

H. BARKOUKI (Marrakech)
A. H. BENTBIB (Marrakech)
K. JBILOU (Calais)

A MODIFIED NONSYMMETRIC RATIONAL BLOCK LANCZOS METHOD FOR MODEL REDUCTION IN LARGE SCALE LTI DYNAMICAL SYSTEMS

Abstract. We propose an adaptive model reduction algorithm for computing a reduced-order model of dynamical multi-input and multi-output (MIMO) linear time independent (LTI) dynamical systems. The process is based on multipoint moment matching. Moreover, we develop new simple Lanczos-like equations for the rational block case, and we use them to derive simple residual error expressions. An adaptive method for choosing the interpolation points is also proposed. Finally, some numerical experiments are reported to show the effectiveness of the new adaptive modified rational block Lanczos (AMRBL) process when applied to stable linear LTI dynamical systems.

1. Introduction. We consider the following multi-input and multi-output (MIMO) linear time invariant (LTI) dynamical system:

$$(1) \quad \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t), \end{cases}$$

where $x(t) \in \mathbb{R}^n$ is the state vector and $u(t), y(t) \in \mathbb{R}^p$ are the input and the output vectors of (1). The matrix $A \in \mathbb{R}^{n \times n}$ is assumed to be large and sparse, and $B, C^T \in \mathbb{R}^{n \times p}$. A classical way for relating the input to the output is to use the transfer function of (1). If we apply the Laplace

2010 *Mathematics Subject Classification:* Primary 65F; Secondary 15A.

Key words and phrases: rational block Lanczos, moment matching, model-order reduction, transfer function.

Received 20 April 2016; revised 12 May 2016.

Published online 4 November 2016.

transform to (1), we obtain

$$\begin{cases} sX(s) = AX(s) + BU(s), \\ Y(s) = CX(s), \end{cases}$$

where $X(s)$, $Y(s)$ and $U(s)$ are the Laplace transforms of $x(t)$, $y(t)$ and $u(t)$, respectively. If we eliminate $X(s)$ from the previous two equations, we obtain $Y(s) = H(s)U(s)$, where $H(s)$ is called the *transfer function* of (1) and is defined as

$$H(s) = C(sI_n - A)^{-1}B.$$

In many applications, the number of state variables is very high, and therefore it is not practical to use the full system for simulation or run-on-time control. Then the aim of model reduction problems is to produce a lower-dimensional system of the form

$$(2) \quad \begin{cases} \dot{x}_m(t) = A_m x_m(t) + B_m u(t), \\ y_m(t) = C_m x_m(t), \end{cases}$$

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

$$H_m(s) = C_m(sI_r - A_m)^{-1}B_m.$$

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 [2, 18, 23] and the references therein). One of the most common approaches of the first category is the so-called balanced model reduction, introduced by Moore [30]. For Krylov subspace methods, the Lanczos process has been extensively used for SISO (the case $p = 1$) and MIMO dynamical systems (see [7, 8, 13, 14, 25, 27] and the references therein). Unfortunately, the standard version of the Lanczos algorithm builds reduced order models that poorly approximate low frequency dynamics. In order to address this problem, rational Krylov subspace methods have recently been developed [4, 15, 16, 22, 21, 35, 33] for SISO and MIMO dynamical systems. The major problem in these methods is the selection of interpolation points that have to guarantee a good convergence of the process. Various methods for construction of good interpolation points have been proposed [9, 6, 5, 22, 21, 28, 34].

The paper is organized as follows. We first review, in Section 2, some moment matching techniques for model reduction. In Section 3, we propose a

modified nonsymmetric rational block Lanczos process, and derive Lanczos-like equations for the rational case. Simple residual error expressions are developed in Section 4. In Section 5, we propose an adaptive choice for selecting the shifts used in the adaptive rational block Lanczos algorithm. The last section is devoted to some numerical experiments.

2. Moment matching based methods

2.1. The moment matching problem. Let $H(s) = C(sI_n - A)^{-1}B$ be the transfer function of a linear dynamical system as described in (1). Expanding $H(s)$ in a power series around a given point $\sigma_0 \in \mathbb{R}$, we get

$$H(s) = h_0 + h_1(s - \sigma_0) + h_2(s - \sigma_0)^2 + \dots,$$

where the coefficients h_j for $j \geq 0$ are known as *moments* of the system around σ_0 and they are given by

$$h_j(\sigma_0) = C(\sigma_0 I_n - A)^{-(j+1)}B.$$

These moments are the values of the transfer function of (1) and of its derivatives at σ_0 , and they are also called shifted moments. The model reduction using the moment matching method consists in finding a lower-order transfer function $H_m(s)$ having a power series expansion

$$H_m(s) = \hat{h}_0 + \hat{h}_1(s - \sigma_0) + \hat{h}_2(s - \sigma_0)^2 + \dots$$

such that $2m$ moments are matched, i.e.,

$$(3) \quad h_j(\sigma_0) = \hat{h}_j(\sigma_0), \quad j = 0, \dots, 2m - 1,$$

for an appropriate $m \ll n$. The resulting reduced-order model is known as *rational interpolation*. If $\sigma_0 = 0$, the moments satisfy $h_j = -CA^{-(j+1)}B$ for $j \geq 0$, and the problem is known as *Padé approximation* [1, 35]. The Laurent series of the transfer function H around $\sigma_0 = \infty$ is

$$H(s) = \sum_{i=0}^{\infty} h_i s^{-i},$$

where $h_i = CA^i B$ for $i \geq 0$ are called the *Markov parameters* and the corresponding problem is known as a *partial realization* [20]. We can also generate a reduced-order model using multipoint Padé approximation or multipoint rational interpolation [10, 21, 36]. By multipoint rational interpolation, we mean that the reduced system matches the moments of the original system at multiple interpolation points.

Let $\mathbb{V}_m, \mathbb{W}_m \in \mathbb{R}^{n \times mp}$. Then the reduced-order system (2) can be constructed by applying the oblique projector $\Pi = \mathbb{V}_m \mathbb{W}_m^T$ to (1) and obtaining lower-dimensional matrices

$$(4) \quad A_m = \mathbb{W}_m^T A \mathbb{V}_m, \quad B_m = \mathbb{W}_m^T B, \quad C_m = C \mathbb{V}_m.$$

A careful selection of \mathbb{V}_m and \mathbb{W}_m as bases of certain Krylov subspaces results in moment matching. Let $\sigma \in \mathbb{R}$ and $m \in \mathbb{Z}_{>0}$. Then the block Krylov subspace $\mathbb{K}_m(A, B, \sigma)$ is given by

$$\mathbb{K}_m(A, B, \sigma) = \begin{cases} \text{Range}\{(A - \sigma I_n)^{-1}B, \dots, (A - \sigma I_n)^{-m}B\} & \text{if } \sigma \neq \infty, \\ \text{Range}\{B, AB, A^2B, \dots, A^{m-1}B\} & \text{if } \sigma = \infty. \end{cases}$$

These are subspaces of \mathbb{R}^n generated by the columns of the matrices appearing in the expressions of $\mathbb{K}_m(A, B, \sigma)$. The block Krylov subspaces corresponding to $\sigma = \infty$ were widely used in many topics of linear algebra and applications to control theory and other engineering problems. Different algorithms such as block Arnoldi or block Lanczos were proposed to get orthonormal bases for these block subspaces. In this paper, we focus on the block Lanczos process [3, 7, 19, 12, 31] and give an extension to the rational case.

