Numer Algor (2017) 75:285–304
DOI 10.1007/s11075-016-0207-7

ORIGINAL PAPER

On some properties of the extended block and global Arnoldi methods with applications to model reduction

O. Abidi[1] · M. Heyouni[2] · K. Jbilou[1]

Received: 23 November 2015 / Accepted: 11 September 2016 / Published online: 22 September 2016

Abstract The aim of the present paper is to give some new algebraic properties of the extended block and the extended global Arnoldi algorithms. These results are then applied on moment matching methods for model reductions in large-scale dynamical systems to get low-order models that approximate the original models by matching moments and Markov parameters at the same time. Some numerical examples are given to show the effectiveness of the methods on some benchmark tests.

Keywords Extended block Arnoldi · Extended global Arnoldi · Low order · Model reduction · Transfer functions

Mathematics Subject Classification (2010) MSC 65F · MSC 15A

1 Introduction

The extended Arnoldi method was first proposed by Druskin and Knizhnerman in [3] for functions of matrices in the symmetric, large, and sparse case. The method was

✉ K. Jbilou
jbilou@lmpa.univ-littoral.fr

O. Abidi
abidi@lmpa.univ-littoral.fr

M. Heyouni
mohammed.heyouni@gmail.com

1 Université du Littoral, Côte d'Opale, batiment H. Poincarré, 50 rue F. Buisson, F-62280 Calais Cedex, France

2 Equipe EMMA, ENSA Al-Hoceima, UMP, Oujda, Morocco

then generalized to the nonsymmetric case by Simoncini in [16] and applied for solving large-scale Lyapunov matrix equations [13, 16] with low rank right-hand sides. In [11], the extended block Arnoldi method was used for computing approximate solutions to large-scale continuous-time algebraic Riccati equations while in [12] the extended global Arnoldi method was defined and used for solving large Sylvester matrix equations. If $A \in \mathbb{R}^{n\times n}$ is nonsingular, $v \in \mathbb{R}^n$ and m is a fixed integer, the classical extended Arnoldi Krylov subspace $K_m(A, v)$, is the subspace of $\mathbb{R}^n$ spanned by the vectors $A^{-m}\,v, \ldots, A^{-2}\,v, A^{-1}\,v, v, A\,v, A^2\,v, \ldots, A^{m-1}\,v$.
A convergence analysis of the extended Krylov subspace was recently developed in [13] where new general estimates for the convergence rate were obtained with real nonsymmetric and nonsingular matrices A.
For $V \in \mathbb{R}^{n\times r}$, the extended block Krylov subspace $\mathbb{K}_m(A, V)$ is the subspace of $\mathbb{R}^n$ spanned by the columns of the matrices $A^k\,V, k = -m, \ldots, m-1$. This subspace is denoted by

$$\mathbb{K}_m(A, V) = \text{colspan}(\{A^{-m}\,V, \ldots, A^{-2}\,V, A^{-1}\,V, V, A\,V, A^2\,V, \ldots, A^{m-1}\,V\}).$$

The subspace $\mathbb{K}_m(A, V)$ is the sum of the simple extended Krylov subspaces $K_m(A, V^{(i)})$, $i = 1, \ldots, r$ where $V^{(i)}$ is the ith column of the matrix V. Notice that $Z \in \mathbb{K}_m(A, V)$ means that

$$Z = \sum_{i=-m}^{m-1} A^i\,V\,\Omega_i, \text{ where } \Omega_i \in \mathbb{R}^{r\times r},\ i = -m, \ldots m-1.$$

On the other hand, the extended matrix or global Krylov subspace $\mathcal{K}_m(A, V) \subset \mathbb{R}^{n\times r}$ is the subspace of matrices in $\mathbb{R}^{n\times r}$ spanned by $A^k\,V$, $k = -m, \ldots, m-1$, i.e.,

$$\mathcal{K}_m(A, V) = span\{A^{-m}\,V, \ldots, A^{-2}\,V, A^{-1}\,V, V, A\,V, A^2\,V, \ldots, A^{m-1}\,V\},$$

and hence $Z \in \mathcal{K}_m(A, V)$ iff $Z = \sum_{i=-m}^{m-1} \alpha_i A^i\,V,\ \alpha_i \in \mathbb{R}$.
In this work, we give some new algebraic properties of the extended block and the extended global Arnoldi algorithms. The new properties use algebraic relations with the matrix A^{-1}. These new relations could be used in moment matching techniques for model reduction in large scale multiple input multiple output (MIMO) dynamical systems. In particular, we will show that some moments of the original transfer function are matched when using the approximated transfer function. The advantage of the extended block or global methods is the fact that they allow to approximate the original transfer functions by low order one for low and high frequencies at the same time.
The paper is organized as follow. In Section 2, we recall the extended block and global Arnoldi algorithms and give some new algebraic properties. The application of these methods to model order reduction is considered in Section 3. We show how to apply the extended block and global Arnoldi processes to dynamical MIMO sys-

tems to obtain low order models such that the Markov parameters and the moments of the original transfer function are approximated by the ones of the projected transfer function. The last section is devoted to some numerical experiments.

Preliminaries and notations We review some notations and definitions that will be used throughout this paper. For two matrices Y and Z in $\mathbb{R}^{n\times r}$, we define the Frobenius inner product $\langle Y, Z\rangle_F = \mathrm{Tr}(Y^T Z)$ where $\mathrm{Tr}(Y^T Z)$ denotes the trace of the square matrix $Y^T Z$. The associated Frobenius norm is given by $\|Y\|_F = \mathrm{Tr}(Y^T Y)^{\frac{1}{2}}$. A system $\{V_1, V_2, \ldots, V_m\}$ of elements of $\mathbb{R}^{n\times r}$ is said to be $F-$orthonormal if it is orthonormal with respect to the inner product $\langle .\,,.\rangle_F$, i.e., $\langle V_i, V_j\rangle_F = \delta_{i,j}$. For $Y \in \mathbb{R}^{n\times r}$, we denote by $\mathrm{vec}(Y)$ the vector of $\mathbb{R}^{nr}$ obtained by stacking the columns of Y. For two matrices A and B, $A \otimes B = [a_{i,j}B]$ denotes the Kronecker product of the matrices A and B. In the sequel, we give some properties of the Kronecker product assuming that all the sizes are in agreement.

1. $(A \otimes B)^T = A^T \otimes B^T$.
2. $(A \otimes B)(C \otimes D) = (AC \otimes BD)$.
3. If A and B are non singular matrices of size $n \times n$ and $p \times p$ respectively, then the $np \times np$ matrix $A \otimes B$ is non singular and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
4. $\mathrm{vec}(A)^T \mathrm{vec}(B) = \mathrm{Tr}(A^T B)$.

Definition 1 Let $A = [A_1, \ldots, A_q]$ and $B = [B_1, \ldots, B_l]$ be matrices of dimension $n \times qp$ and $n \times lp$, respectively, where A_i and B_j $(i = 1, \ldots, q;\; j = 1, \ldots, l)$ are $n \times p$. Then the $q \times l$ matrix $A^T \diamond B$ is defined by:

$$A^T \diamond B = [\langle A_i, B_j\rangle_F]_{1\le i\le q;\ 1\le j\le l}.$$

Remark 1 The following relations were established in [1].

1. The matrix $A = [A_1, \ldots, A_q]$ is $F-$orthonormal if and only if $A^T \diamond A = I_q$.
2. For all $X \in \mathbb{R}^{n\times p}$, we have $X^T \diamond X = \|X\|_F^2$.
3. $(DA)^T \diamond B = A^T \diamond (D^T B)$.
4. $A^T \diamond \big(B\,(L \otimes I_p)\big) = (A^T \diamond B)\,L$.
5. $\|A^T \diamond B\|_F \le \|A\|_F \|B\|_F$.

In the next proposition, we recall the global QR (gQR) factorisation of an $n \times kr$ matrix Z.

Proposition 1 [1] *Let* $\mathcal{Z} = \left[\, Z_1,\ Z_2,\ \ldots,\ Z_k \,\right]$ *be an* $n \times kr$ *matrix with* $Z_i \in \mathbb{R}^{n\times r}$, $i = 1, \ldots, k$. *Then, the matrix* $\mathcal{Z}$ *can be factored as*

$$\mathcal{Z} = \mathcal{Q}\,(R \otimes I_r),$$

where $\mathcal{Q} = [Q_1, \ldots, Q_k]$ *is an* $n \times kr$ *F-orthonormal matrix satisfying* $\mathcal{Q}^T \diamond \mathcal{Q} = I_k$ *and* R *is an upper triangular matrix of dimension* $k \times k$.

