{(3)} = V^T \mathcal{X}^{(3)}$ and obtain $\mathcal{I}^{(1)}$.
3. Restructure $\mathcal{Y} \in \mathbb{R}^{r\times n\times n}$ such that $\mathcal{Y}^{(1)} = Y_1^T \mathcal{H}^{(1)}$.
4. Restructure $\mathcal{Z} \in \mathbb{R}^{r\times r\times n}$ such that $\mathcal{Z}^{(2)} = X_1^T \mathcal{Y}^{(2)}$.
5. Restructure $\mathcal{K} \in \mathbb{R}^{r\times r\times r}$ such that $\mathcal{K}^{(3)} = X_1^T \mathcal{Z}^{(3)}$ and obtain $\mathcal{K}^{(1)}$.
6. Compute $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T$.

In this paper, we give another computationally efficient strategy using the CP decomposition of tensors. Here, we still take $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T$ as the target to show the concrete procedure. For this, we first introduce some properties about tensors in [38, 44]. Since the tensor $\mathcal{H}$ is symmetric, the CP decomposition of $\mathcal{H}$ by a sum of R rank-1 tensors is written as

$$\mathcal{H} = \sum_{j=1}^{R} a_j \bullet b_j \bullet b_j, \tag{6.1}$$

where R is a positive constant and "$\bullet$" denotes the tensor product of vectors [38]. Denote by $A_c = [a_1\ a_2\ \cdots\ a_R] \in \mathbb{R}^{n\times R}$ and $B_c = [b_1\ b_2\ \cdots\ b_R] \in \mathbb{R}^{n\times R}$. Then, utilizing the Khatri–Rao product, one gets

$$\mathcal{H}^{(1)} = A_c(B_c \odot B_c)^T,\ \mathcal{H}^{(2)} = \mathcal{H}^{(3)} = B_c(B_c \odot A_c)^T.$$

The following relation between the Kronecker product and the Khatri–Rao product holds with $A_1 \in \mathbb{R}^{n_1\times m_1}$, $A_2 \in \mathbb{R}^{n_2\times m_2}$, $B_1 \in \mathbb{R}^{m_1\times q}$, and $B_2 \in \mathbb{R}^{m_2\times q}$:

$$(A_1 \otimes A_2)(B_1 \odot B_2) = (A_1B_1 \odot A_2B_2).$$

We thus have

$$\begin{aligned}\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T &= H(V \otimes V)(Y_1^T H(X_1 \otimes X_1))^T \\ &= A_c(B_c \odot B_c)^T(V \otimes V)(X_1^T \otimes X_1^T)(B_c \odot B_c)A_c^T Y_1 \\ &= A_c(V^T B_c \odot V^T B_c)^T(X_1^T B_c \odot X_1^T B_c)A_c^T Y_1.\end{aligned}$$

The steps of computing $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T$ are now listed in Algorithm 3.

Algorithm 3: Computation of $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T$ based on the CP decomposition

---

Input: The matrices H, V, Y_1, and X_1.
Output: $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T$.

1. Deal with $\mathcal{H}$ by the CP decomposition according to (6.1).
2. Obtain the matrices A_c and B_c.
3. Compute $\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T = A_c(V^T B_c \odot V^T B_c)^T(X_1^T B_c \odot X_1^T B_c)A_c^T Y_1$.

For $\hat{H} = V^T H(V \otimes V)$, we observe that its tensor $\hat{\mathcal{H}}$ is written as

$$\hat{\mathcal{H}} = \mathcal{H} \times_1 V^T \times_2 V^T \times_3 V^T = \sum_{j=1}^{R} (V^T a_i) \bullet (V^T b_i) \bullet (V^T b_i). \tag{6.2}$$

Hence, the reduced matrix $\hat{H}$ can be efficiently computed by $\hat{H} = V^T A_c (V^T B_c \odot V^T B_c)^T$. Using Algorithm 3 and (6.2), all the computations related to H can be obtained. We compare the computational costs of the terms related to H as they are computed by a direct computation, Algorithm 2, and Algorithm 3 in Table 6.1.

Table 6.1
Comparison of the computational costs.

	Direct method	Algorithm 2	Algorithm 3
$\hat{H}$	$O(n^3r + n^2r^3)$	$O(n^3r + n^2r^2)$	$O(nrR + Rr^3)$
$\mathcal{I}^{(1)}(\mathcal{K}^{(1)})^T$	$O(n^3r^2)$	$O(n^3r)$	$O(n^2R + n^2r + nRr^2)$
$\mathcal{I}^{(1)}(\mathcal{S}^{(1)})^T$	$O(n^3r^2)$	$O(n^3r)$	$O(n^2R + n^2r)$
$\mathcal{P}^{(2)}(\mathcal{K}^{(2)})^T$	$O(n^3r^2)$	$O(n^3r)$	$O(n^2R + n^2r + nRr^2)$
$\mathcal{P}^{(2)}(\mathcal{S}^{(2)})^T$	$O(n^3r^2)$	$O(n^3r)$	$O(n^2R + n^2r)$
$H(X_1 \otimes X_1)\hat{H}^T$	$O(n^3r^2)$	$O(n^3r)$	$O(n^2R + n^2r + nRr^2)$
$\mathcal{H}^{(2)}(X_1 \otimes Y_1)(\hat{H}^{(2)})^T$	$O(n^3r^2)$	$O(n^3r)$	$O(n^2R + n^2r + nRr^2)$

In practice, the matrix H in some systems, e.g., the system modeled by Burgers' equation, usually is sparse and only has $O(n)$ nonzero columns. Thus, on the basis of Algorithm 2, one can also use the structure of H to further reduce the computational costs. Since H is sparse, its mode-2 matricization $\mathcal{H}^{(2)}$ has the same property. Thereby, in numerical computation, one can first use the routines $size(\cdot)$ and $find(\cdot)$ in MATLAB to search for the nonzero columns of $\mathcal{H}^{(2)}$ and the nonzero elements in each nonzero column. One may then perform the steps in Algorithm 2.

6.2. Computation related to the Sylvester equations. We now discuss the computational issues for solving the Sylvester equations in Algorithm 1. They may cause great computational burden when the dimension n is large. However, as we know, various algorithms are available to solve them efficiently, e.g., the ADI iterations and the projection methods [45, 46]. It should be taken into account that in practice, the value of r is much smaller than the value of n. Thus, the Sylvester equations to be solved in Algorithm 1 usually have large A but small $\hat{A}$; for such case some effective computational methods have been discussed in [45].