2.2. The standard block Lanczos algorithm. Let V and W be two initial blocs of $\mathbb{R}^{n \times p}$. Then the nonsymmetric block Lanczos algorithm applied to the pairs (A, V, ∞) and (A^T, W, ∞) generates two sequences of biorthonormal $n \times p$ matrices $\{V_i\}$ and $\{W_j\}$, such that

$$\begin{aligned} \mathbb{K}_m(A, V, \infty) &= \text{Range}([V_1, \dots, V_m]), \\ \mathbb{K}_m(A^T, W, \infty) &= \text{Range}([W_1, \dots, W_m]). \end{aligned}$$

The matrices V_i and W_j that are generated by the block Lanczos algorithm satisfy the biorthogonality condition

$$(5) \quad \begin{cases} W_j^T V_i = 0_p & \text{if } i \neq j, \\ W_j^T V_i = I_p & \text{if } i = j. \end{cases}$$

The classical versions of the block Lanczos process are generally unstable numerically. In [3], a more stable version of the nonsymmetric block Lanczos algorithm was proposed. This algorithm, named Adaptive Block Lanczos (ABLE), is summarized as follows.

ALGORITHM 1. The nonsymmetric block Lanczos algorithm

Inputs: $A \in \mathbb{R}^{n \times n}$, $V, W \in \mathbb{R}^{n \times p}$ and $m \in \mathbb{R}$

Initialize: $V_0 = W_0 = 0_p$ and $C_1 = B_1 = 0_p$

for $j = 1, \dots, m$

$S_j = AV_j$ and $R_j = A^T W_j$;

$A_j = W_j^T S_j, B_j = W_{j-1}^T S_j$ and $C_j^T = V_{j-1}^T R_j$;

$S_j = S_j - V_j A_j - V_{j-1} B_j$ and $R_j = R_j - W_j A_j^T - W_{j-1} C_j^T$;

Compute the QR decomposition $S_j = V_{j+1} C_{j+1}^T$ and $R_j = W_{j+1} B_{j+1}^T$;

Compute the SVD method $W_{j+1}^T V_{j+1} = P_j D_j Q_j^T$;

$V_{j+1} = V_{j+1} Q_j D_j^{-1/2}$ and $W_{j+1} = W_{j+1} P_j D_j^{-1/2}$;

$B_{j+1} = D_j^{1/2} Q_j^T B_{j+1}$ and $C_{j+1} = D_j^{1/2} P_j^T C_{j+1}$

end.

interpolates the original transfer function $H(s)$ and its derivatives at given interpolation frequencies.

THEOREM 1. *Let $\Sigma = \{\sigma_1, \dots, \sigma_K\}$ and $\tilde{\Sigma} = \{\tilde{\sigma}_1, \dots, \tilde{\sigma}_K\}$ be two sets of interpolation points, with multiplicities m_1, \dots, m_K and $\tilde{m}_1, \dots, \tilde{m}_K$, respectively. Assume that $\mathbb{V}_m, \mathbb{W}_m \in \mathbb{R}^{n \times mp}$ satisfy*

$$\bigcup_{k=1}^K \mathbb{K}_{m_k}(A, B, \sigma_k) \subseteq \text{Range}(\mathbb{V}_m),$$

$$\bigcup_{k=1}^K \mathbb{K}_{\tilde{m}_k}(A^T, C^T, \tilde{\sigma}_k) \subseteq \text{Range}(\mathbb{W}_m),$$

where $\sum_{k=1}^K m_k = \sum_{k=1}^K \tilde{m}_k = m$, and $(A - \sigma I_n)^{-1}$ exists for all $\sigma \in \Sigma \cup \tilde{\Sigma}$.

- If $\sigma_k = \tilde{\sigma}_k$, then $H_m(s)$ matches the first $m_k + \tilde{m}_k$ moments of $H(s)$ at σ_k .
- If $\sigma_k \neq \tilde{\sigma}_k$, then $H_m(s)$ matches the first m_k moment of $H(s)$ at σ_k and the first \tilde{m}_k moments of $H(s)$ at $\tilde{\sigma}_k$, respectively.

The modified rational block Lanczos process constructs two biorthonormal bases and is summarized in Algorithm 2. It is a generalization of the one given in [15] to the block case. For simplicity of presentation we assume that $m_k = \tilde{m}_k$, and also that $\sigma_i \neq \sigma_j$ and $\tilde{\sigma}_i \neq \tilde{\sigma}_j$ for $i \neq j$.

ALGORITHM 2. The modified rational block Lanczos algorithm

Input: $\Sigma = \{\sigma_1, \dots, \sigma_K\}$, $\tilde{\Sigma} = \{\tilde{\sigma}_1, \dots, \tilde{\sigma}_K\}$, A, B, C and $m_k = \tilde{m}_k$

Initialize: $\mathbb{V}_m = [], \mathbb{W}_m = []$ and $i = 0$

for $k = 1, \dots, K$

if $\sigma_k = \infty$ **then** $S_{i+1} = B$ **else** $S_{i+1} = (A - \sigma_k I_n)^{-1} B$ **end**

if $\tilde{\sigma}_k = \infty$ **then** $R_{i+1} = C^T$ **else** $R_{i+1} = (A - \tilde{\sigma}_k I_n)^{-T} C^T$ **end**

$S_{i+1} = S_{i+1} - \mathbb{V}_m \mathbb{W}_m^T S_{i+1}$; $R_{i+1} = R_{i+1} - \mathbb{W}_m \mathbb{V}_m^T R_{i+1}$;

$S_{i+1} = V_{i+1} H_{i+1,i}$; $R_{i+1} = W_{i+1} G_{i+1,i}$ (QR factorization);

$W_{i+1}^T V_{i+1} = P_i D_i Q_i^T$ (Singular Value Decomposition);