2 Some algebraic properties on the extended block and global Arnoldi processes

2.1 The block case

The extended block Arnoldi algorithm generates a sequence of blocks $\{V_1^b, \ldots, V_m^b\}$ of size $n \times 2r$ such that their columns form an orthonormal basis of the extended block Krylov subspace $\mathbb{K}_m(A, V)$. The algorithm is defined as follows [11, 16].

Algorithm 1 The extended block Arnoldi algorithm

- Inputs: $A \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{n\times r}$, m.
- Compute $[V_1^b, \Lambda] = QR([V, A^{-1}V])$, $\mathbb{V}_1 = [V_1^b]$.
- For $j = 1, \ldots, m$

  1. Set $V_j^{b(1)}$: first r columns of V_j^b; $V_j^{b(2)}$: second r columns of V_j^b.
  2. $\widetilde{V}_{j+1}^b = [A\, V_j^{b(1)}, A^{-1}\, V_j^{b(2)}]$.
  3. Orthogonalize $\widetilde{V}_{j+1}^b$ with respect to to $\mathbb{V}_1^b, \ldots, \mathbb{V}_j^b$ to get V_{j+1}^b, i.e.,
     for $i = 1, 2, \ldots, j$
       $H_{i,j}^b = (V_i)^{bT}\, \widetilde{V}_{j+1}^b$;
       $\widetilde{V}_{j+1}^b = \widetilde{V}_{j+1}^b - V_i^b\, H_{i,j}^b$;
     end for
  4. $[V_{j+1}^b,\ H_{j+1,j}^b] = QR(\widetilde{V}_{j+1}^b)$.
  5. $\mathbb{V}_{j+1} = [\mathbb{V}_j,\ V_{j+1}^b]$.

  End For.

The blocks $\mathbb{V}_m = \left[V_1^b, V_2^b, \ldots, V_m^b\right]$ with $V_i^b \in \mathbb{R}^{n\times 2r}$ have their columns mutually orthogonal provided that none of the upper triangular matrices $H_{j+1,j}^b$ are rank deficient.

Hence, after m steps, Algorithm 1 builds an orthonormal basis $\mathbb{V}_m$ of the extended block Krylov subspace $\mathbb{K}_m(A, V)$ and a upper block Hessenberg matrix $\mathbb{H}_m$ whose non zero blocks are the $H_{i,j}^b$. Note that each submatrix $H_{i,j}^b$ $(1 \le i \le j \le m)$ is of order $2r$.

Let $T_{i,j}^b = (V_i^b)^T\, A\, V_j^b \in \mathbb{R}^{2r\times 2r}$ and $\mathbb{T}_m = [T_{i,j}^b] \in \mathbb{R}^{2mr\times 2mr}$ be the restriction of the matrix A to the extended Krylov subspace $\mathbb{K}_m(A, V)$, i.e.,

$$\mathbb{T}_m = \mathbb{V}_m^T\, A\, \mathbb{V}_m.$$

It is shown in [16] that $\mathbb{T}_m$ is also upper block Hessenberg with $2r \times 2r$ blocks. Moreover, a recursion is derived to compute $\mathbb{T}_m$ from $\mathbb{H}_m$ without requiring matrix-vector products with A. For more details, on how to compute $\mathbb{T}_m$ from $\mathbb{H}_m$, we refer to [16]. We note that for large problems, the inverse of the matrix A is not computed explicitly and in this case we can use iterative solvers with preconditioners to solve linear systems with A. Next, we give the extended block Arnoldi relations

$$\begin{aligned} A\,\mathbb{V}_m &= \mathbb{V}_{m+1}\,\overline{\mathbb{T}}_m, \\ &= \mathbb{V}_m\,\mathbb{T}_m + V_{m+1}^b\,T_{m+1,m}^b\,\mathbb{E}_m^T, \end{aligned}$$

where $\overline{\mathbb{T}}_m = \mathbb{V}_{m+1}^T\,A\,\mathbb{V}_m$, and $\mathbb{E}_m = [O_{2r\times 2(m-1)r}, I_{2r}]^T$ is the matrix of the last $2r$ columns of the $2mr \times 2mr$ identity matrix I_{2mr}. We will also consider the matrix defined as

$$\mathbb{L}_m = \mathbb{V}_m^T\,A^{-1}\,\mathbb{V}_m.$$

Notice that we can check that the matrix $\mathbb{L}_m = [L_{i,j}^b]$ is also an upper block Hessenberg matrix. Moreover, the sub-matrices $L_{i+1,i}^b \in \mathbb{R}^{2r\times 2r}$ are such that the r first columns are zero. Hence, $L_{m+1,m}^b$ is partitioned under the form

$$L_{m+1,m}^b = \begin{bmatrix} 0_r & L_{m+1,m}^{b(1,2)} \\ 0_r & L_{m+1,m}^{b(2,2)} \end{bmatrix}. \quad (1)$$

In the sequel, we give some new properties that would be useful for building model order reduction for large-scale dynamical systems defined by (19).

Proposition 2 *Assume that m steps of Algorithm 1 have been run and let* $\overline{\mathbb{L}}_m = \mathbb{V}_{m+1}^T\,A^{-1}\,\mathbb{V}_m$, *then we have the following relations*

$$\begin{aligned} A^{-1}\,\mathbb{V}_m &= \mathbb{V}_{m+1}\,\overline{\mathbb{L}}_m \\ &= \mathbb{V}_m\,\mathbb{L}_m + V_{m+1}^b\,L_{m+1,m}^b\,\mathbb{E}_m^T. \end{aligned} \quad (2)$$

Proof As $\mathbb{V}_{m+1} = [\mathbb{V}_m, V_{m+1}^b]$, we have

$$\begin{aligned} \mathbb{L}_{m+1} &= \mathbb{V}_{m+1}^T\,A^{-1}\,\mathbb{V}_{m+1} \\ &= \begin{bmatrix} \mathbb{V}_m^T\,A^{-1}\,\mathbb{V}_m & \mathbb{V}_m^T\,A^{-1}\,V_{m+1}^b \\ V_{m+1}^{bT}\,A^{-1}\,\mathbb{V}_m & V_{m+1}^{bT}\,A^{-1}\,V_{m+1}^b \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{L}_m & \mathbb{V}_m^T\,A^{-1}\,V_{m+1}^b \\ V_{m+1}^{bT}\,A^{-1}\,\mathbb{V}_m & V_{m+1}^{bT}\,A^{-1}\,V_{m+1}^b \end{bmatrix}. \end{aligned}$$

Now, since $\mathbb{L}_{m+1}$ is an upper block Hessenberg matrix, we also have

$$(V_{m+1}^b)^T\,A^{-1}\,\mathbb{V}_m = L_{m+1,m}^b\,\mathbb{E}_m^T,$$

and so the upper block Hessenberg matrix $\overline{\mathbb{L}}_m$ can be expressed as

$$\overline{\mathbb{L}}_m = \begin{bmatrix} \mathbb{L}_m \\ L_{m+1,m}^b\,\mathbb{E}_m^T \end{bmatrix}.$$

Using the fact that $A^{-1}\,\mathbb{K}_m(A, V) \subseteq \mathbb{K}_{m+1}(A, V)$ and $\mathbb{V}_{m+1}$ is orthogonal, it follows that there exists an upper block Hessenberg matrix L such that $A^{-1}\,\mathbb{V}_m = \mathbb{V}_{m+1}\,L$. Then, $\mathbb{V}_{m+1}^T\,A^{-1}\,\mathbb{V}_m = L$, which shows that $L = \overline{\mathbb{L}}_m$. Hence, we obtain

$$A^{-1}\,\mathbb{V}_m = \mathbb{V}_{m+1}\,L = \mathbb{V}_{m+1}\,\overline{\mathbb{L}}_m = \mathbb{V}_m\,\mathbb{L}_m + V_{m+1}^b\,L_{m+1,m}^b\,\mathbb{E}_m^T,$$

which completes the proof. □

Now using (1) and the fact that $V^b_{m+1} = [V^{b(1)}_{m+1}, V^{b(2)}_{m+1}]$, the relation (2) becomes

$$A^{-1}\,\mathbb{V}_m = \mathbb{V}_m\,\mathbb{L}_m + [O_{n\times(2m-1)r}, V^{b(1)}_{m+1}\,L^{b(1,2)}_{m+1,m} + V^{b(2)}_{m+1}\,L^{b(2,2)}_{m+1,m}].$$

Next, to show how to compute the columns of the matrix $\overline{\mathbb{L}}_m$ without using A^{-1}, we have to give some notations:

– Let $[V,\ A^{-1}V] = V^b_1\Lambda$ be the QR decomposition of $[V,\ A^{-1}V]$ which can be written as

$$[V, A^{-1}V] = V^b_1\Lambda = [V^{b(1)}_1,\ V^{b(2)}_1]\begin{bmatrix}\Lambda_{1,1} & \Lambda_{1,2}\\ 0 & \Lambda_{2,2}\end{bmatrix}. \tag{3}$$

– For $k = 1,\dots,m$, let us partition the lower triangular matrix $H^b_{k+1,k}$ under the form

$$H^b_{k+1,k} = \begin{bmatrix} H^{b(1,1)}_{k+1,k} & H^{b(1,2)}_{k+1,k}\\ 0 & H^{b(2,2)}_{k+1,k}\end{bmatrix}.$$

The following result enables us to compute $\overline{\mathbb{L}}_m$ directly from the columns of the upper block Hessenberg matrix $\overline{\mathbb{H}}_m$ obtained from Algorithm 1.