7. Numerical examples. In this section, we illustrate the numerical performance of our algorithm by two standard test examples for MOR. Since the RMPRP process is also applicable for bilinear systems, the test examples consist of bilinear and QB systems. All the simulation results are obtained by using MATLAB 9.1.0.441655(R2016b) on two Intel Xeon E5-2650 CPUs with 2.2GHz, and all the original and reduced systems are integrated by the routine ode15s in MATLAB.

The proposed algorithm with the Armijo-type rule and the Wolfe condition are denoted by $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$, respectively. We compare them with some existing methods, e.g., the proper orthogonal decomposition (POD) method in [1] and some moment-matching methods, including the QLMOR method in [15] and the one-sided interpolatory projection method in [16] for QB systems, and their counterparts in the bilinear case. For the POD method, we simply collect the uniformly distributed snapshots of the true solution for a given training input $u(t)$ and use r dominant modes to generate the POD-based projection matrix. As shown in [5, 16], the linear $\mathcal{H}_2$ optimal points are applied to generate interpolation points of the one-sided interpolatory projection method.

Parareal (parallel in time) usually chooses approximate values with low accuracy at coarse time points as initial guesses [47, 48]. Following the idea behind parareal, we first construct an initial reduced system with low accuracy before we run $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$. Concretely, the starting points are generated by using some linear MOR methods on the corresponding linear part, such as the Krylov subspace method and the iterative rational Krylov algorithm (IRKA). IRKA can be achieved by the code from https://gitlab.mpi-magdeburg.mpg.de/saak/best_practice_IRKA. By observation, it is seen that the above initialization strategy takes less computational cost and computational time since the bilinear and quadratic terms of the original systems are not involved. In our numerical tests, both $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ are initialized by the same starting points and stopped with the same stopping criterion, where $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ are run until the norm of the gradient becomes smaller than a given tolerance or until a maximum of 180 iterations is reached. We use the algorithms [49, A 6.3.1, A 6.3.1mod] based on quadratic or cubic polynomial interpolation to find the step-size satisfying the Wolfe condition (4.29). Moreover, the computational strategies in section 6 are adopted to compute the Riemannian gradient, where the Sylvester equations are solved by using the method in [45].

7.1. A nonlinear transmission line circuit. We first consider a nonlinear transmission line circuit, i.e., the RC circuit [15]. A detailed derivation therein is given to rewrite the RC circuit system in QB form with order $n = 2k$, where k denotes the number of capacitors. In this example, we set $k = 600$, and a QB system of size $n = 1200$ is then obtained. Notice that A is unstable. Since the stability of A and $\hat{A}$ is necessary for the existences of the truncated Gramians and the truncated $\mathcal{H}_2$ error, we replace the matrix A with $\tilde{A} := A - 0.0095 I_n$ and project the QB system with $\tilde{A}$ to construct its reduced systems.

In this example, reduced systems are constructed by POD, QLMOR, $\mathrm{RMPRP_A}$, $\mathrm{RMPRP_W}$, and the one-sided interpolatory projection method. The linear $\mathcal{H}_2$ optimal interpolation points used in the one-sided interpolatory projection method are generated by IRKA. The starting points of $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ are chosen from the QR decomposition of the combination of the projection matrices constructed by IRKA to the linear part of the QB system.

First, we set the reduced order $r = 10$. The POD-based reduced system is obtained by using 250 snapshots for $u(t) = \sin(t)\cos(t)$. For $\mathrm{RMPRP_A}$, the parameters involved in computing the step-size are chosen as $\rho = 0.65$ and $\varrho = 1.0$, while the parameters in $\mathrm{RMPRP_W}$ are set to be $\delta_1 = 0.41$ and $\delta_2 = 0.55$. With the given parameters, we initialize $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ and iterate them until the norm of the gradient becomes smaller than 2.0e-4. Due to the different line-search procedures, a potential difference between $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ is the computation numbers of $f([V_i])$ and grad $f([V_i])$ at each iteration. It finds that $\mathrm{RMPRP_A}$ takes once to compute $f([V_i])$ per iteration, while it takes l times to compute grad $f([V_i])$. Here, l is the nonnegative integer shown above (4.28). From the procedure used to achieve (4.29), we observe that the numbers taken to compute $f([V_i])$ and grad $f([V_i])$ in $\mathrm{RMPRP_W}$ are equal per iteration. For convenience,

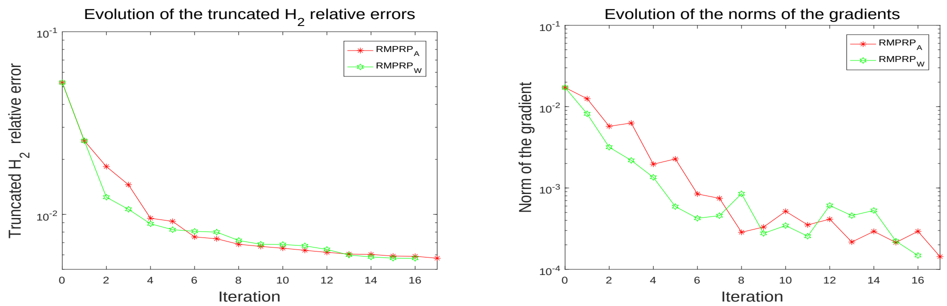

FIG. 7.1. *Evolution of the truncated $\mathcal{H}_2$ relative errors and the norms of the gradients, $r = 10$.*

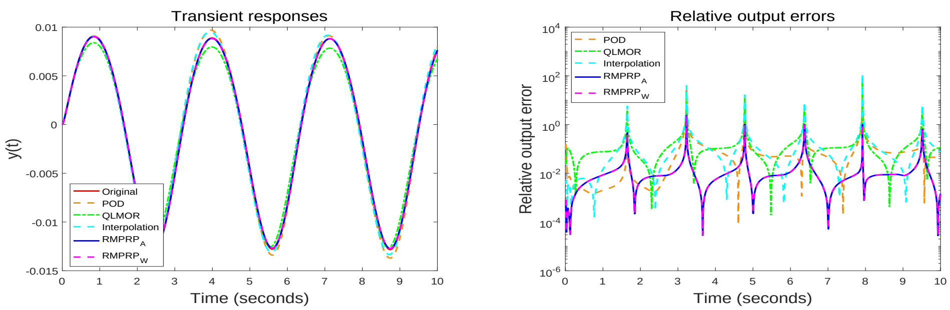

FIG. 7.2. *Comparison of the output responses for $u(t) = \sin(t)\cos(t)$, $r = 10$.*

we compute their average computation numbers by counting the values at all iterations. In $\mathrm{RMPRP_A}$, we see that the average computation number for grad $f([V_i])$ is 5, and the average numbers to compute $f([V_i])$ and grad $f([V_i])$ in $\mathrm{RMPRP_W}$ both are 4. Figure 7.1 plots the convergence behavior of $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$, where the truncated $\mathcal{H}_2$ relative error is defined as $\|\Sigma - \hat{\Sigma}\|_{\mathcal{H}_2^{\mathcal{T}}}/\|\Sigma\|_{\mathcal{H}_2^{\mathcal{T}}}$. It is observed that the proposed algorithm with two different choices of step-sizes performs similar convergence behavior. Figure 7.2 shows the transient responses and the corresponding relative output errors of all the reduced systems for $u(t) = \sin(t)\cos(t)$. It is seen that $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ have good performances in simulating this example.

