Volume 27, N. 2, pp. 211–236, 2008
Copyright © 2008 SBMAC
ISSN 0101-8205
www.scielo.br/cam

Model reduction in large scale MIMO dynamical systems via the block Lanczos method

M. HEYOUNI[1], K. JBILOU[1], A. MESSAOUDI[2] and K. TABAA[3]

[1]L.M.P.A, Université du Littoral, 50 rue F. Buisson BP 699, F-62228 Calais Cedex, France
[2]École Normale Supérieure Rabat, Département de Mathématiques, Rabat, Maroc
[3]Faculté des Sciences Agdal, Département de Mathématiques, Rabat, Maroc
E-mails: heyouni@lmpa.univ-littoral.fr / jbilou@lmpa.univ-littoral.fr /
abdou.messaoudi@caramail.com / tabaa.khalid@caramail.com

Abstract. In the present paper, we propose a numerical method for solving the coupled Lyapunov matrix equations

$$A\,P + P\,A^T + B\,B^T = 0 \quad \text{and} \quad A^T\,Q + Q\,A + C^T\,C = 0$$

where A is an $n \times n$ real matrix and B, C^T are $n \times s$ real matrices with $\text{rank}(B) = \text{rank}(C) = s$ and $s \ll n$. Such equations appear in control problems. The proposed method is a Krylov subspace method based on the nonsymmetric block Lanczos process. We use this process to produce low rank approximate solutions to the coupled Lyapunov matrix equations. We give some theoretical results such as an upper bound for the residual norms and perturbation results. By approximating the matrix transfer function $F(z) = C\,(z\,I_n - A)^{-1}\,B$ of a Linear Time Invariant (LTI) system of order n by another one $F_m(z) = C_m\,(z\,I_m - A_m)^{-1}\,B_m$ of order m, where m is much smaller than n, we will construct a reduced order model of the original LTI system. We conclude this work by reporting some numerical experiments to show the numerical behavior of the proposed method.

Mathematical subject classification: 65F10, 65F30.

Key words: coupled Lyapunov matrix equations; Krylov subspace methods; nonsymmetric block Lanczos process; reduced order model; transfer functions.

#746/07. Received: 01/VII/07. Accepted: 05/XII/07.

1 Introduction

Consider a stable linear multi-input multi-output (MIMO) state-space model of the form:

$$\left\{\begin{array}{lcl} \dot{x}(t) & = & A\,x(t) + B\,u(t) \\ y(t) & = & C\,x(t), \end{array}\right. \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B,\ C^T \in \mathbb{R}^{n\times s}$, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^s$ is the input vector and $y(t) \in \mathbb{R}^s$ is the output vector of the system (1.1). When dealing with high-order models, it is reasonable to look for an approximate stable model

$$\left\{\begin{array}{lcl} \dot{x}_m(t) & = & A_m\,x_m(t) + B_m\,u(t) \\ y_m(t) & = & C_m\,x_m(t), \end{array}\right. \tag{1.2}$$

in which $A_m \in \mathbb{R}^{m\times m}$, $B_m,\ C_m^T \in \mathbb{R}^{m\times s}$ and $x_m(t), y_m(t) \in \mathbb{R}^m$, with $m \ll n$. Hence, the reduction problem consists in approximating the triplet $\{A, B, C\}$ by another one $\{\hat{A}, \hat{B}, \hat{C}\}$ of small size. Several approaches in this area have been used as Padé approximation [15, 33, 34], balanced truncation [29, 37], optimal Hankel norm [16, 17] and Krylov subspace methods [3, 6, 11, 12, 21, 22]. These approaches require the solution of coupled Lyapunov matrix equations [1, 13, 25, 27] having the form

$$\left\{\begin{array}{l} A\,P + P\,A^T + B\,B^T = 0 \\ A^T\,Q + Q\,A + C^T\,C = 0, \end{array}\right. \tag{1.3}$$

where P, Q are the controllability and the observability Grammians of the system (1.1). For historical developments, applications and importance of Lyapunov equations and related problems, we refer to [10, 13] and the references therein. Throughout the paper, we will assume that $\lambda_i(A) + \bar{\lambda}_j(A) \neq 0$ for all $i, j = 1, \ldots, n$ where $\lambda_k(A)$ and it's conjugate $\bar{\lambda}_k(A)$ are eigenvalues of A. In this case, the equations (1.3) have unique solutions [26].