Proposition 3 *Let $\overline{\mathbb{L}}_m$ and $\overline{\mathbb{H}}_m$ be the upper block Hessenberg matrices defined earlier. Then we have the following relations*

$$\overline{\mathbb{L}}_m\,\widetilde{e}_1 = \big[\widetilde{e}_1\,\Lambda_{1,2} + \widetilde{e}_1\,\Lambda_{2,2}\big](\Lambda_{1,1})^{-1}, \tag{4}$$

and for $k = 1,\dots,m$

$$\overline{\mathbb{L}}_m\,\widetilde{e}_{2k} = \overline{\mathbb{H}}_m\,\widetilde{e}_{2k}, \tag{5}$$

and

$$\mathbb{L}_m\,\widetilde{e}_{2k+1} = \left(\widetilde{e}_{2k-1} - \begin{bmatrix}\overline{\mathbb{L}}_k\\ 0_{2(m-k)r\times 2kr}\end{bmatrix}\mathbb{H}_k\,\widetilde{e}_{2k-1}\right)(H^{b(1,1)}_{k+1,k})^{-1}, \tag{6}$$

where $\widetilde{e}_i = e_i \otimes I_r$ and the e_i's are the vectors of the canonical basis.

Proof To prove (4), we start from the QR decomposition of $[V,\ A^{-1}V]$ given in (3):

$$[V,\ A^{-1}V] = [V^{b(1)}_1\,\Lambda_{1,1}, V^{b(1)}_1\,\Lambda_{1,2} + V^{b(2)}_1\,\Lambda_{2,2}].$$

Then, if $\Lambda_{1,1}$ is nonsingular, we obtain

$$A^{-1}\,V^{b(1)}_1 = A^{-1}\,V\,\Lambda^{-1}_{1,1} = [V^{b(1)}_1\,\Lambda_{1,2} + V^{(2)}_1\,\Lambda_{2,2}]\,\Lambda^{-1}_{1,1}.$$

Then, we get (4) by pre-multiplying the above equality on the left by $\mathbb{V}^T_{m+1}$ and using the facts that $\mathbb{V}^T_{m+1}\,V^{b(i)}_1 = (e_i \otimes I_r) = \widetilde{e}_i$ for $i = 1,\ 2$ and $\mathbb{V}^T_{m+1}\,A^{-1}\,V^{b(1)}_1 = \overline{\mathbb{L}}_m\,(e_1 \otimes I_r) = \overline{\mathbb{L}}_m\,\widetilde{e}_1$.

To prove (5) and (6), we notice that for $k \geq 1$, $V_k = [V^{b(1)}_k, V^{b(2)}_k] \in \mathbb{R}^{n\times 2r}$ and from Algorithm 1, we have

$$\widehat{V}^b_{k+1} = [A\,V^{b(1)}_k, A^{-1}\,V^{b(2)}_k] - \mathbb{V}_k\,\mathbb{H}_k\,[\widetilde{e}_{2k-1}, \widetilde{e}_{2k}], \tag{7}$$

and

$$V_{k+1}^{b}\, H_{k+1,k}^{b} = \widehat{V}_{k+1}^{b}. \tag{8}$$

Using the relations (7) and (8), we obtain

$$\begin{aligned} A^{-1}\, V_{k}^{b(2)} = \widehat{V}_{k+1}^{b}\, \widetilde{e}_2 + \mathbb{V}_k\, \mathbb{H}_k\, \widetilde{e}_{2k} &= V_{k+1}^{b}\, H_{k+1,k}^{b}\, \widetilde{e}_2 + \mathbb{V}_k\, \mathbb{H}_k\, \widetilde{e}_{2k} \\ &= \mathbb{V}_{k+1} \overline{\mathbb{H}}_k\, \widetilde{e}_{2k}. \end{aligned}$$

Now, multiplying on the left by $\mathbb{V}_{m+1}^{T}$, we get

$$\mathbb{V}_{m+1}^{T}\, A^{-1}\, V_{k}^{b(2)} = \mathbb{V}_{m+1}^{T}\, \mathbb{V}_{k+1}\, \overline{\mathbb{H}}_k\, \widetilde{e}_{2k},$$

hence,

$$\mathbb{V}_{m+1}^{T}\, A^{-1}\, \mathbb{V}_m\, \widetilde{e}_{2k} = \begin{bmatrix} I_{2(k+1)r} \\ 0_{2(m-k)r \times 2(k+1)r} \end{bmatrix} \overline{\mathbb{H}}_k\, \widetilde{e}_{2k}$$

and so

$$\overline{\mathbb{L}}_m\, \widetilde{e}_{2k} = \begin{bmatrix} \overline{\mathbb{H}}_k \\ 0_{2(m-k)r \times 2kr} \end{bmatrix} = \overline{\mathbb{H}}_m\, \widetilde{e}_{2k},$$

which gives the relation (5).
Now, for the even blocks, we multiply (7) on the left by A^{-1} and we consider only the first r-columns of each block. We obtain the following relation

$$A^{-1}\, \widehat{V}_{k+1}^{b(1)} = V_{k}^{b(1)} - A^{-1}\, \mathbb{V}_k\, \mathbb{H}_k\, \widetilde{e}_{2k-1}.$$

Notice that since $\widehat{V}_{k+1}^{b} = V_{k+1}^{b}\, H_{k+1,k}^{b}$, we also have

$$\widehat{V}_{k+1}^{b(1)} = V_{k+1}^{b(1)}\, H_{k+1,k}^{b(1,1)},$$

where $H_{k+1,k}^{b(1,1)}$ is the first $r \times r$ block of the upper $2r \times 2r$ triangular matrix $H_{k+1,k}^{b}$. Then, if $H_{k+1,k}^{b(1,1)}$ is nonsingular, we obtain

$$A^{-1}\, V_{k+1}^{b(1)} = A^{-1}\, \widehat{V}_{k+1}^{b(1)}\, (H_{k+1,k}^{b(1,1)})^{-1} = \big(V_{k}^{b(1)} - A^{-1}\, \mathbb{V}_k\, \mathbb{H}_k\, \widetilde{e}_{2k-1}\big)(H_{k+1,k}^{b(1,1)})^{-1}.$$

Multiplying from the left by $\mathbb{V}_{m+1}^{T}$, we get

$$\mathbb{V}_{m+1}^{T}\, A^{-1}\, V_{k+1}^{b(1)} = \left(\mathbb{V}_{m+1}^{T}\, V_{k}^{b(1)} - \mathbb{V}_{m+1}^{T}\, A^{-1}\, \mathbb{V}_k\, \mathbb{H}_k\, \widetilde{e}_{2k-1}\right)\, (H_{k+1,k}^{b(1,1)})^{-1},$$

and then

$$\begin{aligned} \mathbb{L}_{m+1}\, \widetilde{e}_{2k+1} &= \left(\mathbb{V}_{m+1}^{T}\, \mathbb{V}_{m+1}\, \widetilde{e}_{2k-1} - \mathbb{V}_{m+1}^{T}\, A^{-1}\, \mathbb{V}_m \begin{bmatrix} I_{2kr} \\ 0_{2(m-k)r \times 2kr} \end{bmatrix} \mathbb{H}_k\, \widetilde{e}_{2k-1}\right)\, (H_{k+1,k}^{b(1,1)})^{-1} \\ &= \left(\widetilde{e}_{2k-1} - \overline{\mathbb{L}}_m \begin{bmatrix} I_{2kr} \\ 0_{2(m-k)r \times 2kr} \end{bmatrix} \mathbb{H}_k\, \widetilde{e}_{2k-1}\right)\, (H_{k+1,k}^{b(1,1)})^{-1} \\ &= \left(\widetilde{e}_{2k-1} - \begin{bmatrix} \overline{\mathbb{L}}_k \\ 0_{2(m-k)r \times 2kr} \end{bmatrix} \mathbb{H}_k\, \widetilde{e}_{2k-1}\right)\, (H_{k+1,k}^{b(1,1)})^{-1}, \end{aligned}$$

which gives the second relation (6). □

2.2 The global case

The extended global Arnoldi process was first described in [12]. The algorithm is summarized as follows