The approximation performances of the reduced systems with $r = 2, 6$ are also tested by us. Similarly, we start $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ by IRKA. For the case of $r = 2$, the evolution of the truncated $\mathcal{H}_2$ relative errors and the norms of the gradients is shown in Figure 7.3. From Figure 7.3, it can be seen that both $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ converge in 10 iterations, and they converge toward the local minimum. For the case of $r = 6$, Figure 7.4 plots the transient responses and the relative output errors for $u(t) = \exp(\sin(t))$, where the POD approximation is yielded by using 350 snapshots. The truncated $\mathcal{H}_2$ relative errors for different reduced orders are listed in Table 7.1. It is observed that $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ can yield higher accuracy in the sense of the truncated $\mathcal{H}_2$ norm. With the same initial points and the stopping conditions, the truncated $\mathcal{H}_2$ relative errors of $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ are slightly different since they are achieved in two different line-search ways.

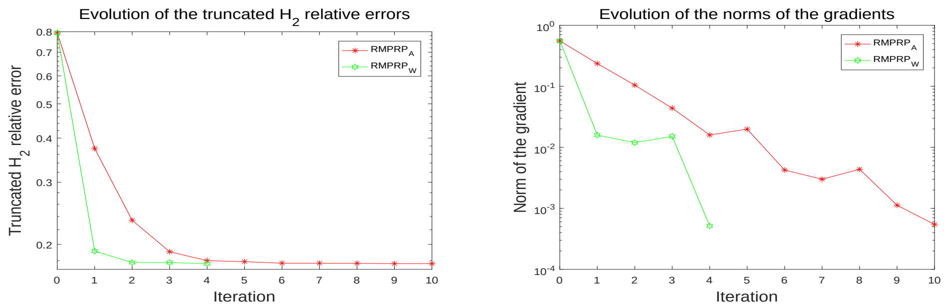

FIG. 7.3. *Evolution of the truncated $\mathcal{H}_2$ relative errors and the norms of the gradients, $r = 2$.*

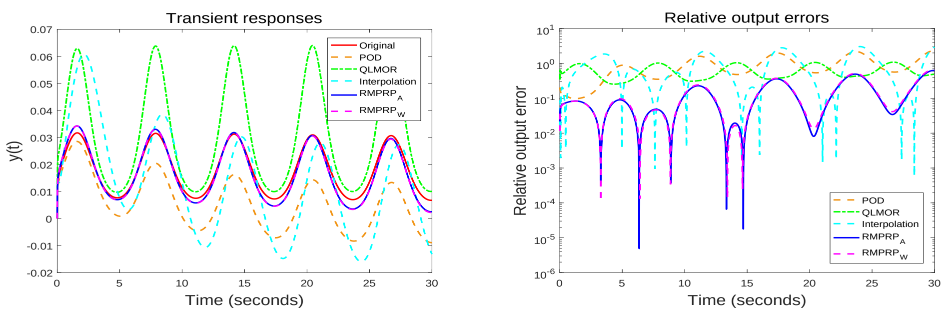

FIG. 7.4. *Comparison of the output responses for $u(t) = \exp(\sin(t))$, $r = 6$.*

TABLE 7.1
Comparison of the truncated $\mathcal{H}_2$ relative errors.

	POD	QLMOR	Interpolatory	RMPRP$_\mathrm{A}$	RMPRP$_\mathrm{W}$
$r = 2$	5.48e-1	3.53e-1	1.15e+1	1.76e-1	1.76e-1
$r = 6$	3.07e-1	2.99e-1	3.80e-1	2.90e-2	2.80e-2
$r = 10$	2.37e-1	1.76e-1	4.03e-2	5.70e-3	5.75e-3

7.2. A heat transfer model. The second example is the heat transfer model

$$\begin{aligned} y_t &= \Delta y && \text{in } (0,1)\times(0,1), \\ n\cdot\nabla y &= 0.5\cdot u_{1,2}(y-1) && \text{on } \Gamma_1, \Gamma_2, \\ y &= 0 && \text{on } \Gamma_3, \Gamma_4, \end{aligned}$$

where Γ_1, Γ_2, Γ_3, and Γ_4 denote the boundaries of region [36]. We consider κ^2 grid points in the spatial domain and discretize this model by a finite difference scheme. We set $\kappa = 35$. Thus, a 2-input and 1-output bilinear system of order 1225 is obtained.

In this example, POD, QLMOR, RMPRP$_\mathrm{A}$, RMPRP$_\mathrm{W}$, the balanced truncation method and the one-sided interpolatory projection method are implemented to construct the reduced systems. The QLMOR method is proposed for QB systems, yet it degenerates to the Krylov subspace method for bilinear systems as the quadratic term of QB systems is eliminated. Hence, we use this method to reduce this bilinear system and still denote it by "QLMOR." The one-sided interpolatory projection method is a moment-matching-type method presented for single-input and single-output QB systems. For bilinear systems, this method is also suitable for multiple-input cases. Thus, we apply it to the second model, where the linear $\mathcal{H}_2$ optimal interpolation points are chosen as the matching points. To generate the reduced system by POD, 450 snapshots for $u(t) = [-2\exp(\sin(\pi t)),\ \sin(2\pi t)]^T$ are collected.