$V_{i+1} = V_{i+1} Q_i D_i^{-1/2}$; $W_{i+1} = W_{i+1} P_i D_i^{-1/2}$;

$\mathbb{V}_m = [\mathbb{V}_m, V_{i+1}]$; $\mathbb{W}_m = [\mathbb{W}_m, W_{i+1}]$; $i = i + 1$

for $j = 1, \dots, m_k - 1$

if $\sigma_k = \infty$ **then** $S_{i+1} = A V_i$ **else** $S_{i+1} = (A - \sigma_k I_n)^{-1} V_i$ **end**

if $\tilde{\sigma}_k = \infty$ **then** $R_{i+1} = A^T W_i$ **else** $R_{i+1} = (A - \tilde{\sigma}_k I_n)^{-T} W_i$ **end**

$S_{i+1} = S_{i+1} - \mathbb{V}_m \mathbb{W}_m^T S_{i+1}$; $R_{i+1} = R_{i+1} - \mathbb{W}_m \mathbb{V}_m^T R_{i+1}$;

$S_{i+1} = V_{i+1} H_{i+1,i}$; $R_{i+1} = W_{i+1} G_{i+1,i}$ (QR factorization);

$W_{i+1}^T V_{i+1} = P_i D_i Q_i^T$ (Singular Value Decomposition);

$V_{i+1} = V_{i+1} Q_i D_i^{-1/2}$; $W_{i+1} = W_{i+1} P_i D_i^{-1/2}$;

$\mathbb{V}_m = [\mathbb{V}_m, V_{i+1}]$; $\mathbb{W}_m = [\mathbb{W}_m, W_{i+1}]$; $i = i + 1$

end

if $\{k = K\}$ **then**

$S_{i+1} = A^{m_\infty} B$ and $R_{i+1} = (A^{\tilde{m}_\infty})^T C^T$;

$$\begin{aligned}
 S_{i+1} &= S_{i+1} - \mathbb{V}_m \mathbb{W}_m^T S_{i+1}; R_{i+1} = R_{i+1} - \mathbb{W}_m \mathbb{V}_m^T R_{i+1}; \\
 S_{i+1} &= V_{i+1} H_{i+1,i}; R_{i+1} = W_{i+1} G_{i+1,i} \quad (\text{QR factorization}); \\
 W_{i+1}^T V_{i+1} &= P_i D_i Q_i^T \quad (\text{Singular Value Decomposition}); \\
 V_{i+1} &= V_{i+1} Q_i D_i^{-1/2}; W_{i+1} = W_{i+1} P_i D_i^{-1/2}; \\
 \mathbb{V}_{m+1} &= [\mathbb{V}_m, V_{m+1}]; \mathbb{W}_{m+1} = [\mathbb{W}_m, W_{m+1}]
 \end{aligned}$$

end

end.

3.2. The Lanczos-like equations for the rational block case.

In [4], the authors give relations between $\mathbb{V}_m, \mathbb{W}_m$ and the matrix A in the rational form for the block Lanczos algorithm. Here, we show that the modified rational block Lanczos process proposed in the last section yields Lanczos-like equations for the rational case. This result is given in [14, 15] for the standard rational Lanczos algorithm and extended here to the block case.

THEOREM 2. *Let \mathbb{V}_{m+1} and \mathbb{W}_{m+1} be the matrices generated by the modified rational block Lanczos algorithm (Algorithm 2). Then*

$$\begin{aligned}
 \text{Range}[\mathbb{V}_m, A^{m_\infty} B] &\subseteq \text{Range}\{\mathbb{V}_{m+1}\}, \\
 \text{Range}[\mathbb{W}_m, (A^{\tilde{m}_\infty})^T C^T] &\subseteq \text{Range}\{\mathbb{W}_{m+1}\}, \\
 \mathbb{W}_{m+1}^T \mathbb{V}_{m+1} &= I_{m+1},
 \end{aligned}$$

where m_∞ and \tilde{m}_∞ are the multiplicities of ∞ in Σ and $\tilde{\Sigma}$, respectively.

Moreover, we have the Lanczos-like relations

$$\begin{aligned}
 (6) \quad & A \mathbb{V}_m = \mathbb{V}_m A_m + V_{m+1} P_{m+1}, \\
 (7) \quad & B = \mathbb{V}_m B_m + V_{m+1} b_m, \\
 (8) \quad & A^T \mathbb{W}_m = \mathbb{W}_m A_m^T + W_{m+1} Q_{m+1}, \\
 (9) \quad & C^T = \mathbb{W}_m C_m^T + W_{m+1} c_m^T,
 \end{aligned}$$

where $b_m = W_{m+1}^T B, c_m = C V_{m+1}, P_{m+1} = W_{m+1}^T A \mathbb{V}_m$ and $Q_{m+1} = V_{m+1}^T A^T \mathbb{W}_m$. Furthermore, $b_m = 0$ if $m_\infty > 0$, and $c_m = 0$ if $\tilde{m}_\infty > 0$.

Proof. We first prove the result for $m_\infty = 0$. Let \mathbb{V}_m be as in Theorem 2. We extend \mathbb{V}_m to $\mathbb{V}_{m+1} = [\mathbb{V}_m, V_{m+1}]$ such that

$$\text{Range}[\mathbb{V}_m, B] \subseteq \text{Range}\{\mathbb{V}_{m+1}\}$$

by biorthogonalizing B against all previous columns of \mathbb{W}_m with the Lanczos algorithm. Then

$$\begin{aligned}
 (10) \quad & \text{Range}\{(A - \sigma_1 I_n)^{-1} B\} \subset \text{Range}\{\mathbb{V}_1\}, \\
 & \text{Range}\{(A - \sigma_k I_n)^{-(i-1)} B\} \subset \text{Range}\{\mathbb{V}_{j-1}\}, \\
 (11) \quad & \text{Range}\{(A - \sigma_k I_n)^{-i} B\} \subset \text{Range}\{\mathbb{V}_j\}, \\
 & \text{Range}\{B\} \subset \text{Range}\{\mathbb{V}_{m+1}\},
 \end{aligned}$$

where $\mathbb{V}_m = [V_1, \dots, V_m] \in \mathbb{R}^{n \times mp}$, $m = \sum_{i=1}^K m_i$ and \mathbb{V}_j is the matrix constructed at the k th interpolation point and for the i th multiplicity, i.e., $j = \sum_{l=1}^{k-1} m_l + i$. We start by proving the theorem for the first block column of \mathbb{V}_m . Multiply (10) by $A - \sigma_1 I_n$ from the left and rearrange to get