Direct methods for solving the Lyapunov matrix equations (1.3) such as those proposed in [5, 18, 24] are attractive if the matrices are of moderate size. These methods are based on the Schur or the Hessenberg decomposition. For large problems, several iterative methods have been proposed; see [14, 22, 23, 32]. These methods use Galerkin projection technics to produce low-dimensional Sylvester or Lyapunov matrix equations that are solved by using direct methods.

For the single-input single-output (SISO) case, i.e., $s = 1$, two approaches based on Arnoldi and Lanczos processes were proposed in [21, 22] to solve large Lyapunov matrix equations. The Arnoldi and Lanczos processes were also applied in order to give an approximate reduced order model to (1.1) [2, 6, 11, 13].

Our purpose in this paper is to describe a method based on the nonsymmetric block Lanczos process [4, 15, 19, 36] for solving the coupled Lyapunov matrix equations (1.3). In this method, we project the initial equations onto block Krylov subspaces generated by the block Lanczos process to produce low-dimensional Lyapunov matrix equations that are solved by direct methods. By approximating the transfer function $F(z) = C\,(z\,I_n - A)^{-1}\,B$ by another one $F_m(z) = C_m\,(z\,I_m - A_m)^{-1}\,B_m$ where $A_m \in \mathbb{R}^{m\times m}$, $B_m, C_m^T \in \mathbb{R}^{m\times s}$, and m is much smaller than n, we will construct an approximate reduced order model of the continuous time linear system (1.1).

The remainder of the paper is organized as follows. In the following section, we review the nonsymmetric block Lanczos process and give the exact solutions of the coupled Lyapunov matrix equations. In Section 3, we first show how to extract low rank approximate solutions to (1.3). Then, we give some theoretical results on the residual norms and demonstrate that the low rank approximate solutions are exact solutions to a pair of perturbed Lyapunov matrix equations. In Section 4, we consider the problem of obtaining reduced order models to LTI systems by approximating the associated transfer function. This approach is based on the nonsymmetric block Lanczos process. Finally, we will present some numerical experiments.

2 The block Lanczos method and coupled Lyapunov matrix equations

2.1 The nonsymmetric block Lanczos process

Let $V \in \mathbb{R}^{n\times s}$ and consider the block matrix Krylov subspace $K_m(A, V) = \text{span}\{V, A\,V, \ldots, A^{m-1}\,V\}$. Notice that $Z \in K_m(A, V)$ means that

$$Z = \sum_{i=0}^{m-1} A^i\,V\,\Omega_i, \quad \Omega_i \in \mathbb{R}^{s\times s},\ i = 0, \ldots m-1.$$

We also recall that the minimal polynomial $\mathcal{P}$ of A with respect to V is the nonzero monic polynomial of lowest degree q such that

$$\mathcal{P}(A) \circ V = \sum_{i=0}^{q} A^i \, V \, \Omega_i ,$$

where $\Omega_i \in \mathbb{R}^{s\times s}$ and $\Omega_q = I_s$. The grade of V denoted by $\text{grad}(V)$ is the degree of the minimal polynomial, hence $\text{grad}(V) = q$.

In the sequel, we suppose that given two matrices V, $W \in \mathbb{R}^{n\times s}$ of full rank, we compute initial block vectors V_1 and W_1 using the QR decomposition of $W^T \, V$. Hence, if $W^T \, V = \delta \, \beta$ where $\delta \in \mathbb{R}^{s\times s}$ is an orthogonal matrix (i.e., $\delta^T \, \delta = \delta \, \delta^T = I_s$) and $\beta \in \mathbb{R}^{s\times s}$ is an upper triangular matrix, then

$$V_1 = V \, \beta^{-1} \quad \text{and} \quad W_1 = W \, \delta. \tag{2.1}$$

Given an $n \times n$ matrix A and the initial $n \times s$ block vectors V, W, the block Lanczos process applied to the triplet (A, V, W) and described by Algorithm 1, generates sequences of $n \times s$ right and left block Lanczos vectors $\{V_1, \ldots, V_m\}$ and $\{W_1, \ldots, W_m\}$. These block vectors form biorthonormal bases of the block Krylov subspaces $K_m(A, V_1)$ and $K_m(A^T, W_1)$.