We now construct the reduced systems of order $r = 6$. The parameters in $\mathrm{RMPRP_A}$ are set to be $\rho = 0.55$ and $\varrho = 0.00191$, and the parameters in $\mathrm{RMPRP_W}$ are chosen as $\delta_1 = 0.0591$ and $\delta_2 = 0.0699$. We first try to initialize $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ by IRKA. The stopping criterion is $\|\mathrm{grad}\ f([V_i])\|_{[V_i]} \leq 1.0\mathrm{e}-3$. However, it is observed that they converge slowly. For $\mathrm{RMPRP_A}$, it requires 155 iterations to converge, while it does not satisfy the stopping condition for $\mathrm{RMPRP_W}$ even when the maximum number of iterations is reached. To reduce the number of iterations, we start $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ by applying the Krylov subspace method on the linear part of this bilinear system. Using this initial point, $\mathrm{RMPRP_A}$ converges after 35 iterations and $\mathrm{RMPRP_W}$ takes 28 iterations to converge. Table 7.2 displays the truncated $\mathcal{H}_2$ relative errors. The computation numbers taken to evaluate $f([V_i])$ and grad $f([V_i])$ per iteration are also measured. For $\mathrm{RMPRP_A}$, it takes once to compute $f([V_i])$ and the average computation number related to grad $f([V_i])$ is 2. For $\mathrm{RMPRP_W}$, the average computation numbers to compute $f([V_i])$ and grad $f([V_i])$ are equal, and both are 6.

We then generate the reduced systems of order $r = 14$ using the above six methods. We run $\mathrm{RMPRP_A}$ and $\mathrm{RMPRP_W}$ until $\|\mathrm{grad}\ f([V_i])\|_{[V_i]} \leq 1.0\mathrm{e}\text{-}4$. Figure 7.5 plots the evolution of the truncated $\mathcal{H}_2$ relative errors and the norms of the gradients. We report the corresponding truncated $\mathcal{H}_2$ relative errors in Table 7.2 as well. Additionally, the accuracy of these reduced systems for the input $u(t) = [-2\exp(\sin(\pi t)),\ \sin(2\pi t)]^T$ are compared, where the relative output errors are given in Figure 7.6.

From these simulation results, it is seen that the balanced truncated method, $\mathrm{RMPRP_A}$, and $\mathrm{RMPRP_W}$ can yield good reduced systems. Moreover, Table 7.2 shows that they produce better reduced systems in the truncated $\mathcal{H}_2$ norm measure.

TABLE 7.2
Comparison of the truncated $\mathcal{H}_2$ relative errors.

	POD	QLMOR	Interpolation	BT	$\mathrm{RMPRP_A}$	$\mathrm{RMPRP_W}$
$r = 6$	1.24e-1	1.46e-1	1.67e-1	3.50e-2	4.73e-2	4.69e-2
$r = 14$	9.61e-2	2.26e-2	7.70e-2	1.32e-3	7.50e-3	7.41e-3

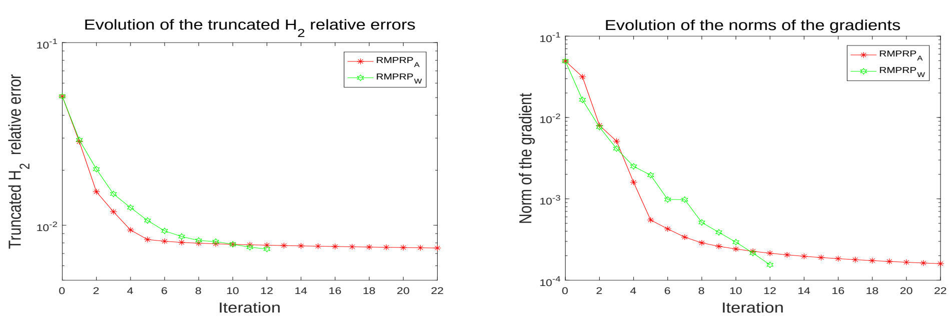

FIG. 7.5. *Evolution of the truncated $\mathcal{H}_2$ relative errors and the norms of the gradients, $r = 14$.*

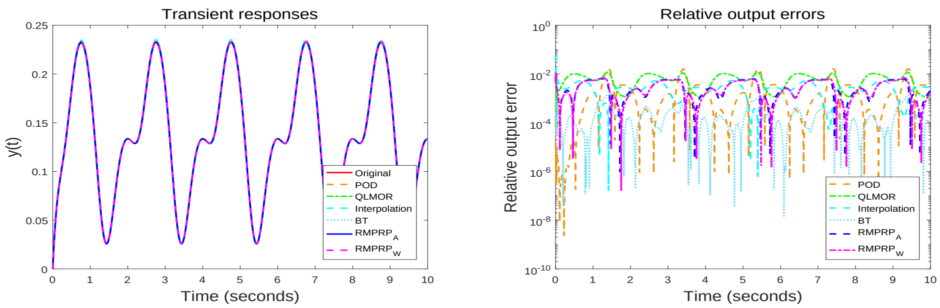

FIG. 7.6. *Comparison of the output responses for* $u(t) = [-2\exp(\sin(\pi t)),\ \sin(2\pi t)]^T$, $r = 14$.

8. Conclusions. In this paper, the truncated $\mathcal{H}_2$ optimal MOR problem for QB systems has been explored. We reformulate this optimal problem as a Riemannian optimization problem on the Grassmann manifold. Its compactness ensures the existence of the minimum solution of our problem. By exploiting the Riemannian gradient of the cost function and the geometric properties, we design a new Riemannian MPRP conjugate gradient scheme. The RMPRP algorithm is then presented to construct the reduced systems. We show that the search direction in RMPRP is the descent direction of the cost function and prove the global convergence of the proposed algorithm. The numerical experiments show that the geometry chosen of the Grassmann manifold makes an efficient implementation of the proposed algorithm.