$$\text{Range}\{A\mathbb{V}_1\} \subset \text{Range}\{\mathbb{V}_1, B\}.$$

Then

$$\text{Range}\{A\mathbb{V}_1\} \subset \text{Range}\{\mathbb{V}_{m+1}\}.$$

We now proceed by induction. We assume that the result holds for an arbitrary interpolation point σ_k of the \mathbb{K}_{σ_k} Krylov subspace up to the $(i - 1)$ th multiplicity. We will prove the result for the next multiplicity. Therefore we assume

$$(12) \quad \text{Range}\{A\mathbb{V}_{j-1}\} \subset \text{Range}\{\mathbb{V}_{m+1}\},$$

and we will prove that

$$\text{Range}\{A\mathbb{V}_j\} \subset \text{Range}\{\mathbb{V}_{m+1}\}.$$

Multiply (11) from the left by $A - \sigma_k I_n$ and rearrange to get

$$\text{Range}\{A\mathbb{V}_j\} \subset \text{Range}\{\mathbb{V}_j, A\mathbb{V}_{j-1}\},$$

which gives

$$\text{Range}\{A\mathbb{V}_j\} \subset \text{Range}\{\mathbb{V}_{m+1}\}.$$

Combining the last relation with (12) gives

$$(13) \quad \text{Range}\{A\mathbb{V}_j\} \subset \text{Range}\{\mathbb{V}_{m+1}\}.$$

It is easy to see that the result in (13) holds for all columns in \mathbb{V}_m , i.e.,

$$\text{Range}\{A\mathbb{V}_m\} \subset \text{Range}\{\mathbb{V}_{m+1}\},$$

and so there exists a matrix $Y \in \mathbb{R}^{(m+1)p \times mp}$ such that

$$A\mathbb{V}_m = \mathbb{V}_{m+1}Y.$$

Setting $Y = \begin{pmatrix} \tilde{A}_m \\ P_{m+1} \end{pmatrix}$, we obtain

$$A\mathbb{V}_m = \mathbb{V}_{m+1} \begin{pmatrix} \tilde{A}_m \\ P_{m+1} \end{pmatrix}.$$

Using the biorthogonality between \mathbb{W}_{m+1} and \mathbb{V}_{m+1} gives $\tilde{A}_m = A_m = \mathbb{W}_m^T A \mathbb{V}_m$ and $P_{m+1} = W_{m+1}^T A \mathbb{V}_m$. Therefore

$$A\mathbb{V}_m = \mathbb{V}_m A_m + V_{m+1} P_{m+1}.$$

Similarly, to prove (7) we proceed as follows: As $\text{Range}\{B\} \subset \text{Range}\{\mathbb{V}_{m+1}\}$, there exists a matrix $Z \in \mathbb{R}^{(m+1)p \times p}$ such that

$$(14) \quad B = \mathbb{V}_{m+1}Z.$$

Setting $Z = \begin{pmatrix} \tilde{B}_m \\ b_m \end{pmatrix}$ and multiplying (14) by \mathbb{W}_{m+1}^T from the left, we get

$$(15) \quad \mathbb{W}_{m+1}^T B = \begin{pmatrix} \tilde{B}_m \\ b_m \end{pmatrix}.$$

Then $\tilde{B}_m = B_m = \mathbb{W}_m^T B$, $b_m = W_{m+1}^T B$ and consequently

$$B = \mathbb{V}_m B_m + V_{m+1} b_m.$$

Assume now that $m_\infty > 0$. The matrix \mathbb{V}_m satisfies

$$\text{Range}[B \ AB \ \dots \ A^{p-1}B \ \mathbb{V}_{m-p}] \subset \text{Range}\{\mathbb{V}_m\}$$

for $p < m$. Since B is already in the range of \mathbb{V}_m , it is easy to see that (6) will be satisfied if

$$(16) \quad \text{Range}\{AV_1, \dots, AV_p\} \subset \text{Range}\{\mathbb{V}_{m+1}\}.$$

This can be shown by setting $\mathbb{V}_{m+1} = [\mathbb{V}_m, V_{m+1}]$. Then

$$\text{Range}[\mathbb{V}_m, A^p B] \subset \text{Range}\{\mathbb{V}_{m+1}\} \quad \text{and} \quad \mathbb{W}_{m+1}^T \mathbb{V}_{m+1} = I_{m+1},$$

which is obtained by biorthogonalizing $A^p B$ against all the previous columns of \mathbb{W}_m . It follows that (16) holds, since by construction we have

$$(17) \quad \text{Range}\{A^k B\} \subset \begin{cases} \text{Range}\{V_1, \dots, V_{k+1}\} & \text{for } 0 < k < p, \\ \text{Range}\{V_1, \dots, V_{m+1}\} & \text{for } k = p, \end{cases}$$

which completes the proof of (6) and (7).

To prove the last part, note that if $m_\infty > 0$ then $B \in \text{Range}\{\mathbb{V}_m\}$, from which it follows that $b_m = 0$.

In a similar way, (8) and (9) can be shown.

4. Error estimations and residual error expressions. The computation of the norm of the error

$$(18) \quad \varepsilon(s) = H(s) - H_m(s)$$

is important to measure the accuracy of the resulting reduced-order model. Unfortunately, the exact error $\varepsilon(s)$ is not available during the process, because the high dimension of the original system yields the computation of $H(s)$ difficult. To remedy this, various approaches have been explored. In [21], a method was proposed of computing the modeling error in terms of two residual vectors in the case of single-input and single-output systems.

These results are extended here to the multi-input and multi-output case. Let

$$(19) \quad \begin{cases} R_B(s) = B - (sI_n - A)\mathbb{V}_m \tilde{X}_B(s), \\ R_C(s) = C^T - (sI_n - A)^T \mathbb{W}_m \tilde{X}_C(s) \end{cases}$$

be the residual expressions where $\tilde{X}_B(s)$ and $\tilde{X}_C(s)$ are the solutions of the multiple linear systems

$$\begin{cases} (sI_{mp} - A_m)\tilde{X}_B(s) = B_m, \\ (sI_{mp} - A_m)^T \tilde{X}_C(s) = C_m^T, \end{cases}$$

and satisfy the Petrov–Galerkin condition

$$\begin{cases} R_B(s) \perp \text{Range}(W_1, \dots, W_m), \\ R_C(s) \perp \text{Range}(V_1, \dots, V_m), \end{cases}$$

which means that $\mathbb{W}_m^T R_B(s) = \mathbb{V}_m^T R_C(s) = 0$. Then we can state the following theorem.