Algorithm 1. The nonsymmetric block Lanczos process [4]

- *Inputs :* A an $n \times n$ matrix, V, W two $n \times s$ matrices and m an integer.
- *Step 0*. Compute the QR decomposition of $W^T \, V$, i.e., $W^T \, V = \delta \, \beta$;
  $V_1 = V \, \beta^{-1}$; $W_1 = W \, \delta$; $\tilde{V}_2 = A \, V_1$; $\tilde{W}_2 = A^T \, W_1$;
- *Step 1*. For $j = 1, \ldots, m$
  $\alpha_j = W_j^T \, \tilde{V}_{j+1}$; $\tilde{V}_{j+1} = \tilde{V}_{j+1} - V_j \, \alpha_j$; $\tilde{W}_{j+1} = \tilde{W}_j - W_j \, \alpha_j^T$;
  Compute the QR decompositions of $\tilde{V}_{j+1}$ and $\tilde{W}_{j+1}$, i.e.,
  $\tilde{V}_{j+1} = V_{j+1} \, \beta_{j+1}$; $\tilde{W}_{j+1} = W_{j+1} \, \delta_{j+1}^T$;
  Compute the singular value decomposition of $W_{j+1}^T \, V_{j+1}$, i.e.,
  $W_{j+1}^T \, V_{j+1} = U_j \, \Sigma_j \, Z_j^T$;
  $\delta_{j+1} = \delta_{j+1} \, U_j \, \Sigma_j^{1/2}$; $\beta_{j+1} = \Sigma_j^{1/2} \, V_j^T \, \beta_{j+1}$;
  $V_{j+1} = V_{j+1} \, Z_j \, \Sigma_j^{-1/2}$; $W_{j+1} = W_{j+1} \, U_j \, \Sigma_j^{-1/2}$;
  $\tilde{V}_{j+2} = A \, V_{j+1} - V_j \, \delta_{j+1}$; $\tilde{W}_{j+2} = A^T \, W_{j+1} - W_j \, \beta_{j+1}^T$;
  end For.

Specifically, after m steps, the block Lanczos procedure determines two block matrices $\mathcal{V}_m = (V_1, \ldots, V_m) \in \mathbb{R}^{n \times ms}$, $\mathcal{W}_m = (W_1, \ldots, W_m) \in \mathbb{R}^{n \times ms}$ and an $ms \times ms$ block-tridiagonal matrix

$$\mathcal{T}_m = \begin{pmatrix} \alpha_1 & \delta_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \delta_m \\ & & \beta_m & \alpha_m \end{pmatrix}, \tag{2.2}$$

that satisfy the following relations

$$A\,\mathcal{V}_m = \mathcal{V}_m\,\mathcal{T}_m + V_{m+1}\,\beta_{m+1}\,E_m^T = \mathcal{V}_m\,\mathcal{T}_m + \tilde{V}_{m+1}\,E_m^T, \tag{2.3}$$

$$A^T\,\mathcal{W}_m = \mathcal{W}_m\,\mathcal{T}_m^T + W_{m+1}\,\delta_{m+1}^T\,E_m^T = \mathcal{W}_m\,\mathcal{T}_m^T + \tilde{W}_{m+1}\,E_m^T, \tag{2.4}$$

and

$$\mathcal{W}_m^T\,A\,\mathcal{V}_m = \mathcal{T}_m, \tag{2.5}$$

$$\mathcal{W}_m^T\,\mathcal{V}_m = I_{ms}, \tag{2.6}$$

where $E_m^T = (0_s, \ldots, 0_s, I_s) \in \mathbb{R}^{s \times ms}$.

Notice that, a breakdown may occur in Algorithm 1 if $\tilde{W}_j^T\,\tilde{V}_j$ is singular, or if $\tilde{V}_j$ or $\tilde{W}_j$ is not full rank. In [4], several strategies are proposed to overcome breakdowns and near-breakdowns in order to preserve the numerical stability of the block Lanczos process for eigenvalue problems. In the sequel, we assume that $m \leq \min\{q, r\}$ where q and r are the degrees of the minimal polynomials of A with respect to V and of A^T with respect to W respectively.

2.2 Exact solutions of coupled Lyapunov matrix equations

Before deriving exact expressions for the solutions of the coupled Lyapunov matrix equations, let us give the following result which will be useful in the sequel.