Appendix A. Proof of Lemma 2.1. We begin this proof by deriving

$$
\begin{aligned}
\mathcal{Z}^{(1)} &= \begin{bmatrix} \mathcal{X}^{(1)} T_{(n_1,n_2)} \\ \mathcal{Y}^{(1)} \widehat{T}_{(n_1,n_2)} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathcal{X}_1 & \cdots & \mathcal{X}_{n_1} \end{bmatrix} \left(\begin{bmatrix} I_{n_1} & 0_{n_1\times n_2} \end{bmatrix} \otimes \begin{bmatrix} I_{n_1} & 0_{n_1\times n_2} \end{bmatrix} \right) \\ \begin{bmatrix} \mathcal{Y}_1 & \cdots & \mathcal{Y}_{n_2} \end{bmatrix} \left(\begin{bmatrix} 0_{n_2\times n_1} & I_{n_2} \end{bmatrix} \otimes \begin{bmatrix} 0_{n_2\times n_2} & I_{n_2} \end{bmatrix} \right) \end{bmatrix} \\
&= \begin{bmatrix} [\mathcal{X}_1\ 0_{n_1\times n_2}] & \cdots & [\mathcal{X}_{n_1}\ 0_{n_1\times n_2}] & 0_{n_1\times(n_1+n_2)} & \cdots & 0_{n_1\times(n_1+n_2)} \\ 0_{n_2\times(n_1+n_2)} & \cdots & 0_{n_2\times(n_1+n_2)} & [0_{n_2\times n_1}\ \mathcal{Y}_1] & \cdots & [0_{n_2\times n_1}\ \mathcal{Y}_{n_2}] \end{bmatrix}.
\end{aligned}
$$

From the above formula, we find that the frontal slices of $\mathcal{Z}$ can be given by

$$
\mathcal{Z}_i = \begin{cases} \begin{bmatrix} [\mathcal{X}_i \quad 0_{n_1\times n_2}] \\ 0_{n_2\times(n_1+n_2)} \end{bmatrix}, & i = 1, 2, \ldots, n_1; \\ \begin{bmatrix} 0_{n_1\times(n_1+n_2)} \\ [0_{n_2\times n_1} \quad \mathcal{Y}_{i-n_1}] \end{bmatrix}, & i = n_1+1, \ldots, n_1+n_2. \end{cases}
$$

When $i = 1, 2, \ldots, n_1$, it follows that

$$
\mathcal{Z}_i^T = \begin{bmatrix} \begin{bmatrix} \mathcal{X}_i^T \\ 0_{n_2\times n_1} \end{bmatrix} & 0_{(n_1+n_2)\times n_2} \end{bmatrix} = \begin{bmatrix} [\mathcal{X}_i^T\ 0_{n_1\times n_2}] \\ 0_{n_2\times(n_1+n_2)} \end{bmatrix},
$$

and for $i = n_1+1, \ldots, n_1+n_2$, it follows that

$$
\mathcal{Z}_i^T = \begin{bmatrix} 0_{(n_1+n_2)\times n_1} & \begin{bmatrix} 0_{n_1\times n_2} \\ \mathcal{Y}_{i-n_1}^T \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0_{n_1\times(n_1+n_2)} \\ [0_{n_2\times n_1}\ \mathcal{Y}_{i-n_1}^T] \end{bmatrix}.
$$

Based on this, one has

$$
\begin{aligned}
\mathcal{Z}^{(2)} &= \begin{bmatrix} \mathcal{Z}_1^T & \cdots & \mathcal{Z}_{(n_1+n_2)}^T \end{bmatrix} \\
&= \begin{bmatrix} [\mathcal{X}_1^T \ 0_{n_1\times n_2}] & \cdots & [\mathcal{X}_{n_1}^T \ 0_{n_1\times n_2}] & 0_{n_1\times(n_1+n_2)} & \cdots & 0_{n_1\times(n_1+n_2)} \\ 0_{n_2\times(n_1+n_2)} & \cdots & 0_{n_2\times(n_1+n_2)} & [0_{n_2\times n_1} \ \mathcal{Y}_1^T] & \cdots & [0_{n_2\times n_1} \ \mathcal{Y}_{n_2}^T] \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} \mathcal{X}_1^T & \mathcal{X}_2^T & \cdots & \mathcal{X}_{n_1}^T \end{bmatrix} T_{(n_1,n_2)} \\ \begin{bmatrix} \mathcal{Y}_1^T & \mathcal{Y}_2^T & \cdots & \mathcal{Y}_{n_2}^T \end{bmatrix} \widehat{T}_{(n_1,n_2)} \end{bmatrix} = \begin{bmatrix} \mathcal{X}^{(2)} T_{(n_1,n_2)} \\ \mathcal{Y}^{(2)} \widehat{T}_{(n_1,n_2)} \end{bmatrix}.
\end{aligned}
$$

Next, we prove the result about $\mathcal{Z}^{(3)}$. Let $\mathcal{X}_{ij}$ be the ith column of the frontal slice $\mathcal{X}_j$ $(i,j=1,2,\ldots,n_1)$ and $\mathcal{Y}_{qk}$ be the qth column of the frontal slice $\mathcal{Y}_k$ $(q,k=1,2,\ldots,n_2)$. Then, it holds that

$$
\begin{aligned}
\mathcal{X}^{(3)} T_{(n_1,n_2)} &= \begin{bmatrix} \operatorname{vec}(\mathcal{X}_1)^T \\ \vdots \\ \operatorname{vec}(\mathcal{X}_{n_1})^T \end{bmatrix} T_{(n_1,n_2)} = \begin{bmatrix} \mathcal{X}_{11}^T & \mathcal{X}_{21}^T & \cdots & \mathcal{X}_{n_1 1}^T \\ \vdots & \vdots & & \vdots \\ \mathcal{X}_{1n_1}^T & \mathcal{X}_{2n_1}^T & \cdots & \mathcal{X}_{n_1 n_1}^T \end{bmatrix} T_{(n_1,n_2)} \\
&= \begin{bmatrix} \mathcal{X}_{11}^T & 0_{1\times n_2} & \cdots & \mathcal{X}_{n_1 1}^T & 0_{1\times n_2} & 0_{1\times(n_1+n_2)} & \cdots & 0_{1\times(n_1+n_2)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mathcal{X}_{1n_1}^T & 0_{1\times n_2} & \cdots & \mathcal{X}_{n_1 n_1}^T & 0_{1\times n_2} & 0_{1\times(n_1+n_2)} & \cdots & 0_{1\times(n_1+n_2)} \end{bmatrix} \\
&= \begin{bmatrix} \operatorname{vec}\left(\begin{bmatrix} \mathcal{X}_1 \\ 0_{n_2\times n_1} \end{bmatrix}\right)^T & \operatorname{vec}\left(0_{(n_1+n_2)\times n_2}\right)^T \\ \vdots & \vdots \\ \operatorname{vec}\left(\begin{bmatrix} \mathcal{X}_{n_1} \\ 0_{n_2\times n_1} \end{bmatrix}\right)^T & \operatorname{vec}\left(0_{(n_1+n_2)\times n_2}\right)^T \end{bmatrix} = \begin{bmatrix} \operatorname{vec}(\mathcal{Z}_1)^T \\ \vdots \\ \operatorname{vec}(\mathcal{Z}_{n_1})^T \end{bmatrix}.
\end{aligned}
$$