THEOREM 3. *The error between the frequency responses of the original and reduced-order systems is given by*

$$(20) \quad \varepsilon(s) = R_C^T(s)(sI_n - A)^{-1}R_B(s).$$

The proof is similar to the one given in [21] for SISO dynamical systems. In [15] the authors developed simple Lanczos-like equations for the standard rational algorithm and used them to derive simple residual errors expressions. Such results for the rational block case are given in the next subsection.

4.1. Simplified Lanczos residual errors in the rational block Lanczos. Using the results of Theorem 2, we first give a new expression of the residuals $R_B(s)$ and $R_C(s)$. In fact, from (19) we get

$$\begin{aligned} R_B(s) &= B - (sI - A)\mathbb{V}_m(sI_{mp} - A_m)^{-1}B_m. \\ &= B - \mathbb{V}_m(sI_{mp} - A_m)(sI_{mp} - A_m)^{-1}B_m \\ &\quad + V_{m+1}P_{m+1}(sI_{mp} - A_m)^{-1}B_m, \\ &= V_{m+1}(P_{m+1}(sI_{mp} - A_m)^{-1}B_m + b_m), \\ &= V_{m+1}\tilde{R}_B(s), \end{aligned}$$

where $\tilde{R}_B(s) = P_{m+1}(sI_{mp} - A_m)^{-1}B_m + b_m$ is the frequency-dependent term of the residual error $R_B(s)$.

In a similar way, we can obtain the relations

$$\begin{aligned} R_C(s) &= C^T - (sI - A)^T\mathbb{W}_m(sI_{mp} - A_m)^{-T}C_m^T \\ &= W_{m+1}\tilde{R}_C(s), \end{aligned}$$

where $\tilde{R}_C(s) = Q_{m+1}(sI_{mp} - A_m)^{-T}C_m^T + c_m^T$ is the frequency-dependent term of the residual error $R_C(s)$.

Then the error expression in (20) becomes

$$(21) \quad \varepsilon_m(s) = \tilde{R}_C(s)^T \tilde{C}(sI_n - A)^{-1} \tilde{B} \tilde{R}_B(s) = \tilde{R}_C(s)^T \tilde{H}(s) \tilde{R}_B(s).$$

The transfer function $\tilde{H}(s) = \tilde{C}(sI_n - A)^{-1}\tilde{B}$ contains terms related to the original system, which makes the computation of $\|\tilde{R}_C^T \tilde{H} \tilde{R}_B\|_\infty$ expensive. Therefore, instead of using $\tilde{H}(s)$, we can use an approximation of $\tilde{H}(s)$.

Table 1. Various estimations of the error

$$\begin{array}{l}
 \hat{\varepsilon}_m(s) = \tilde{R}_B(s) \\
 \hat{\varepsilon}_m(s) = \tilde{R}_C(s)^T \\
 \hat{\varepsilon}_m(s) = \tilde{H}_m(s)\tilde{R}_B(s) \\
 \hat{\varepsilon}_m(s) = \tilde{H}_m(s) \\
 \hat{\varepsilon}_m(s) = \tilde{R}_C^T(s)\tilde{H}_m(s) \\
 \hat{\varepsilon}_m(s) = \tilde{R}_C^T(s)\tilde{H}_m(s)\tilde{R}_B(s)
 \end{array}$$

Various possible approximations are listed in Table 1. The simple ones are the first two for which the computations require small work as compared to the other choices. In the last section, we will give some numerical tests comparing these approximations.

5. An adaptive modified rational block Lanczos algorithm. Model-order reduction using multipoint rational interpolation generally gives a more accurate reduced-order model than interpolation around a single point. Unfortunately, the selection of interpolation points is not an automated process and it requires an appropriate choice of a more accurate rational Krylov subspace. In [5, 24] the iterative rational Krylov algorithm (IRKA) has been proposed in the context of H_2 -optimal model-order reduction by using a specific way to choose the interpolation points σ_i , $i = 1, \dots, m$. Starting from an initial set of interpolation points, a reduced-order system is determined and a new set of interpolation points is chosen which are the Ritz values $-\lambda_i(A_m)$, $i = 1, \dots, m$, i.e., $\lambda_i(A_m)$ are the eigenvalues of A_m . The process continues until the Ritz values from consecutive reduced-order models stabilize.

In contrast to the IRKA method, the authors of [22] proposed an adaptive method for choosing the interpolation points. This approach is based on the residual expression derived for the rational Lanczos algorithm such that the interpolation points are selected where the residual error is large. At each iteration of the algorithm, a residual function is computed and a new interpolation point is selected so as to correspond to the maximum of this residual function.

For more adaptive interpolation points methods, we refer the readers to [6, 11, 26, 28, 34]. In [6] the authors proposed an adaptive computation of the shifts in the case of the standard rational Arnoldi. This method is based on approximation of an upper bound of the norm of the error between the

original and the reduced transfer functions. Here, we use the same approach for our rational block Lanczos algorithm. The following result is the key ingredient of this adaptive method.

PROPOSITION 1. *Let $\Sigma_m = \{\sigma_1, \dots, \sigma_m\}$ denote a given set of interpolation points and let $\mathbb{V}_m, \mathbb{W}_m \in \mathbb{R}^{n \times mp}$ be the biorthogonal matrices computed by the rational block Lanczos algorithm. Then*

$$(22) \quad \|H(s) - H_m(s)\|_2 \leq \|CP^{-1}(s)\|_2 \|R_B(s)\|_2,$$

where $R_B(s) = B - P(s)\mathbb{V}_m P_m^{-1}(s)B_m$, $P(s) = sI - A$ and $B_m = \mathbb{W}_m^T B$.

Proof. From the expressions of $H(s)$ and $H_m(s)$, we have

$$\begin{aligned} \|H(s) - H_m(s)\|_2 &= \|C(sI_n - A)^{-1}B - C_m(sI_{mp} - A_m)^{-1}B_m\|_2 \\ &= \|CP(s)^{-1}(B - P(s)\mathbb{V}_m P_m^{-1}(s)B_m)\|_2 \\ &\leq \|CP(s)^{-1}\|_2 \|R_B(s)\|_2. \end{aligned}$$

Now, we can approximate the upper bound by employing the reduced order matrix triplet (A_m, B_m, C_m) ; then $\|CP(s)^{-1}\|$ could be approximated by $\|C_m P_m(s)^{-1}\|$, where $P_m(s) = sI_{mp} - A_m$ and $A_m = \mathbb{W}_m^T A \mathbb{V}_m$.