Lemma 2.1. *Let $q = \mathrm{grad}(B)$, $r = \mathrm{grad}(C^T)$ and $m \leq \min\{q, r\}$. Assume that m steps of Algorithm 1 are applied to the triplet (A, B, C^T), then we have*

$$A^j\,V_1 = \mathcal{V}_m\,\mathcal{T}_m^j\,E_1, \qquad \textit{for } j = 0, \ldots, m-1, \tag{2.7}$$

$$(A^T)^j\,W_1 = \mathcal{W}_m\,(\mathcal{T}_m^T)^j\,E_1, \qquad \textit{for } j = 0, \ldots, m-1, \tag{2.8}$$

where $E_1^T = (I_s, 0_s, \dots, 0_s) \in \mathbb{R}^{s\times ms}$.

Proof. Since $\mathcal{T}_m$ is a block tridiagonal matrix, then for $0 \le j \le m-2$, $\mathcal{T}_m^j$ is a block band matrix with j upper-diagonals and j lower-diagonals. Therefore, letting $E_m^T = (0_s, \dots, 0_s, I_s)$ we have

$$E_m^T\,\mathcal{T}_m^j\,E_1 = E_m^T\,(\mathcal{T}_m^T)^j\,E_1 = 0, \qquad \text{for } j = 0, \dots, m-2.$$

Using this last relation and the fact that $E_m^T\,E_1 = 0$, we can proof (2.7) and (2.8) by induction.

$$\begin{aligned} A^{j+1}\,V_1 &= A\,(A^j\,V_1) = A\,(\mathcal{V}_m\,\mathcal{T}_m^j\,E_1) \\ &= (\mathcal{V}_m\,\mathcal{T}_m + \tilde{V}_{m+1}\,E_m^T)\,\mathcal{T}_m^j\,E_1 = \mathcal{V}_m\,\mathcal{T}_m^{j+1}\,E_1. \end{aligned}$$

Using the same arguments, we obtain (2.8). □

Using the previous lemma, we next give the low rank approximate solutions to (1.3). Let $\mathcal{Q}_q$ be the minimal polynomial of A with respect to B and $q = \text{grad}(B)$, i.e.,

$$\mathcal{Q}_q(A)\circ B = \sum_{i=0}^{q} A^i\,B\,\Omega_i = 0, \quad \text{with} \quad \Omega_i \in \mathbb{R}^{s\times s} \quad \text{and} \quad \Omega_q = I_s,$$

and let $\mathcal{R}_r$ be the minimal polynomial of A^T for C^T where $r = \text{grad}(C^T)$, i.e.,

$$\mathcal{R}_r\big(A^T\big)\circ C^T = \sum_{i=0}^{r} \big(A^T\big)^i\,C^T\,\Psi_i = 0, \quad \text{with} \quad \Psi_i \in \mathbb{R}^{s\times s} \quad \text{and} \quad \Psi_r = I_s.$$

Define

$$K_q = \begin{pmatrix} 0_s & 0_s & \cdots & 0_s & -\Omega_0 \\ I_s & 0_s & \ddots & \vdots & -\Omega_1 \\ 0_s & \ddots & \ddots & 0_s & \vdots \\ \vdots & \ddots & \ddots & 0_s & \vdots \\ 0_s & \dots & 0_s & I_s & -\Omega_{q-1} \end{pmatrix} \text{ and } K_r = \begin{pmatrix} 0_s & 0_s & \cdots & 0_s & -\Psi_0 \\ I_s & 0_s & \ddots & \vdots & -\Psi_1 \\ 0_s & \ddots & \ddots & 0_s & \vdots \\ \vdots & \ddots & \ddots & 0_s & \vdots \\ 0_s & \dots & 0_s & I_s & -\Psi_{r-1} \end{pmatrix},$$

as the block companion polynomials of $\mathcal{Q}_q$ and $\mathcal{R}_r$. Denote by M_q and N_r the following Krylov matrices

$$M_q = \big(B, A\,B, \dots, A^{q-1}\,B\big), \quad N_r = \big(C^T, A^T C^T, \dots, \big(A^T\big)^{r-1} C^T\big).$$

Then

$$A\,M_q = M_q\,K_q, \quad \text{and} \quad A^T\,N_r = N_r\,K_r. \tag{2.9}$$

We now give the following theorem

Theorem 2.2. *The general solutions P and Q of the coupled Lyapunov matrix equation* (1.3) *are given by*

$$P = M_q\,X\,M_q^T, \tag{2.10}$$

$$Q = N_r\,Y\,N_r^T, \tag{2.11}$$

where $X \in \mathbb{R}^{qs\times qs}$ and $Y \in \mathbb{R}^{rs\times rs}$ satisfy

$$K_q\,X + X\,K_q^T + E_1\,E_1^T = 0, \tag{2.12}$$

$$K_r\,Y + Y\,K_r^T + \bar{E}_1\,\bar{E}_1^T = 0, \tag{2.13}$$

with $E_1^T = \big(I_s, 0_s, \ldots, 0_s\big) \in \mathbb{R}^{s\times qs}$ *and* $\bar{E}_1^T = \big(I_s, 0_s, \ldots, 0_s\big) \in \mathbb{R}^{s\times rs}$.