Similarly, according to the mode-3 matricization $\mathcal{Y}^{(3)}$, one obtains

$$
\begin{aligned}
\mathcal{Y}^{(3)} \hat{T}_{(n_1,n_2)} &= \begin{bmatrix} \operatorname{vec}(\mathcal{Y}_1)^T \\ \vdots \\ \operatorname{vec}(\mathcal{Y}_{n_2})^T \end{bmatrix} \hat{T}_{(n_1,n_2)} = \begin{bmatrix} \mathcal{Y}_{11}^T & \mathcal{Y}_{21}^T & \cdots & \mathcal{Y}_{n_2 1}^T \\ \vdots & \vdots & & \vdots \\ \mathcal{Y}_{1n_2}^T & \mathcal{Y}_{2n_2}^T & \cdots & \mathcal{Y}_{n_2 n_2}^T \end{bmatrix} \hat{T}_{(n_1,n_2)} \\
&= \begin{bmatrix} 0_{1\times(n_1+n_2)} & \cdots & 0_{1\times(n_1+n_2)} & 0_{1\times n_1} & \mathcal{Y}_{11}^T & \cdots & 0_{1\times n_1} & \mathcal{Y}_{n_2 1}^T \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0_{1\times(n_1+n_2)} & \cdots & 0_{1\times(n_1+n_2)} & 0_{1\times n_1} & \mathcal{Y}_{1n_2}^T & \cdots & 0_{1\times n_1} & \mathcal{Y}_{n_2 n_2}^T \end{bmatrix} \\
&= \begin{bmatrix} \operatorname{vec}\left(0_{(n_1+n_2)\times n_1}\right)^T & \operatorname{vec}\left(\begin{bmatrix} 0_{n_1\times n_2} \\ \mathcal{Y}_1 \end{bmatrix}\right)^T \\ \vdots & \vdots \\ \operatorname{vec}\left(0_{(n_1+n_2)\times n}\right)^T & \operatorname{vec}\left(\begin{bmatrix} 0_{n_1\times n_2} \\ \mathcal{Y}_{n_2} \end{bmatrix}\right)^T \end{bmatrix} = \begin{bmatrix} \operatorname{vec}(\mathcal{Z}_{n_1+1})^T \\ \vdots \\ \operatorname{vec}(\mathcal{Z}_{n_1+n_2})^T \end{bmatrix},
\end{aligned}
$$

which yields

$$
\mathcal{Z}^{(3)} = \begin{bmatrix} \operatorname{vec}(\mathcal{Z}_1)^T \\ \vdots \\ \operatorname{vec}(\mathcal{Z}_{n_1+n_2})^T \end{bmatrix} = \begin{bmatrix} \mathcal{X}^{(3)} T_{(n_1,n_2)} \\ \mathcal{Y}^{(3)} \hat{T}_{(n_1,n_2)} \end{bmatrix} = \begin{bmatrix} \mathcal{X}^{(2)} T_{(n_1,n_2)} \\ \mathcal{Y}^{(2)} \hat{T}_{(n_1,n_2)} \end{bmatrix} = \mathcal{Z}^{(2)}.
$$

Thus, the proof is completed.

Appendix B. (1) Some explanations for (3.12)–(3.14). We first show the derivation of (3.12)–(3.14). Using Lemma 2.1, one gets the explicit expression of $\mathcal{H}_e^{(2)}$, namely,

$$\mathcal{H}_e^{(2)} = \begin{bmatrix} \mathcal{H}^{(2)} T_{(n,r)} \\ \hat{\mathcal{H}}^{(2)} \hat{T}_{(n,r)} \end{bmatrix}. \tag{B.1}$$

We substitute the partition of $Q_{\mathcal{T}e}$ in (3.3) into the second equation of (3.4) and also substitute the partition of Q_{1e} in (3.5) into the second equation of (3.6). Combining with (B.1), it is clear that (3.12)–(3.14) hold.

(2) Some explanations for (4.3)–(4.6). In subsection 2.2, we have noted the relations given in (2.7). For the error system Σ_e, the relation $Q_{\mathcal{T}e} = \sum_{i=1}^{3} Q_{ie}$ also holds, where Q_{1e} has been given in (3.6), and Q_{ie} for $i = 2, 3$ respectively satisfy

$$A_e^T Q_{2e} + Q_{2e} A_e + \sum_{k=1}^{m} N_{k,e}^T Q_{1e} N_{k,e} = 0, \tag{B.2a}$$

$$A_e^T Q_{3e} + Q_{3e} A_e + \mathcal{H}_e^{(2)} (P_{1e} \otimes Q_{1e}) (\mathcal{H}_e^{(2)})^T = 0. \tag{B.2b}$$

Similar to (3.5), we partition Q_{2e} and Q_{3e} as

$$Q_{2e} = \begin{bmatrix} Q_2 & Y_2 \\ Y_2^T & \hat{Q}_2 \end{bmatrix}, \ Q_{3e} = \begin{bmatrix} Q_3 & Y_3 \\ Y_3^T & \hat{Q}_3 \end{bmatrix}. \tag{B.3}$$

Substituting (B.1) and (B.3) into (B.2), one can find that Y_i and $\hat{Q}_i$ for $i = 2, 3$ satisfy (4.3)–(4.5). Moreover, combining with the partitions in (3.3), (3.5), and (B.3), it is observed that $Y_{\mathcal{T}} = \sum_{i=1}^{3} Y_i$ and $\hat{Q}_{\mathcal{T}} = \sum_{i=1}^{3} \hat{Q}_i$.

Appendix C. Proof of Theorem 5.9. In order to prove the global convergence of Algorithm 1 under the Wolfe step-size control (4.29), we first give the Zoutendijk condition [42], i.e.,

$$\sum_{i \geq 0} \frac{\|\text{grad } f([V_i])\|_{[V_i]}^4}{\|\eta_i\|_{[V_i]}^2} < +\infty. \tag{C.1}$$

In fact, according to Lemma 5.4 and (4.29b) on the Wolfe condition, one gets

$$(\delta_2 - 1) \frac{d}{dt} f(\mathcal{R}_{[V_i]}(t\eta_i))|_{t=0} \leq \frac{d}{dt} f(\mathcal{R}_{[V_i]}(t\eta_i))|_{t=\alpha_i} - \frac{d}{dt} f(\mathcal{R}_{[V_i]}(t\eta_i))|_{t=0} \leq l_1 \alpha_i \|\eta_i\|_{[V_i]}^2,$$

which implies $(\delta_2 - 1)\langle \text{grad } f([V_i]), \eta_i \rangle_{[V_i]} \leq l_1 \alpha_i \|\eta_i\|_{[V_i]}^2$. This yields

$$\alpha_i \geq \frac{\delta_2 - 1}{l_1} \cdot \frac{\langle \text{grad } f([V_i]), \eta_i \rangle_{[V_i]}}{\|\eta_i\|_{[V_i]}^2}.$$

By (4.29a) and (4.30), it further holds that

$$f([V_i]) - f(\mathcal{R}_{[V_i]}(\alpha_i \eta_i)) \geq \frac{\delta_1(1 - \delta_2)}{l_1} \cdot \frac{\|\text{grad } f([V_i])\|_{[V_i]}^4}{\|\eta_i\|_{[V_i]}^2}.$$

Summing the above formula over all i, the formula (C.1) follows by the boundedness of f.