Algorithm 2 The extended global Arnoldi algorithm

- Inputs: $A \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{n\times r}$, m.
- Compute $[V_1^g, \Omega] = gQR([V, A^{-1}V])$, (global QR decomposition) $\mathcal{V}_1 = \left[V_1^g\right]$.
- For $j = 1, \ldots, m$

  1. Set $V_j^{g(1)}$: first r columns of V_j^g; $V_j^{g(2)}$: second r columns of V_j^g.
  2. $\widetilde{V}_{j+1}^g = [A\, V_j^{g(1)}, A^{-1}\, V_j^{g(2)}]$.
  3. F-Orthogonalize $\widetilde{V}_{j+1}^g$ with respect to $\mathcal{V}_1^g, \ldots, \mathcal{V}_j^g$ to get V_{j+1}^g, i.e.,
     for $i = 1, 2, \ldots, j$
     $H_{i,j}^g = (V_i)^{gT} \diamond \widetilde{V}_{j+1}^g$;
     $\widetilde{V}_{j+1}^g = \widetilde{V}_{j+1}^g - V_i^g\,(H_{i,j}^g \otimes I_r)$;
     end for
  4. $[V_{j+1}^g,\ H_{j+1,j}^g] = gQR(\widetilde{V}_{j+1}^g)$.
  5. $\mathcal{V}_{j+1} = [\mathcal{V}_j,\ V_{j+1}^g]$.

  End For.

We point out that if the upper 2×2 triangular matrices $H_{j+1,j}^g$ $(j = 1, \ldots, m)$ are full rank, Algorithm 2 constructs an $n \times 2mr$ F-orthonormal matrix $\mathcal{V}_m = [V_1^g, \ldots, V_m^g]$ with $V_i^g \in \mathbb{R}^{n\times 2r}$ $(i = 1, \ldots, m)$ and a $2(m+1) \times 2m$ upper block Hessenberg matrix $\overline{\mathcal{H}}_m = \left[H_{i,j}^g\right] = [h_{p,q}]$ with $H_{i,j}^g \in \mathbb{R}^{2\times 2}$ for $i = 1, \ldots, m+1$, $j = 1, \ldots, m$ and $h_{p,q} \in \mathbb{R}$ for $p = 1, \ldots, 2(m+1)$, $q = 1, \ldots, 2m$.

Now, setting $T_{i,j}^g = V_i^{gT} \diamond (A\, V_j^g) \in \mathbb{R}^{2\times 2}$, for $i, j = 1, \ldots, m$ and introducing the matrices

$$\mathcal{T}_m = \mathcal{V}_m^T \diamond (A\,\mathcal{V}_m) = \left[T_{i,j}^g\right] \quad \text{and} \quad \overline{\mathcal{T}}_m = \mathcal{V}_{m+1}^T \diamond (A\,\mathcal{V}_m),$$

a recursive relation was given in [12] allowing the computation of $\bar{\mathcal{T}}_m$ without requiring additional matrix-vector products with A. Moreover, it was also shown that after m steps of Algorithm 2, we have

$$\begin{aligned} A\,\mathcal{V}_m &= \mathcal{V}_{m+1}\,(\overline{\mathcal{T}}_m \otimes I_r) \\ &= \mathcal{V}_m\,(\mathcal{T}_m \otimes I_r) + V_{m+1}^g\,(T_{m+1,m}^g\,\mathcal{E}_m^T \otimes I_r), \end{aligned}$$

where $\mathcal{E}_m^T = [O_{2\times 2(m-1)}, I_2]$ is the matrix of the last two rows of the $2m \times 2m$ identity matrix I_{2m}.

Now, as for the block case seen in the previous subsection, we consider the matrix $\mathcal{L}_m = \left[L^g_{i,j}\right] = [l_{p,q}]$ defined by

$$\mathcal{L}_m = \mathcal{V}_m^T \diamond (A^{-1}\mathcal{V}_m),$$

where $L^g_{i,j} \in \mathbb{R}^{2\times 2}$ for $i, j = 1, \ldots, m$ and $l_{p,q} \in \mathbb{R}$ for $p, q = 1, \ldots, 2m$. We mention that it can be easily verified that the matrix $\mathcal{L}_m$ is a $2m \times 2m$ upper block Hessenberg matrix and that the sub-matrices $L^g_{i+1,i} \in \mathbb{R}^{2\times 2}$ are such that the first column is zero. So, the sub-matrix $L^g_{m+1,m}$ is such that $l_{2m+1,2m-2} = l_{2m+2,2m-2} = 0$, i.e.,

$$L^g_{m+1,m} = \begin{bmatrix} 0 & l_{2m+1,2m} \\ 0 & l_{2m+2,2m} \end{bmatrix}. \tag{9}$$

In order to build order model reductions for large-scale dynamical systems, we give some new properties of the extended global Arnoldi process.

Proposition 4 *Assume that m steps of Algorithm 2 have been run and let $\overline{\mathcal{L}}_m = \mathcal{V}_{m+1}^T A^{-1}\mathcal{V}_m$, then we have the following relations*

$$\begin{aligned} A^{-1}\mathcal{V}_m &= \mathcal{V}_{m+1}(\overline{\mathcal{L}}_m \otimes I_r) \\ &= \mathcal{V}_m(\mathcal{L}_m \otimes I_r) + V^g_{m+1}(L^g_{m+1,m}\mathcal{E}_m^T \otimes I_r). \end{aligned} \tag{10}$$

The proof is similar to the one given for Proposition 2.
Notice that since $V^g_{m+1} = [V^{g(1)}_{m+1}, V^{g(2)}_{m+1}]$, then by using (9), the Arnoldi relation (10) becomes

$$A^{-1}\mathcal{V}_m = \mathcal{V}_m \mathcal{L}_m + [O_{n\times(2m-1)}, \, l_{2m+1,2m} V^{g(1)}_{m+1} + l_{2m+2,2m} V^{g(2)}_{m+1}].$$

Now, in order to update progressively the columns of the matrix $\overline{\mathcal{L}}_m$ without inverting A or solving linear systems with A, we recall some elementary results:

- Let $[V, \; A^{-1}V] = V^g_1(\Gamma \otimes I_r)$ be the global QR decomposition of $[V, \; A^{-1}V]$ which can be written as

$$[V, A^{-1}V] = V^g_1(\Gamma \otimes I_r) = [V^{g(1)}_1, \; V^{g(2)}_1] \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} \\ 0 & \gamma_{2,2} \end{bmatrix}.$$

- For $k = 1, \ldots, m$, let us partition the lower triangular matrix $H^g_{k+1,k}$ under the form

$$H^g_{k+1,k} = \begin{bmatrix} h_{2k+1,2k-1} & h_{2k+1,2k} \\ 0 & h_{2k+2,2k} \end{bmatrix}.$$

As in the block case, the following result enables us to compute $\overline{\mathbb{L}}_m$ directly from the columns of the upper block Hessenberg matrix $\overline{\mathbb{H}}_m$ obtained from Algorithm 1.

Proposition 5 *Let $\overline{\mathcal{L}}_m = [l_{:,1}, \ldots, l_{:,2m}]$ and $\overline{\mathcal{H}}_m = [h_{:,1}, \ldots, h_{:,2m}]$ be the upper block Hessenberg matrices defined earlier. Then, we have the following relations*

$$l_{:,1} = (\gamma_{1,2}\, e_1 + \gamma_{2,2}\, e_2)/\gamma_{1,1},$$

and for $k = 1, \ldots, m$, *we have*

$$\bar{l}_{:,2k} = h_{:,2k} \tag{11}$$

and

$$l_{2k+1} = \left(e_{2k-1} - \begin{bmatrix} \overline{\mathbb{L}}_k \\ 0_{2(m-k)\times 2k} \end{bmatrix} \mathbb{H}_k \, e_{2k-1} \right) / \, h_{2k+1,2k},$$

where the e_i*'s are the vectors of the canonical basis.*

The proof can be obtained in a similar way as the one for Proposition 3 in the block case.
The results of the previous two subsections are used to prove other properties in the next section which is devoted to the application of the extended block and global Arnoldi methods to obtain reduced order models in large-scale dynamical systems. As we will see, the methods allow one to approximate low and high frequencies of the corresponding transfer function at the same time.

Next, we give some properties that are used to show that the first m moments of the transfer function F are matched.