Proof. Premultiplying (2.10) by A, postmultiplying (2.11) by A^T, using (2.9) and the fact that $B = M_q\,E_1$, we have

$$\begin{aligned}
A\,P + P\,A^T + B\,B^T &= A\,M_q\,X\,M_q^T + M_q\,X\,M_q^T\,A^T + B\,B^T \\
&= M_q\,K_q\,X\,M_q^T + M_q\,X\,K_q^T\,M_q^T + M_q\,E_1\,E_1^T\,M_q^T \\
&= M_q\,(K_q\,X + X\,K_q^T + E_1\,E_1^T)M_q^T \\
&= 0.
\end{aligned}$$

By the same way, we prove (2.11). □

Again using Lemma 2.1, we have the following result which states that the solutions of (1.3) could be obtained by using the block Lanczos process.

Theorem 2.3. *Suppose that $l = \mathrm{grad}(B) = \mathrm{grad}(C^T)$ and assume that l steps of Algorithm* 1 *have been run. Then, the solutions P and Q of the coupled Lyapunov matrix equations* (1.3) *could be expressed as follows*

$$P = \mathcal{V}_l\,\Gamma\,\mathcal{V}_l^T, \tag{2.14}$$

$$Q = \mathcal{W}_l\,\hat{\Gamma}\,\mathcal{W}_l^T, \tag{2.15}$$

where Γ *and* $\hat{\Gamma}$ *are the solutions of the following reduced Lyapunov matrix equations*

$$\mathcal{T}_l\,\Gamma + \Gamma\,\mathcal{T}_l^T + E_1\,\beta\,\beta^T\,E_1^T \;=\; 0, \tag{2.16}$$

$$\mathcal{T}_l^T\,\hat{\Gamma} + \hat{\Gamma}\,\mathcal{T}_l + E_1\,E_1^T \;=\; 0, \tag{2.17}$$

and $E_1^T = \left(I_s, 0_s, \ldots, 0_s\right) \in \mathbb{R}^{s\times ls}$.

Proof. Since $B = V_1\,\beta$, the general solution P of the first Lyapunov equation in (1.3) can be expressed as follows

$$\begin{aligned} P &= M_l\,X\,M_l^T = \left(B,\; A\,B, \ldots, A^{l-1}\,B\right) X \begin{pmatrix} B^T \\ B^T\,A^T \\ \vdots \\ B^T\left(A^T\right)^{l-1} \end{pmatrix}, \\ &= \left(V_1,\; A\,V_1, \ldots, A^{l-1}\,V_1\right)\beta\,X\,\beta^T \begin{pmatrix} V_1^T \\ V_1^T\,A^T \\ \vdots \\ V_1^T\left(A^T\right)^{l-1} \end{pmatrix}. \end{aligned}$$

Using (2.7), we get

$$P = \mathcal{V}_l\left(E_1, \mathcal{T}_l\,E_1, \ldots, \mathcal{T}_l^{l-1}\,E_1\right)\beta\,X\,\beta^T \begin{pmatrix} E_1^T \\ E_1^T\,\mathcal{T}_l^T \\ \vdots \\ E_1^T\left(\mathcal{T}_l^T\right)^{l-1} \end{pmatrix} \mathcal{V}_l^T,$$

where $E_1^T = \left(I_s, 0_s, \ldots, 0_s\right)$ is an $s \times ls$ matrix. Setting

$$\Gamma = \left(E_1, \mathcal{T}_l\,E_1, \ldots, \mathcal{T}_l^{l-1}\,E_1\right)\beta\,X\,\beta^T \begin{pmatrix} E_1^T \\ E_1^T\mathcal{T}_l^T \\ \vdots \\ E_1^T\left(\mathcal{T}_l^T\right)^{l-1} \end{pmatrix},$$

then $P = \mathcal{V}_l \, \Gamma \, \mathcal{V}_l^T$ and we have

$$\begin{aligned} A\,P + P\,A^T + B\,B^T &= A\,\mathcal{V}_l\,\Gamma\,\mathcal{V}_l^T + \mathcal{V}_l\,\Gamma\,\mathcal{V}_l^T A^T + B\,B^T, \\ &= A\,\mathcal{V}_l\,\Gamma\,\mathcal{V}_l^T + \mathcal{V}_l\,\Gamma\,\mathcal{V}_l^T\,A^T + V_1\,\beta\,\beta^T\,V_1^T, \\ &= 0. \end{aligned}$$