Next, a contradiction is used to show the global convergence. We assume that the opposite of (5.2) holds, namely, there is a constant $\varepsilon_3 > 0$ such that $\|\text{grad } f([V_i])\|_{[V_i]} \geq \varepsilon_3$ for all i. By (4.27) and Lemma 5.5, we obtain

$$\begin{aligned}
\|\eta_i\|^2_{[V_i]} &= \| - \text{grad } f([V_i]) + \beta_i^{\text{PRP}} \mathcal{T}_{\alpha_{i-1}\eta_{i-1}} \eta_{i-1} - \theta_i y_{i-1}\|^2_{[V_i]} \\
&= \|\text{grad } f([V_i])\|^2_{[V_i]} + |\beta_i^{\text{PRP}}|^2 \|\mathcal{T}_{\alpha_{i-1}\eta_{i-1}} \eta_{i-1}\|^2_{[V_i]} + |\theta_i|^2 \|y_{i-1}\|^2_{[V_i]} \\
&\quad - 2\theta_i \beta_i^{\text{PRP}} \langle \mathcal{T}_{\alpha_{i-1}\eta_{i-1}} \eta_{i-1}, y_{i-1}\rangle_{[V_i]} \\
&\le \|\text{grad } f([V_i])\|^2_{[V_i]} + 4 \frac{\|\text{grad } f([V_i])\|^2_{[V_i]} \|y_{i-1}\|^2_{[V_i]} \|\mathcal{T}_{\alpha_{i-1}\eta_{i-1}} \eta_{i-1}\|^2_{[V_i]}}{\|\text{grad } f([V_{i-1}])\|^4_{[V_{i-1}]}} \\
&\le \|\text{grad } f([V_i])\|^2_{[V_i]} + 4 \frac{l_2^2 \|\text{grad } f([V_i])\|^2_{[V_i]} \|\alpha_{i-1}\eta_{i-1}\|^2_{[V_{i-1}]} \|\eta_{i-1}\|^2_{[V_{i-1}]}}{\|\text{grad } f([V_{i-1}])\|^4_{[V_{i-1}]}}.
\end{aligned}$$

Hence, it further results in

$$\begin{aligned}
\frac{\|\eta_i\|^2_{[V_i]}}{\|\text{grad } f([V_i])\|^4_{[V_i]}} &\le \frac{1}{\|\text{grad } f([V_i])\|^2_{[V_i]}} + \frac{\|\eta_{i-1}\|^2_{[V_{i-1}]}}{\|\text{grad } f([V_{i-1}])\|^4_{[V_{i-1}]}} \\
&= \sum_{j=0}^{i} \frac{1}{\|\text{grad } f([V_j])\|^2_{[V_j]}} \le \frac{i+1}{\varepsilon_3^2}.
\end{aligned}$$

It suggests that

$$\sum_{i\ge 1} \frac{\|\text{grad } f([V_i])\|^4_{[V_i]}}{\|\eta_i\|^2_{[V_i]}} \ge \sum_{i\ge 1} \frac{\varepsilon_3^2}{i+1} = +\infty,$$

which contradicts the Zoutendijk condition (C.1). Thus, this completes the proof.

Acknowledgments. The authors would like to thank Peter Benner and anonymous reviewers for their valuable suggestions and helpful comments, which greatly improved the content and the presentation of the earlier version of the paper.

REFERENCES

[1] Y. L. Jiang, *Model Order Reduction Methods*, Science Press, Beijing, 2010 (in Chinese).

[2] P. Van Dooren, K. A. Gallivan, and P. A. Absil, *$\mathcal{H}_2$-optimal model reduction of MIMO systems*, Appl. Math. Lett., 21 (2008), pp. 1267–1273.

[3] K. Gallivan, G. Grimme, and P. Van Dooren, *A rational Lanczos algorithms for model reduction*, Numer. Algorithms, 12 (1996), pp. 33–63.

[4] J. Phillips, L. Daniel, and L. M. Silveira, *Guaranteed passive balancing transformations for model order reduction*, IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., 22 (2003), pp. 1027–1041.

[5] S. Gugercin, A. C. Antoulas, and C. Beattie, *$\mathcal{H}_2$ model reduction for large-scale linear dynamical system*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 609–638.

[6] Y. L. Jiang, Z. Z. Qi, and P. Yang, *Model order reduction of linear systems via the cross Gramian and SVD*, IEEE Trans. Circuits Syst. II, 66 (2019), pp. 422–426.

[7] Z. J. Bai and D. Skoogh, *A projection method for model reduction of bilinear dynamical system*, Linear Algebra Appl., 415 (2006), pp. 406–425.

[8] L. Zhang and J. Lam, *On $\mathcal{H}_2$ model reduction of bilinear systems*, Automatica, 38 (2002), pp. 205–216.

[9] T. Breiten and T. Damm, *Krylov subspace methods for model order reduction of bilinear control systems*, Systems Control Lett., 59 (2010), pp. 443–450.

[10] G. Flagg and S. Gugercin, *Multipoint volterra series interpolation and $\mathcal{H}_2$ optimal model reduction of bilinear systems*, SIAM J. Matrix. Anal. Appl., 36 (2015), pp. 549–579.

[11] P. Benner, V. Mehrmann, and D. C. Sorensen, *Dimension Reduction of Large-Scale Systems*, Springer, Berlin, 2005.

[12] M. Rewienski and J. White, *A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices*, IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., 22 (2003), pp. 155–170.

[13] K. Fujimoto and J. M. A. Scherpen, *Balanced realization and model order reduction for nonlinear systems based on singular value analysis*, SIAM J. Control Optim., 48 (2010), pp. 4591–4623.