Proposition 6 *Let* $\mathbb{V}_m = [V_1^b, \ldots, V_m^b]$, $\mathbb{L}_m := \mathbb{V}_m^T A^{-1} \mathbb{V}_m$ *be respectively the matrix and the upper block Hessenberg matrix defined in the extended block Arnoldi process and let* $\mathbb{E}_1 = [I_{2r}, 0_{2r}, \ldots, 0_{2r}]^T$. *Then for* $j = 0 \ldots, m-1$, *we have the following relation*

$$A^{-j} \, \mathbb{V}_m \, \mathbb{E}_1 = \mathbb{V}_m \, \mathbb{L}_m^j \, \mathbb{E}_1, \tag{12}$$

and

$$\mathbb{T}_m^{-1} \, \mathbb{E}_j = \mathbb{L}_m \, \mathbb{E}_j, \;\; j = 1, \ldots, m-1. \tag{13}$$

Proof The relation (12) can be derived directly by multiplying relation (2) from the left by A^{-j+1} and from the right by $\mathbb{E}_1$. Then we obtain

$$A^{-j} \, \mathbb{V}_m \, \mathbb{E}_1 = \mathbb{V}_m \, \mathbb{L}_m^j \, \mathbb{E}_1 + \sum_{i=1}^{j} A^{-(i-1)} \, V_{m+1}^b \, L_{m+1,m}^b \, \mathbb{E}_m^T \mathbb{L}_m^{j-i} \, \mathbb{E}_1.$$

Now, as $\mathbb{L}_m$ is an upper block Hessenberg matrix, it follows that $\mathbb{E}_m^T \, \mathbb{L}_m^{j-i} \, \mathbb{E}_1 = 0$, for $j = 1, \ldots, m-1$.
To prove the relation (13), we multiply (12) from the right by $\mathbb{E}_j$, and then we get

$$A^{-1} \, \mathbb{V}_m \, \mathbb{E}_j = \mathbb{V}_m \, \mathbb{L}_m \, \mathbb{E}_j, \quad \text{for} \quad j = 1, \ldots, m-1.$$

Since, $\mathbb{V}_m$ is orthogonal, then pre-multiplying the above equality by $\mathbb{V}_m^T A$, we get $\mathbb{E}_j = \mathbb{T}_m \, \mathbb{L}_m \, \mathbb{E}_j$ for $j = 1, \ldots, m-1$ and finally we obtain (13), if we assume that $\mathbb{T}_m$ is nonsingular. □

Next, we establish a similar result for the extended global Arnoldi process.

Proposition 7 *Let* $\mathcal{V}_m = [V_1^g, \ldots, V_m^g]$, $\mathcal{L}_m := \mathcal{V}_m^T \diamond (A^{-1} \, \mathcal{V}_m)$ *be respectively the matrix and the upper block Hessenberg matrix defined by the extended global Arnoldi* *process and let* $\mathcal{E}_1 = [I_2, 0_2, \ldots, 0_2]^T$. *Then for* $j = 1\ldots, m-1$, *we have the following relation*

$$A^{-j}\,\mathcal{V}_m\,\mathbb{E}_1 = A^{-j}\,\mathcal{V}_m\,(\mathcal{E}_1 \otimes I_r) = \mathcal{V}_m\,(\mathcal{L}_m^j\,\mathcal{E}_1 \otimes I_r), \tag{14}$$

and for $j = 1, \ldots, m-1$ *we have*

$$\mathcal{T}_m^{-1}\,\mathcal{E}_j = \mathcal{L}_m\,\mathcal{E}_j. \tag{15}$$

Proof Pre-multiplying (10) by A^{-j+1} and using the properties of the Kronecker product, we get

$$A^{-j}\,\mathcal{V}_m = \mathcal{V}_m\,(\mathcal{L}_m^j \otimes I_r) + \sum_{i=1}^{j} A^{-(i-1)}\,V_{m+1}^g\,(L_{m+1,m}^g\,\mathcal{E}_m^T\,\mathcal{L}_m^{j-i} \otimes I_r).$$

Post-multiplying the above equality by $(\mathcal{E}_1 \otimes I_r)$, we obtain

$$\begin{aligned} A^{-j}\,\mathcal{V}_m\,\mathbb{E}_1 &= A^{-j}\,\mathcal{V}_m\,(\mathcal{E}_1 \otimes I_r) \\ &= \mathcal{V}_m\,(\mathcal{L}_m^j\ \mathcal{E}_1 \otimes I_r) + \sum_{i=1}^{j} A^{-(i-1)}\,V_{m+1}^g\,(L_{m+1,m}^g\,\mathcal{E}_m^T\,\mathcal{L}_m^{j-i}\,\mathcal{E}_1 \otimes I_r), \\ &= \mathcal{V}_m\,(\mathcal{L}_m^j\ \mathcal{E}_1 \otimes I_r). \end{aligned}$$

In the last equality, we used the fact that $\mathcal{E}_m^T\,\mathcal{L}_m^{j-i}\,\mathcal{E}_1 = 0$ since $\mathcal{L}_m$ is an upper block Hessenberg matrix. Using again the properties of the $\otimes$ and $\diamond$ products, the proof of the second relation (15) can be derived in a similar fashion to that of (13). □

Next, we give a general result that is satisfied by upper Hessenberg matrices. This result will be used when establishing moment matching properties for the block and global Arnoldi processes.

Proposition 8 *Let* $T = (T_{i,j})$ *and* $L = (L_{i,j})$ *be two upper block Hessenberg matrices with blocks* $T_{i,j},\ L_{i,j} \in \mathbb{R}^{r\times r}$ *for* $i, j = 1, \ldots, m$ *and suppose that*

$$T\,\mathbb{E}_j = L\,\mathbb{E}_j, \quad \textit{for} \quad j = 1, \ldots, m-1 \tag{16}$$

where $\mathbb{E}_j = [0_r, \ldots, 0_r, I_r, 0_r, \ldots, 0_r]^T$, *is the* $mr \times r$ *matrix whose columns are the column* $j, \ldots, j+r$ *of the identity matrix* I_{mr}. *Then*

$$T^k\,\mathbb{E}_1 = L^k\,\mathbb{E}_1, \ \textit{for} \quad k = 1, \ldots, m-1. \tag{17}$$

Proof For $k = 1, \ldots, m-1$, we denote by $T_{i,j}^{(k)}$ and $L_{i,j}^{(k)}$ the (i, j)th block of T^k, L^k respectively, i.e.,

$$T^k = \left(T_{i,j}^{(k)}\right) \quad \text{and} \quad L^k = \left(L_{i,j}^{(k)}\right).$$

Since T and L are upper block Hessenberg matrices, we can easily verify that

$$T_{i,j}^{(k)} = L_{i,j}^{(k)} = 0_r, \ \text{for} \quad i > j+k. \tag{18}$$

Now, we proceed by induction on $k \in \{1, 2, \ldots, m-1\}$.

We first remark that the property is verified for $k = 1$ and we suppose that the property is valid for $k \in \{1, \ldots, m-2\}$.
Obviously, we have

$$T^{k+1}\,\mathbb{E}_1 = T\left(T^k\,\mathbb{E}_1\right) = T\left(L^k\,\mathbb{E}_1\right) = \begin{pmatrix} \sum_{p=1}^{m} T^{(1)}_{1,p}\,L^{(k)}_{p,1} \\ \sum_{p=1}^{m} T^{(1)}_{2,p}\,L^{(k)}_{p,1} \\ \vdots \\ \sum_{p=1}^{m} T^{(1)}_{m,p}\,L^{(k)}_{p,1} \end{pmatrix},$$

and thanks to (18) and (16), we have $L^{(k)}_{p,1} = 0_r$ for $p = m$ and $T^{(1)}_{i,p} = L^{(1)}_{i,p}$ for $i = 1, \ldots, m$ and $p = 1, \ldots, m-1$. So,

$$T^{k+1}\,\mathbb{E}_1 = \begin{pmatrix} \sum_{p=1}^{m-1} T^{(1)}_{1,p}\,L^{(k)}_{p,1} \\ \sum_{p=1}^{m-1} T^{(1)}_{2,p}\,L^{(k)}_{p,1} \\ \vdots \\ \sum_{p=1}^{m-1} T^{(1)}_{m,p}\,L^{(k)}_{p,1} \end{pmatrix} = \begin{pmatrix} \sum_{p=1}^{m-1} L^{(1)}_{1,p}\,L^{(k)}_{p,1} \\ \sum_{p=1}^{m-1} L^{(1)}_{2,p}\,L^{(k)}_{p,1} \\ \vdots \\ \sum_{p=1}^{m-1} L^{(1)}_{m,p}\,L^{(k)}_{p,1} \end{pmatrix},$$

and since $L^{(k)}_{p,1} = 0_r$ for $p = m$, we finally get

$$T^{k+1}\,\mathbb{E}_1 = \begin{pmatrix} \sum_{p=1}^{m} L^{(1)}_{1,p}\,L^{(k)}_{p,1} \\ \sum_{p=1}^{m} L^{(1)}_{2,p}\,L^{(k)}_{p,1} \\ \vdots \\ \sum_{p=1}^{m} L^{(1)}_{m,p}\,L^{(k)}_{p,1} \end{pmatrix} = L^{k+1}\,\mathbb{E}_1.$$