Using the above approximation, the next shift σ_{k+1} can be selected as

$$(23) \quad \sigma_{k+1} = \arg \max_{s \in S} \|C_m P_m(s)^{-1}\|_2 \|R_B(s)\|_2.$$

We notice here that another simple way of choosing the shifts is to consider in (23) only the second part, which gives

$$(24) \quad \sigma_{k+1} = \arg \max_{s \in S} \|R_B(s)\|_2;$$

however, for some problems, choosing the interpolation points by using (23) gives more accurate results than those obtained with (24). The adaptive order rational block Lanczos algorithm for the computation of the reduced-order system using the modified rational block Lanczos process (Algorithm 2) and the adaptive approach given by (23) for selecting the interpolation points can be summarized as follows.

ALGORITHM 3. The Adaptive Modified Rational Block Lanczos (AMRBL) algorithm for model-order reduction

Input: The original system (A, B, C) , the initial values $\sigma_1 = \tilde{\sigma}_1$, choose a tolerance tol and set $H_0 = I_p$

Output: The reduced system (A_m, B_m, C_m)

Initialize: $\Sigma = \{\sigma_1\}$; $\tilde{\Sigma} = \{\tilde{\sigma}_1\}$; $m_1 = \tilde{m}_1 = 3$; $\hat{\varepsilon} = 1$ and $K = 1$

While ($\hat{\varepsilon} > tol$) **do**

$[\mathbb{V}_m, \mathbb{W}_m] = \mathbf{Modified_rational_Block_Lanczos}(A, B, C, \Sigma, \tilde{\Sigma});$

Compute the reduced model $A_m = \mathbb{W}_m^T A \mathbb{V}_m$, $B_m = \mathbb{W}_m^T B$, $C_m = C \mathbb{V}_m$ and the corresponding transfer function H_m ;

Compute the next interpolation point $\sigma_{K+1} = \tilde{\sigma}_{K+1}$ using (23);
 $\Sigma = \{\Sigma, \sigma_{K+1}\}$; $\tilde{\Sigma} = \{\tilde{\Sigma}, \tilde{\sigma}_{K+1}\}$; $m_{K+1} = \tilde{m}_{K+1} = 3$;
 Compute the error estimation $\hat{\varepsilon}_m = \|H_m - H_{m-1}\|_\infty$;
 $K = K + 1$

end.

REMARK. To choose the interpolation points, we can also use one of the error approximation expressions listed in Table 1. In this case the interpolation points are selected to be the frequencies $\sigma \in \Sigma$ and $\tilde{\sigma} \in \tilde{\Sigma}$ at which one of the approximated error expressions achieves its maximum, i.e.,

$$\Sigma = \{\sigma : |\hat{\varepsilon}_m(\sigma)| = \|\hat{\varepsilon}_m\|_\infty\} \quad \text{and} \quad \tilde{\Sigma} = \{\tilde{\sigma} : |\hat{\varepsilon}_m(\tilde{\sigma})| = \|\hat{\varepsilon}_m\|_\infty\}.$$

6. Numerical results. In this section, we give some experimental results to show the effectiveness of the adaptive modified rational block Lanczos (AMRBL) algorithm for model reduction in large LTI dynamical systems. All the experiments were performed on an Intel Core i5 computer at 1.3GHz and 8Gb of RAM. The algorithms were coded in Matlab 8.0. In all the experiments, we used $tol = 10^{-5}$, and the while-loop in Algorithm 3 is stopped when

$$\hat{\varepsilon}_m = \|H_m - H_{m-1}\|_\infty < tol$$

where

$$\|H_m - H_{m-1}\|_\infty = \sup_{\omega \in \mathbb{R}} \|H_m(j\omega) - H_{m-1}(j\omega)\|_2,$$

with $\omega \in [10^{-6}, 10^6]$ and $j = \sqrt{-1}$.

In all the experiments, we considered the special case where the sequences of shifts $\{\sigma_i\}_{i=1}^m$ and $\{\tilde{\sigma}_i\}_{i=1}^m$ are equal. To compute the set of frequencies ω , we used the function `lp_lgfrq` from LYAPACK [32]. This function generates a set of logarithmically distributed frequency sampling points.

EXAMPLE 1. The first model used in this example is the modified FOM model from [29]. We notice that originally, the FOM model is a SISO system of order $n = 1006$, and we modified the inputs and outputs to get a MIMO system. The matrices B and C are then given by

$$B = [b_1, \dots, b_6], \quad C^T = [c_1, \dots, c_6],$$

where $b_1^T = c_1 = (\underbrace{10, \dots, 10}_6, \underbrace{1, \dots, 1}_{1000})$ and $b_2, \dots, b_6, c_2, \dots, c_6$ are random column vectors. We applied AMRBL to get a reduced order model of dimension $m = 40$.

In the second experiment, we considered the International Space Station (ISS) model from [29]. This system is of dimension $n = 270$, with three inputs and three outputs. The reduced second order system was of dimension $m = 45$. The left plots of Figure 1 (modified FOM) and Figure 2 (ISS) show

the frequency responses of the original system (circles) compared with the frequency responses of its approximation (solid plot). The right plots of these figures represent the exact error $\|H(j\omega) - H_m(j\omega)\|_2$ for different frequencies $\omega \in [10^{-6}, 10^6]$.

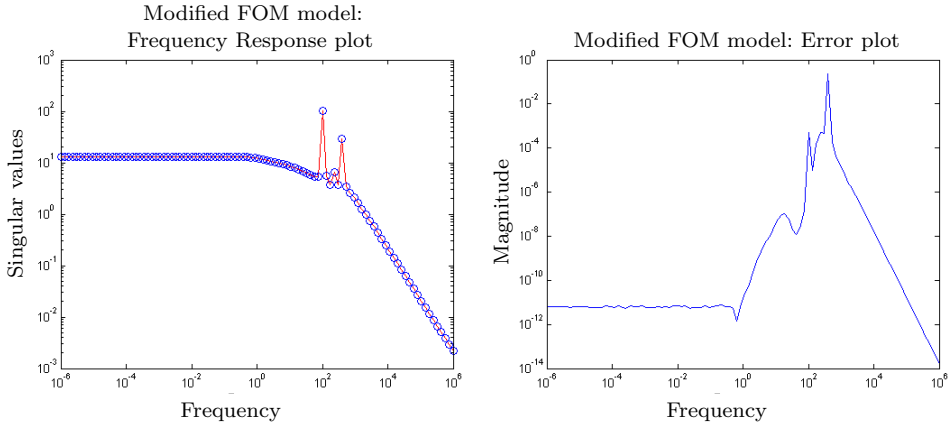


Fig. 1. Left: $\|H(j\omega)\|_2$ and its approximations $\|H_m(j\omega)\|_2$. Right: the exact error $\|H(j\omega) - H_m(j\omega)\|_2$ for the modified FOM model with $m = 40$.