[14] H. Liu, L. Daniel, and N. Wong, *Model reduction and simulation of nonlinear circuits via tensor decomposition*, IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., 34 (2015), pp. 1059–1069.

[15] C. Gu, *QLMOR: A projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems*, IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., 30 (2011), pp. 1307–1320.

[16] P. Benner and T. Breiten, *Two-sided projection methods for nonlinear model order reduction*, SIAM J. Sci. Comput., 37 (2015), pp. B239–B260.

[17] Y. Zhang, H. Liu, Q. Wang, N. Fong, and N. Wong, *Fast nonlinear model order reduction via associated transforms of high-order Volterra transfer functions*, in Proceedings of the 49th ACM/EDAC/IEEE Design Automation Conference, 2012, pp. 289–294.

[18] Z. Li, Y. L. Jiang, and K. L. Xu, *Non-linear model order reduction based on tensor decomposition and matrix product*, IET Control Theory Appl., 12 (2018), pp. 2253–2262.

[19] P. Benner, P. Goyal, and S. Gugercin, *$\mathcal{H}_2$-quasi-optimal model order reduction for quadratic-bilinear control systems*, SIAM J. Matrix. Anal. Appl., 39 (2018), pp. 983-1032.

[20] P. A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifold*, Princeton University Press, Princeton, NJ, 2008.

[21] S. T. Smith, *Geometric Optimization Methods for Adaptive Filtering*, Ph.D. thesis, Division of Applied Sciences, Harvard University, Cambridge, MA, 1993.

[22] L. Eldén and B. Savas, *A Newton-Grassmann method for computing the best multilinear rank-(r_1, r_2, r_3) approximation of a tensor*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 248–271.

[23] M. Ishteva, P. A. Absil, S. V. Huffel, and L. D. Lathauwer, *Best low multilinear rank approximation of higher-order tensors, based on the Riemannian trust-region scheme*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 115–135.

[24] H. D. Sterck and A. Howse, *Nonlinearly preconditioned optimization on Grassmann manifolds for computing approximate Tucker tensor decompositions*, SIAM J. Sci. Comput., 38 (2016), pp. A997–A1018.

[25] B. Savas and L. H. Lim, *Quasi-Newton methods on Grassmannians and multilinear approximations of tensors*, SIAM J. Sci. Comput., 32 (2010), pp. 3352–3393.

[26] Z. Zhao, Z. J. Bai, and X. Q. Jin, *A Riemannian Newton algorithm for nonlinear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 752–774.

[27] B. Vandereycken and S. Vandewalle, *A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2553–2579.

[28] W. Dai, E. Kerman, and O. Milenkovic, *A geometric approach to low-rank matrix completion*, IEEE Trans. Inform. Theory, 58 (2012), pp. 237–247.

[29] W. Y. Yan and J. Lam, *An approximate approach to $\mathcal{H}_2$ optimal model reduction*, IEEE Trans. Automat. Control, 44 (1999), pp. 1341–1358.

[30] H. Sato and K. Sato, *Riemannian optimal control and model matching of linear port-Hamiltonian systems*, IEEE Trans. Automat. Control, 62 (2017), pp. 6575–6581.

[31] Y. L. Jiang and K. L. Xu, *Model order reduction of port-Hamiltonian systems by Riemannian modified Fletcher-Reeves scheme*, IEEE Trans. Circuits Syst. II, 66 (2019), pp. 1825–1829.

[32] Y. L. Jiang and K. L. Xu, *$\mathcal{H}_2$ optimal reduced models of general MIMO LTI systems via the cross Gramian on the Stiefel manifold*, J. Franklin Inst., 354 (2017), pp. 3210–3224.

[33] H. Sato and K. Sato, *Structure preserving $\mathcal{H}_2$ optimal model reduction based on Riemannian trust-region method*, IEEE Trans. Automat. Control, 63 (2018), pp. 505–512.

[34] T. Zeng and C. Lu, *Two-sided Grassmann manifold algorithm for optimal $\mathcal{H}_2$ model reduction*, Internat. J. Numer. Methods Engrg., 104 (2015), pp. 928–943.

[35] D. Amsallem, J. Cortial, K. Carlberg, and C. Farhat, *A method for interpolating on manifolds structural dynamics reduced-order models*, Internat. J. Numer. Methods Engrg., 80 (2009), pp. 1241–1258.

[36] P. Yang, Y. L. Jiang, and K. L. Xu, *A trust-region method for $\mathcal{H}_2$ model reduction of bilinear systems on the Stiefel manifold*, J. Franklin Inst., 356 (2019), pp. 2258–2273.

[37] W. Huang, P. A. Absil, and K. A. Gallivan, *A Riemannian BFGS method without differentiated retraction for nonconvex optimization problems*, SIAM J. Optim., 28 (2018), pp. 470–495.

[38] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.

[39] Y. L. Jiang and X. Kong, *On the uniqueness and perturbation to the best rank-one approximation of a tensor*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 775–792.

[40] P. Benner and P. Goyal, *Balanced Truncation Model Order Reduction for Quadratic-Bilinear Control Systems*, https://arxiv.org//abs/1705.00160, 2017.
[41] L. Zhang, W. Zhou, and D. H. Li, *A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence*, IMA. J. Numer. Ananl., 26 (2006), pp. 629–640.
[42] W. Huang, *Optimization Algorithms on Riemannian Manifolds with Applications*, Ph.D. thesis, Department of Mathematics, Florida State University, 2013.
[43] H. Sato and T. Iwai, *A new, globally convergent Riemannian conjugate gradient method*, Optimization, 64 (2015), pp. 1011–1031.
[44] Y. L. Jiang, *New Methods in Engineering Mathematics*, Higher Education Press, Beijing, 2013 (in Chinese).
[45] V. Simoncini, *Computational methods for linear matrix equations*, SIAM Rev., 58 (2016), pp. 377–441.
[46] G. Flagg and S. Gugercin, *On the ADI method for the Sylvester equation and the optimal $\mathcal{H}_2$ points*, Appl. Numer. Math., 64 (2013), pp. 50–58.
[47] J. L. Lions, Y. Maday, and G. Turinici, *A "parareal" in time discretization of PDE's*, C. R. Acad. Sci. Ser. I Math., 332 (2001), pp. 661-668.
[48] M. J. Gander, Y. L. Jiang, and B. Song, *A superlinear convergence estimate for the parareal Schwarz waveform relaxation algorithm*, SIAM J. Sci. Comput., 41 (2019), pp. A1148–A1169.
[49] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, PA, 1996.