□

3 Application for model reduction techniques

We consider the following Linear Time Independent (LTI) dynamical system

$$\begin{cases} \dot{x}(t) = A\,x(t) + B\,u(t), \\ y(t) = C\,x(t), \end{cases} \tag{19}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t),\ y(t) \in \mathbb{R}^r$ are the input and the output vectors of the system (19), respectively. The matrices $B,\ C^T$ are in $\mathbb{R}^{n\times r}$ and $A \in \mathbb{R}^{n\times n}$ is assumed to be large and sparse. The transfer function of the original system (19) is given as

$$F(s) = C\,(s\,I_n - A)^{-1}\,B. \tag{20}$$

In many applications, the dimension n of the system (19) is large which makes the computations infeasible in terms of execution time and memory. Then the goal of model reduction problems is to produce a low-order system of the form

$$\begin{cases} \dot{x}_m(t) = A_m\,x_m(t) + B_m\,u(t) \\ y_m(t) = C_m\,x_m(t), \end{cases} \tag{21}$$

where $A_m \in \mathbb{R}^{p\times p}$, $B_m,\ C_m^T \in \mathbb{R}^{p\times r}$. The basic technique is to project the system's state space of dimension n onto a space of lower dimension $p \ll n$, in such a way that the reduced-order model preserves the important properties of the original system like stability and passivity and such that the output y_m is close to the output y of the original system. The associated low-order transfer function is denoted by

$$F_m(s) = C_m\,(s\,I_p - A_m)^{-1}\,B_m.$$

There are two well known sets of model reduction methods for MIMO systems which are currently in use, SVD-based methods and Krylov (moment matching) based methods; see [7, 8] and the references therein. One of the most common approach of the first category is the so-called balanced reduced order model which was introduced by Moore [15]. Krylov subspace methods have been extensively used for SISO (the case $r = 1$) and MIMO dynamical systems; see [2–6, 9, 10, 17] and the references therein. Unfortunately, the standard version of these methods builds reduced order models that poorly approximate low and high frequency dynamics at the same time. In order to address this problem, we consider the extended Arnoldi process associated to the matrices A and A^{-1}. The transfer function F relates the Laplace transform of the output vector to that of the input vector. For that reason, it is called the transfer function matrix of the system. Each entry $F_{i,j}(s)$ is a rational function representing the transfer function between the ith input and the jth output, all other inputs being set equal to zero.
The rational function F can be expressed as a sum of a Taylor series around ($s = \infty$) in the following form

$$F(s) = \frac{1}{s}\,C\,(I_n - \frac{A}{s})^{-1}\,B = \frac{1}{s}\sum_{i=0}^{\infty} M_i\,s^{-i},\ \text{with}\ \ M_i = C\,A^i\,B.$$

Recall that the matrix coefficients M_i are called the Markov parameters of F. Now applying the extended block Arnoldi process to the pair (A, B), we can verify that the original transfer function F can be approximated by

$$\mathbb{F}_m(s) = \mathbb{C}_m\,(s\,I_{2mr} - \mathbb{T}_m)^{-1}\,\mathbb{B}_m,$$

where $\mathbb{T}_m = \mathbb{V}_m^T\,A\,\mathbb{V}_m$, $\mathbb{C}_m = C\,\mathbb{V}_m$, and $\mathbb{B}_m = \mathbb{V}_m^T\,B$. This reduced transfer function is related to the low-order dynamical system (21) with $A_m = \mathbb{T}_m$.
Similarly, if m iterations of the extended global Arnoldi algorithm are applied to the pair (A, B), then we can approximate F by

$$\mathcal{F}_m(s) = \mathcal{C}_m\,(s\,I_{2mr} - (\mathcal{T}_m \otimes I_r))^{-1}\,\mathcal{B}_m,$$

where $\mathcal{T}_m = \mathcal{V}_m^T \diamond (A\,\mathcal{V}_m)$, $\mathcal{C}_m = C\,\mathcal{V}_m$ and $\mathcal{B}_m = \mathcal{V}_m^T \diamond B = \|B\|_F\,(e_1^{(2m)} \otimes I_r)$. In this case, the reduced transfer function is related to the low-order dynamical system (21) with $A_m = \mathcal{T}_m \otimes I_r$.
The developments of $\mathbb{F}_m$ and $\mathcal{F}_m$ around $s = \infty$ give the following expressions

$$\mathbb{F}_m(s) = \frac{1}{s}\sum_{i=0}^{\infty} m_i^b\,s^{-i}, \quad \text{with } m_i^b = \mathbb{C}_m\,\mathbb{T}_m^i\,\mathbb{B}_m,$$

and

$$\mathcal{F}_m(s) = \frac{1}{s}\sum_{i=0}^{\infty} m_i^g\,s^{-i}, \quad \text{with } m_i^g = \mathcal{C}_m\,(\mathcal{T}_m \otimes I_r)^i\,\mathcal{B}_m.$$

In this case, one can show that the first m Markov parameters are matched, i.e., in the block case

$$M_i = m_i^b, \quad i = 0, \ldots, m-1,$$

and in the global case

$$M_i = m_i^g, \quad i = 0, \ldots, m-1.$$

Now, the development of the Neumann series of F around $s = 0$ gives the following expression

$$F(s) = \sum_{i=0}^{\infty} \widetilde{M}_{i+1}\,s^i.$$

The matrix coefficients $\widetilde{M}_i$ are called the moments of F and they are given by

$$\widetilde{M}_j = -C\,A^{-j}\,B, \quad j = 1, 2, \ldots.$$

By considering the Taylor series of $\mathbb{F}_m$ and $\mathcal{F}_m$, we get the following expansion of $\mathbb{F}_m$ around $s = 0$

$$\mathbb{F}_m(s) = \sum_{i=0}^{\infty} \widetilde{m}_{i+1}^b s^i, \ \text{with } \widetilde{m}_i^b = -\mathbb{C}_m\,\mathbb{T}_m^{-i}\,\mathbb{B}_m,$$

while for $\mathcal{F}_m$, we get

$$\mathcal{F}_m(s) = \sum_{i=0}^{\infty} \widetilde{m}_{i+1}^g s^i, \ \text{with } \widetilde{m}_i^b = -\mathcal{C}_m\,(\mathcal{T}_m \otimes I_r)^{-i}\,\mathcal{B}_m.$$

As for the Markov parameters, the following result shows that the first m moments resulting from the Newman series of the transfer function F around $s = 0$ are also matched either by those of $\mathbb{F}_m$ when using the extended block Arnoldi process or by those of $\mathcal{F}_m$ when using the extended global Arnoldi process.

Proposition 9 *Let* $\widetilde{M}_j$ *and* $\widetilde{m_j^b}$ *be the matrix moments given by the Newman expansions of* F *and* $\mathbb{F}_m$*, respectively around* $s = 0$*. Then we have*

$$\widetilde{M}_j = \widetilde{m_j^b}, \quad \textit{for} \;\; j = 0, \dots, m-1.$$

Proof The equality is verified for $j = 0$. For $j \geq 1$, we obtain

$$\begin{aligned}\widetilde{M}_j &= C\, A^{-j}\, B \\ &= C\, A^{-j}\, V_1^b \begin{bmatrix} \Lambda_{1,1} \\ 0 \end{bmatrix} = C\, A^{-j}\, \mathbb{V}_m\, \mathbb{E}_1 \begin{bmatrix} \Lambda_{1,1} \\ 0 \end{bmatrix}.\end{aligned}$$

Therefore, using the result of Proposition 6, we get

$$\widetilde{M}_j = C\, \mathbb{V}_m\, \mathbb{L}_m^j\, \mathbb{E}_1 \begin{bmatrix} \Lambda_{1,1} \\ 0 \end{bmatrix}.$$

On the other hand, applying Proposition 8 to the upper Hessenberg matrices $\mathbb{L}_m$ and $\mathbb{T}_m^{-1}$, we get

$$\mathbb{L}_m^j\, \mathbb{E}_1 = \mathbb{T}_m^{-j}\, \mathbb{E}_1; \quad j = 0, \dots, m-1$$

and this gives for $j = 1, \dots, m-1$

$$\begin{aligned}\widetilde{M}_j &= C\, \mathbb{V}_m\, \mathbb{T}_m^{-j}\, \mathbb{V}_m^T\, V_1^b \begin{bmatrix} \Lambda_{1,1} \\ 0 \end{bmatrix} \\ &= C\, \mathbb{V}_m\, \mathbb{T}_m^{-j}\, \mathbb{V}_m^T\, B = \mathbb{C}_m\, \mathbb{T}_m^{-j}\, \mathbb{B}_m \\ &= \widetilde{m}_j^b.\end{aligned}$$