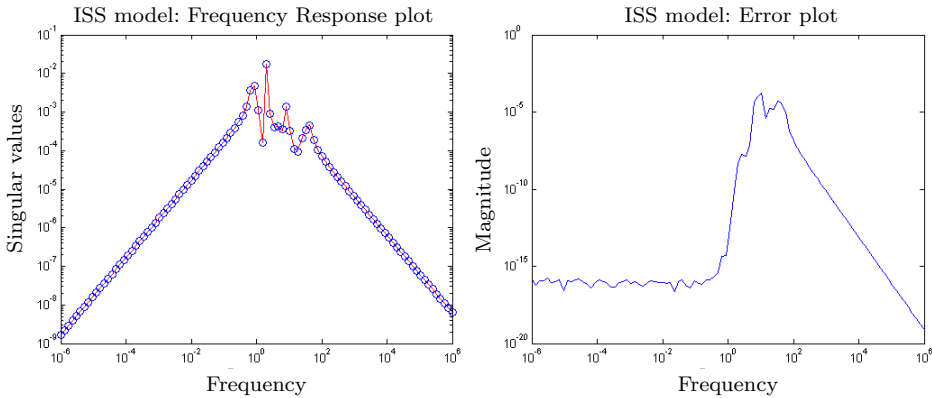


Fig. 2. Left: $\|H(j\omega)\|_2$ and its approximations $\|H_m(j\omega)\|_2$. Right: the exact error $\|H(j\omega) - H_m(j\omega)\|_2$ for the ISS circuit model with $m = 45$.

EXAMPLE 2. For this experiment, we considered the **fdm** [32] and the **Rail821** [29] models. We plotted the H_∞ relative error norm

$$\frac{\|H - H_m\|_\infty}{\|H\|_\infty}$$

versus the number m of iterations. 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

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

The matrices B and C were random matrices with entries uniformly distributed in $[0, 1]$. The number of inner grid points in each direction was $n_0 = 100$, and the dimension of A was $n = n_0^2 = 10.000$. For this experiment, we used $p = 6$. The **Rai1821** model is a first-order system of dimension $n = 821$ and $p = 6$. As can be seen from Figure 3, the relative error decreases rapidly to zero.

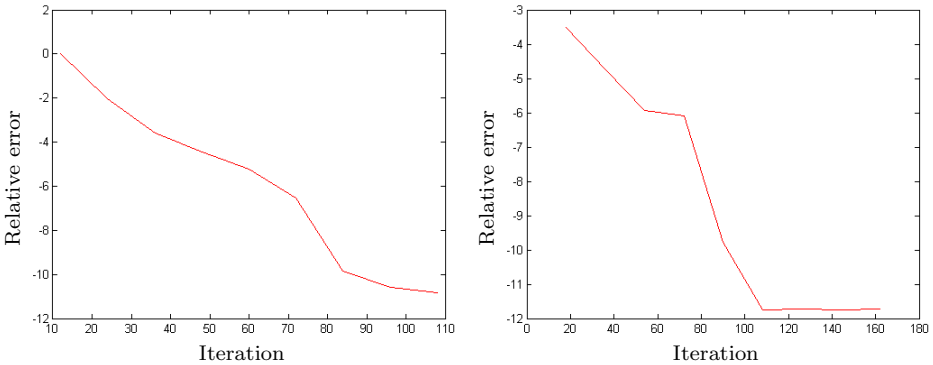


Fig. 3. Relative error norms. Left: the **Rai1821** model; right: the **fdm** model.

In Table 2 we report the results obtained with the different matrix tests. In this table, we list the exact H_∞ -error norm, the corresponding number of iterations (# Itrs.) and the cpu-time obtained.

Table 2. The exact H_∞ -error for different matrix tests

Matrices	# Itrs.	$\ H - \hat{H}\ _\infty$
Rai1821	16	2.12×10^{-11}
Rai13113	26	5.32×10^{-9}
fdm , $n = 10.000$, $p = 6$	40	30.66×10^{-9}

EXAMPLE 3. In the last example, we compare the exact H_∞ -error with different approximations using the methods described in this paper for

choosing the interpolation points (Table 1). For this experiment, the matrix test was Rai1821. As seen in Table 3, the results are similar when using the

Table 3. Results with the Rai1821 model

Error expressions	# Itrs.	$\ H - H_m\ _\infty$
$\Sigma_1 = \{\sigma : \ \tilde{R}_B(\sigma)\ _2 = \ \tilde{R}_B\ _\infty\}$	16	4.5×10^{-10}
$\Sigma_2 = \{\sigma : \ \tilde{R}_C^T(\sigma)\ _2 = \ \tilde{R}_C^T\ _\infty\}$	14	5.2×10^{-11}
$\Sigma_3 = \{\sigma : \ \tilde{H}_m(\sigma)\tilde{R}_B(\sigma)\ _2 = \ \tilde{H}_m\tilde{R}_B\ _\infty\}$	16	1.8×10^{-11}
$\Sigma_4 = \{\sigma : \ \tilde{R}_C^T(\sigma)\tilde{H}_m(\sigma)\ _2 = \ \tilde{R}_C^T\tilde{H}_m\ _\infty\}$	16	2.5×10^{-10}
$\Sigma_5 = \{\sigma : \ \tilde{H}_m(\sigma)\ _2 = \ \tilde{H}_m\ _\infty\}$	50	9.6×10^{-10}
Σ_6	16	2.0×10^{-11}
Σ_7	16	5.3×10^{-11}

different proposed approaches for selecting the shifts except for the set Σ_5 for which one needs many iterations to get a good approximation. Therefore, we can choose simple sets such as Σ_1 or Σ_2 to get good interpolation points that could be used in the adaptive modified rational block Lanczos algorithm with $\Sigma_6 = \{\sigma : \|\tilde{R}_C^T(\sigma)\tilde{H}_m(\sigma)\tilde{R}_B(\sigma)\|_2 = \|\tilde{R}_C^T\tilde{H}_m\tilde{R}_B\|_\infty\}$ and $\Sigma_7 = \{\sigma : \|C_m P_m^{-1}(\sigma)\|_2 \|R_B(\sigma)\|_2 = \|C_m P_m^{-1}\|_\infty \|R_B\|_\infty\}$.