□

Now, using the extended global Arnoldi process, we can also state the following result

Proposition 10 *Let* $\widetilde{M}_j$ *and* $\widetilde{m_j^g}$ *be the matrix moments given by the Newman expansions of* F *and* $\mathcal{F}_m$*, respectively around* $s = 0$*. Then we have*

$$\widetilde{M}_j = \widetilde{m_j^g}, \quad \textit{for} \;\; j = 0, \dots, m-1.$$

Proof The equality is verified for $j = 0$. For $j \geq 1$, we obtain

$$\begin{aligned}\widetilde{M}_j &= C\, A^{-j}\, B \\ &= C\, A^{-j}\, V_1^g \begin{bmatrix} \gamma_{1,1} \\ 0 \end{bmatrix} = C\, A^{-j}\, \mathcal{V}_m\, \mathbb{E}_1 \begin{bmatrix} \gamma_{1,1} \\ 0 \end{bmatrix}.\end{aligned}$$

Therefore, using the result of Proposition 7, we get

$$\widetilde{M}_j = C\, \mathcal{V}_m\, (\mathcal{L}_m^j\, \mathcal{E}_1 \otimes I_r) \begin{bmatrix} \gamma_{1,1} \\ 0 \end{bmatrix}.$$

Table 1 Test matrices

Matrix A	size n	$\|A\|_F$	$cond(A)$
FOM	1006	1.82e+04	1000
RAIL5177	5177	5.64e+03	3.74e+07
CDplayer	120	2.31e+05	1.81e+04
Eady	598	1.26e+02	5.37e+02
MNA3	4863	2.11e+05	1.81e+08
Flow	9669	2.54e+04	1.61e+07
FDM	160,000	2.87e+08	9.79e+04

Now, similarly to the block case, applying Proposition 8 to $\mathcal{L}_m$ and $\mathcal{T}_m^{-1}$, we also have $\mathcal{L}_m^j\,\mathcal{E}_1 = \mathcal{T}_m^{-j}\,\mathcal{E}_1$ for $j = 0, \ldots, m-1$ and so we get

$$\begin{aligned}
\widetilde{M}_j = C\,\mathcal{V}_m\,(\mathcal{T}_m^{-j}\,\mathcal{E}_1 \otimes I_r)\begin{bmatrix}\gamma_{1,1}\\0\end{bmatrix} &= C\,\mathcal{V}_m\,(\mathcal{T}_m^{-j} \otimes I_r)\,(\mathcal{E}_1 \otimes I_r)\begin{bmatrix}\gamma_{1,1}\\0\end{bmatrix}\\
&= C\,\mathcal{V}_m\,(\mathcal{T}_m^{-j} \otimes I_r)\,(\mathcal{V}_m^T \diamond V_1^g)\begin{bmatrix}\gamma_{1,1}\\0\end{bmatrix}\\
&= C\,\mathcal{V}_m\,(\mathcal{T}_m^{-j} \otimes I_r)\left(\mathcal{V}_m^T \diamond \left(V_1^g\begin{bmatrix}\gamma_{1,1}\\0\end{bmatrix} \otimes I_r\right)\right)\\
&= C\,\mathcal{V}_m\,(\mathcal{T}_m^{-j} \otimes I_r)\,(\mathcal{V}_m^T \diamond B)\\
&= \mathcal{C}_m\,\mathcal{T}_m^{-j}\,\mathcal{B}_m = \widetilde{m_j^g}.
\end{aligned}$$

□

We would like to mention here that these moment matching results do not influence the extended Arnoldi algorithms themselves but just to clarify for example why the extended block and global Arnoldi algorithms allow us to match some moments and Markov parameters of transfer functions. This will be shown with some numerical experiments in the next section.

4 Numerical tests

In this section, we give some experimental results to show the effectiveness of the proposed approaches. All the experiments were performed on a computer of Intel Core i5 at 1.3 GHz and 8 GB of RAM. The algorithms were coded in Matlab R2010a. We used different known benchmark models listed in Table 1.
The matrices for the benchmark problems CDplayer, FOM, Eady, MNA3 were obtained from NICONET [14] while the matrices for the Flow and RAIL5177 models are from the Oberwolfach collection [1]. Some informations on these matrices are

[1] Oberwolfach model reduction benchmark collection, 2003. http://www.imtek.de/simulation/benchmark

reported in Table 1. For the FDM model, the corresponding matrix A is obtained from the centred finite difference discretization of the operator

$$L_A(u) = \Delta u - f(x, y)\frac{\partial u}{\partial x} - g(x, y)\frac{\partial u}{\partial y} - h(x, y)u,$$

on the unit square $[0, 1] \times [0, 1]$ with homogeneous Dirichlet boundary conditions with

$$\begin{cases} f(x, y) = \sin(x + 2y), \\ g(x, y) = e^{x+y}, \\ h(x, y) = x + y, \end{cases}$$

and the matrices B and C of sizes $n \times r$ and $r \times n$, respectively, were random matrices with entries uniformly distributed in $[0, 1]$. The number of inner grid points in each direction was $n_0 = 400$ and the dimension of A is $n = n_0^2$.
We notice that in all the figures of Example 1, the parameter m denotes the maximal iteration number for extended block Arnoldi and extended global Arnoldi algorithms. When using balanced truncation, the number m denotes also the maximal iteration number for convergence of the extended block Arnoldi algorithm when applied for solving the coupled Lyapunov equations. We also notice that for the results presented in our plots, the dimension of the reduced models are $2mr$ for extended block and global Arnoldi methods and also for balanced truncation. For Example 2, the sizes of the obtained reduced-order models are given in Table 2.

Example 1 In the first experiment, we considered the models CDplayer and FOM. Although the matrices of these models have small sizes they are usually considered as benchmark examples. The plots of Fig. 1 show the norms of the errors $\|F(j\omega) - F_m(j\omega)\|_2$ (where $j^2 = -1$) for the extended block (dashed), extended global (solid), and the balanced-truncation (dashed-dotted) methods with $\omega \in [10^{-5}, 10^5]$. We denote here that the balanced truncation method needs the solution of two low-rank

Table 2 The H_∞ error-norms $\|F - F_m\|_{H_\infty}$, execution times and reduced space dimensions for extended block, extended global, and balanced-truncation methods with the frequencies $\omega \in [10^{-5}, 10^{-2}]$

Model / Method	Bl-Extended	Gl-Extended	Balanced-Truncation
FDM, $n = 160{,}000$, $r = 5$			
H_∞-Error norms	2.5×10^{-4}	4.6×10^{-4}	6.5×10^{-4}
Times in seconds	88	92	364
Space dimension	100	104	120
Flow , $n = 9669$, $r = 3$			
H_∞-Error norms	2.6×10^{-4}	3.9×10^{-4}	1.5×10^{-6}
Times in seconds	1.9	2.1	4.5
Space dimension	180	190	220
MNA3 , $n = 4863$, $r = 4$			
H_∞-Error norms	2.6×10^{-6}	3.9×10^{-6}	2.4×10^{-7}
Times in seconds	8.1	8.6	14.8
Space dimension	300	320	360

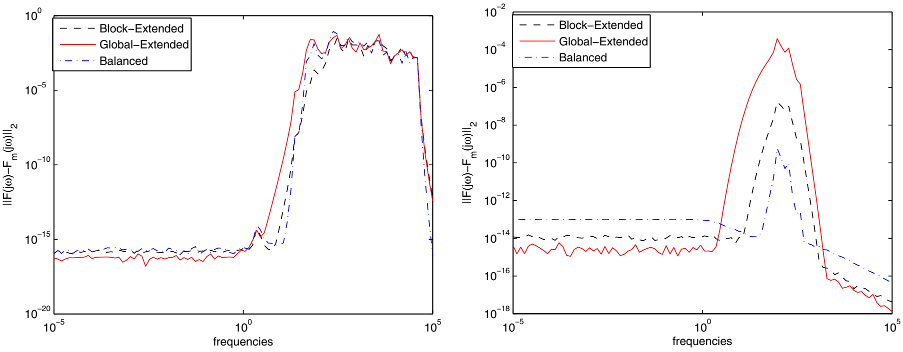

Fig. 1 The norms of the errors $\|F(j\omega) - F_m(j\omega)\|_2$ for the extended block (*dashed*), extended global (*solid*), and the balanced-truncation (*dashed-dotted*) methods with $\omega \in [10^{-5}, 10^5]$. Left the `CDplayer` model with $m = 10$ and $r = 2$. *Right:* the `FOM` model with $m = 15$, $r = 3$

right hand sides Lyapunov matrix equations that we solved by using the extended block Arnoldi method [16] and we stopped the iterations when the norm of the residual was less than 10^{-8}. Figure 1 shows that the three methods return similar results with an advantage, in the right plots of this figure, for balanced truncation for medium frequencies. However, balanced-trucation is generally more expensive as compared to the two other methods.