References

- [1] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, Adv. Design and Control, SIAM, Philadelphia, 2005.
- [2] A. C. Antoulas, D. C. Sorensen and S. Gugercin, *A survey of model reduction methods for large scale systems*, in: Contemp. Math. 280, Amer. Math. Soc., 2001, 193–219.
- [3] Z. Bai, D. Day and Q. Ye, *ABLE: An adaptive block Lanczos method for non-Hermitian eigenvalue problems*, SIAM J. Matrix Anal. Appl. 20 (1999), 1060–1082.
- [4] H. Barkouki, A. H. Bentbib and K. Jbilou, *An adaptive rational block Lanczos-type algorithm for model reduction of large scale dynamical systems*, J. Sci. Comput. 67 (2016), 221–236.
- [5] C. A. Beattie and S. Gugercin, *Krylov-based minimization for optimal H_2 model reduction*, in: 46th IEEE Conference on Decision and Control, 2007, 4385–4390.
- [6] A. Bodendiek and M. Bollhöfer, *Adaptive-order rational Arnoldi-type methods in computational electromagnetism*, BIT 54 (2013), 357–380.
- [7] C. Brezinski, *The block Lanczos and Vorobyev methods*, C. R. Math. Acad. Sci. Paris 331 (2000), 137–142.
- [8] C. Brezinski, *Model reduction by state space projection*, Comm. Appl. Anal. 8 (2004), 323–333.
- [9] A. Bultheel, *Laurent Series and Their Padé Approximations*, Birkhäuser, Basel, 1987.
- [10] C. De Villemaigne and R. Skelton, *Model reduction using a projection formulation*, Int. J. Control 40 (1987), 2141–2169.

- [11] V. Druskin, V. Simoncini and M. Zaslavsky, *Adaptive rational Krylov subspaces for large-scale dynamical systems*, technical report, Dipartimento di Matematica, Univ. di Bologna, 2010.
- [12] A. El Guennouni, K. Jbilou and H. Sadok, *The block Lanczos method for multiple linear systems*, Appl. Numer. Math. 51 (2004), 243–256.
- [13] P. Feldmann and R. W. Freund, *Reduced-order modeling of large linear subcircuits via a block Lanczos algorithm*, in: IEEE/ACM Proc. DAC, 1995, 474–479.
- [14] M. Frangos and I. M. Jaimoukha, *Adaptive rational interpolation: Arnoldi and Lanczos-like equations*, Eur. J. Control 14 (2008), 342–354.
- [15] M. Frangos and I. Jaimoukha, *Adaptive rational Krylov algorithms for model reduction*, in: Proc. Eur. Control Conference, 2007, 4179–4186.
- [16] K. Gallivan, E. Grimme and P. Van Dooren, *A rational Lanczos algorithm for model reduction*, Numer. Algorithms 12 (1996), 33–63.
- [17] K. Gallivan, A. Vandendorpe and P. Van Dooren, *Sylvester equations and projection-based model reduction*, J. Comp. Appl. Math. 162 (2004), 213–229.
- [18] K. Glover, *All optimal Hankel-norm approximation of linear multivariable systems and their L^∞ -error bounds*, Int. J. Control 39 (1984), 1115–1193.
- [19] G. H. Golub and R. R. Underwood, *The block Lanczos method for computing eigenvalues*, in: Mathematical Software 3, J. R. Rice (ed.), Academic Press, New York, 1977, 364–377.
- [20] W. B. Gragg and A. Lindquist, *On the partial realization problem*, Linear Algebra Appl. 50 (1983), 277–319.
- [21] E. Grimme, *Krylov projection methods for model reduction*, PhD thesis, ECE Dept., Univ. of Illinois, Urbana-Champaign, 1997.
- [22] E. Grimme and K. Gallivan, *A rational Lanczos algorithm for model reduction ii: Interpolation point selection*, Numer. Algorithms 12 (1998), 33–63.
- [23] S. Gugercin and A. C. Antoulas, *Model reduction of large scale systems by least squares*, Linear Algebra Appl. 415 (2006), 290–321.
- [24] S. Gugercin, A. Antoulas and C. Beattie, *H2-model reduction for large-scale linear systems*, SIAM J. Matrix Anal. Appl. 30 (2008), 609–638.
- [25] M. Heyouni, K. Jbilou, A. Messaoudi and K. Tabaa, *Model reduction in large scale MIMO dynamical systems via the block Lanczos method*, Comput. Appl. Math. 27 (2008), 211–236.
- [26] A. Köhler and S. Reitzinger, *An adaptive multi-point multi-moment model order reduction algorithm for fast broadband simulation of large-scale 3D electromagnetic models*, in: ANALOG 2010, Entwicklung von Analogschaltungen mit CAE-Methoden, VDE-Verlag, Berlin, 2010, 39–52.
- [27] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Nat. Bur. Standards 45 (1950), 225–280.
- [28] H.-J. Lee, C.-C. Chu and W.-S. Feng, *An adaptive-order rational Arnoldi method for model-order reductions of linear time-invariant systems*, Linear Algebra Appl. 415 (2006), 235–261.
- [29] V. Mehrmann and T. Penzl, *Benchmark collections in SLICOT*, technical report, <http://slicot.org/working-notes/wgs-niconet-reports/7-benchmark-collections-in-slicot>, 1998.
- [30] B. C. Moore, *Principal component analysis in linear systems: controllability, observability, and model reduction*, IEEE Trans. Automat. Control 26 (1981), 17–32.
- [31] D. O’Leary, *The block conjugate gradient algorithm and related methods*, Linear Algebra Appl. 29 (1980), 293–322.

- [32] T. Penzl, *LYAPACK, a MATLAB toolbox for large Lyapunov and Riccati equations, model reduction problems, and linear-quadratic optimal control problems*, <http://www.netlib.org/lyapack/guide.pdf>.
- [33] A. Ruhe, *Rational Krylov algorithms for nonsymmetric eigenvalue problems II: matrix pairs*, *Linear Algebra Appl.* 197 (1994), 283–295.
- [34] A. Soppa, *Krylov-Unterraum basierte Modellreduktion zur Simulation von Werkzeugmaschinen*, PhD thesis, Technical Univ. Braunschweig, 2011.
- [35] P. Van Dooren, *The Lanczos algorithm and Padé approximations, short course*, in: *Benelux Meeting on Systems and Control*, 1995.
- [36] A. Yousouff and R. E. Skelton, *Covariance equivalent realizations with applications to model reduction of large-scale systems*, in: *Control Dynam. Systems* 22, C. T. Leondes (ed.), Academic Press, 1985, 273–348.

H. Barkouki, A. H. Bentbib
Faculté des Sciences et Techniques Guéliz
Laboratoire de Mathématiques Appliquées
et Informatique
Marrakech, Morocco
E-mail: houda.barkouki@gmail.com
a.bentbib@uca.ma

K. Jbilou (corresponding author)
Laboratoire LMPA
ULCO
50 rue F. Buisson
62228 Calais, France
E-mail: jbilou@lmpa.univ-littoral.fr