For the second experiment, we considered the models `RAIL5177` and `MNA3` given in Table 1. In Fig. 2, we plotted the norms of the errors $\|F(j\omega) - F_m(j\omega)\|_2$ for the extended block (dashed), extended global (solid), and the balanced-truncation (dashed-dotted) methods with $\omega \in [10^{-6}, 10^6]$.

As can be seen from Fig. 2, the three methods work well for small and high frequencies with a little advantage for the extended block and global Arnoldi methods for high frequencies.

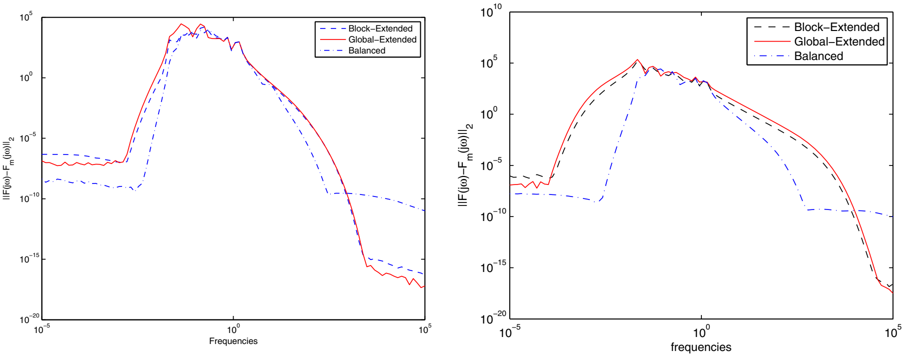

Fig. 2 The norm of the errors $\|F(j\omega) - F_m(j\omega)\|_2$ for the extended block (*dashed*), extended global (*solid*), and the balanced-truncation (*dashed-dotted*) methods . *Left:* the `RAIL5177` model with $m = 40$ and $r = 2$. *Right:* the `MNA3` model with $m = 12$ and $r = 3$

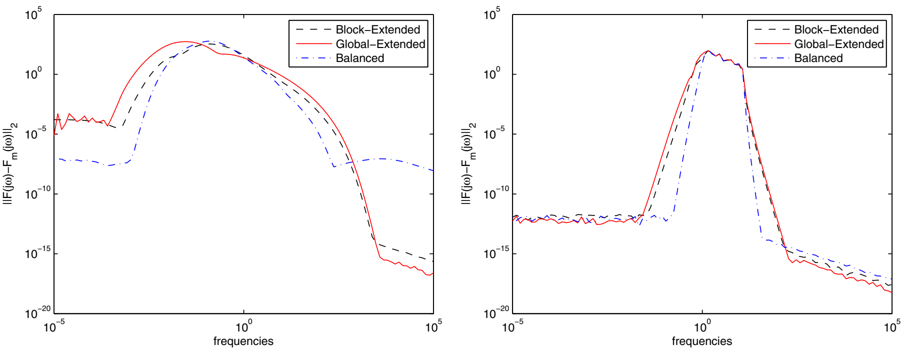

Fig. 3 The norm of the errors $\|F(j\omega) - F_m(j\omega)\|_2$ for the extended block (*dashed*), extended global (*solid*) and the balanced-truncation (*dashed-dotted*) methods with $\omega \in [10^{-5}, 10^5]$. *Left:* the `Flow` model with $m = 15$ and $r = 3$. *Right:* the `Eady` model with $m = 10$ and $r = 3$

The plots in Fig. 3, represent the norms of the errors $\|F(j\omega) - F_m(j\omega)\|_2$ corresponding to the extended block and global Arnoldi methods and to the balanced-truncation method for the models: the `Eady` model with $m = 10$ and $r = 3$ and the `Flow` model with $m = 15$ and $r = 3$ for the frequencies $\omega \in [10^{-5}, 10^5]$.

Example 2 For this example, we compared the obtained H_∞ error-norms $\|F - F_m\|_{H_\infty}$, the execution times and the reduced space dimensions for the extended block and global Arnoldi algorithms with those obtained by the balanced-truncation method in which the two coupled low-rank right hand sides Lyapunov matrix equations were solved by the extended block Arnoldi algorithm. For the latter method, the inner iterations were stopped when the norm of the residual was less than 10^{-8} and the obtained approximate solution was given as a product of a matrix with a low rank with its transpose. We considered three models: `FDM` with $n = 160{,}000$ and $r = 5$, the `flow-meter` model with $n = 9669$ and $r = 3$, and the `MNA3` model with $n = 4863$ and $r = 4$.

The results of Table 2 show that the cost of balanced truncation method is generally higher than the cost of the extended block or global Arnoldi methods. However, some of the obtained H_∞ norms are good when using the balanced truncation method.

5 Conclusion

In this paper, we considered the extended block and global Arnoldi methods. We gave some new algebraic properties of these two algorithms. We also showed how these properties could be used in moment matching methods for model reduction in large-scale dynamical systems. The proposed numerical results on some Benchmark models, show that the extended block and global Arnoldi algorithms are efficient. Generally, the two methods return similar results. One advantage of the extended global Arnoldi is the fact that a beak-down cannot occur which may be the case for the extended block Arnoldi algorithm.

References

1. Bouyouli, R., Jbilou, K., Sadaka, R., Sadok, H.: Convergence properties of some block Krylov subspace methods for multiple linear systems. J. Comput. Appl. Math. **196**, 498–511 (2006)
2. Brezinski, C.: Computational Aspects of Linear Control. Kluwer, Dordrecht (2002)
3. Druskin, V., Knizhnerman, L.: Extended Krylov subspaces: approximation of the matrix square root and related functions, SIAM. J. Matrix Anal. Appl. **3**, 755–771 (1998)
4. Druskin, V., Simoncini, V.: Adaptive rational Krylov subspaces for large-scale dynamical systems. Syst. Control Lett. **60**(8), 546–560 (2011)
5. Druskin, V., Lieberman, C., Zaslavsky, M.: On adaptive choice of shifts in rational Krylov subspace reduction of evolutionary problems. SIAM J. Sci. Comput. **32**(5), 2485–2496 (2010)
6. Gallivan, K., Grimme, E., Van Dooren, P.: A rational Lanczos algorithm for model reduction. Numer. Alg. **12**, 33–63 (1996)
7. Glover, K.: All optimal Hankel-norm approximations of linear multivariable systems and their L-infinity error bounds. Inter. J. Cont. **39**, 1115–1193 (1984)
8. Gugercin, S., Antoulas, A.C.: A survey of model reduction by balanced truncation and some new results. Int. J. Control. **77**, 748–766 (2004)
9. Gugercin, S., Antoulas, A.C., Beattie, C.: h_2 model reduction for large-scale dynamical systems. SIAM J. Matrix Anal. Appl. **30**(2), 609–638 (2008)
10. Heyouni, M., Jbilou, K.: Matrix Krylov subspace methods for large scale model reduction problems. App. Math. Comput. **181**, 1215–1228 (2006)
11. Heyouni, M., Jbilou, K.: An extended block Arnoldi algorithm for large-scale solutions of the continuous-time algebraic Riccati equation. Elect. Trans. Num. Anal. **33**, 53–62 (2009)
12. Heyouni, M.: Extended Arnoldi methods for large Sylvester matrix equations. App. Num. Math. **60**(11), 1171–1182 (2010)
13. Knizhnerman, L., Simoncini, V.: Convergence analysis of the Extended Krylov Subspace Method for the Lyapunov equation. Numer. Math. **118**(3), 567–586 (2011)
14. Mehrmann, V., Penzl, T.: Benchmark Collections in SLICOT. Technical Report SLWN1998- 5, 1 Working Note, ESAT, KU Leuven, K. Mercierlaan 94, Leuven-Heverlee, 3100, Belgium. http://www.win.tue.nl/niconet/NIC2/reports.html (1998)
15. Moore, B.C.: Principal component analysis in linear systems: controllability, observability and model reduction. IEEE Trans. Automatic Contr. **AC-26**, 17–32 (1981)
16. Simoncini, V.: A new iterative method for solving large-scale Lyapunov matrix equations. SIAM J. Sci. Comp. **29**(3), 1268–1288 (2007)
17. Van Dooren, P., Gallivan, K.A., Absil, P.A.: h_2-optimal model reduction of MIMO systems. Appl. Math. Lett. **21**(12), 1267–1273 (2008)