Multiplying the last equalities on the right by $\mathcal{W}_l$, on the left by $\mathcal{W}_l^T$ and using (2.5) and (2.6) with $m = l$ we obtain (2.16).

Note that (2.17) is obtained as above by using (2.8), (2.11) and the fact that $C^T = W_1\,\delta^T$ and $\delta^T\,\delta = \delta\,\delta^T = I_s$. $\square$

Before ending this section, we have to say that in general $\mathrm{grad}(B) \neq \mathrm{grad}(C^T)$. Hence, if $l = \min\{\mathrm{grad}(B), \mathrm{grad}(C^T)\}$ and l steps of the nonsymmetric block Lanczos process have been run, then only (2.14) and (2.16) are satisfied if $l = \mathrm{grad}(B)$. Similarly, only (2.15) and (2.17) are satisfied if $l = \mathrm{grad}(C^T)$.

3 Solving the coupled Lyapunov matrix equations by the block Lanczos process

Let $q = \mathrm{grad}(B)$, $r = \mathrm{grad}(C^T)$ and $m \leq \min\{q, r\}$. Assuming that m steps of the nonsymmetric block Lanczos process have been run, we show how to extract low rank approximate solutions of the coupled Lyapunov matrix equations (1.3).

The results given in the previous section show that the matrices given below could be considered as approximate solutions to (1.3).

$$P_m = \mathcal{V}_m\,X_m\,\mathcal{V}_m^T, \tag{3.1}$$

$$Q_m = \mathcal{W}_m\,Y_m\,\mathcal{W}_m^T, \tag{3.2}$$

where X_m and Y_m are solutions of the following reduced Lyapunov equations

$$\mathcal{T}_m\,X_m + X_m\,\mathcal{T}_m^T + E_1\,\beta\,\beta^T\,E_1^T = 0, \tag{3.3}$$

$$\mathcal{T}_m^T\,Y_m + Y_m\,\mathcal{T}_m + E_1\,E_1^T = 0, \tag{3.4}$$

and $E_1 = \left(I_s, 0_s, \ldots, 0_s\right)^T \in \mathbb{R}^{s \times ms}$.

The low-dimensional Lyapunov equations (3.3) and (3.4) could be solved by direct methods such those described in [5, 18, 24]. In the sequel, we assume that the eigenvalues $\lambda_i(\mathcal{T}_m)$ of the block tridiagonal matrix $\mathcal{T}_m$ constructed by the nonsymmetric block Lanczos process satisfy $\lambda_i(\mathcal{T}_m) + \bar{\lambda}_j(\mathcal{T}_m) \neq 0$, for $i, j = 1, \ldots, m$. This condition ensures the existence and uniqueness of X_m and Y_m the solutions of the reduced Lyapunov equations and that these solutions are symmetric and positive semidefinite.

Next, we show how to compute an upper-bound for the Frobenius residual norms in order to use it as a stopping criterion. Notice that the upper bound given below will allow us to stop the algorithm without having to compute the approximate solutions P_m and Q_m. Hence, letting

$$R(P_m) = A\,P_m + P_m\,A^T + B\,B^T, \tag{3.5}$$

$$R(Q_m) = A^T\,Q_m + Q_m\,A + C^T\,C, \tag{3.6}$$

be the residuals associated to P_m and Q_m respectively, we have the following result

Theorem 3.1. *Let P_m, Q_m be the approximate solutions defined by* (3.1)*,* (3.3) *and* (3.2)*,* (3.4) *respectively. Let $R(P_m)$ and $R(Q_m)$ be the corresponding residuals defined by* (3.5) *and* (3.6) *respectively. Then*

$$\begin{aligned} &\|R(P_m)\|_F \leq 2\,\|\tilde{V}_{m+1}\,\tilde{X}_m\,\mathcal{V}_m^T\|_F \quad \textit{and} \\ &\|R(Q_m)\|_F \leq 2\,\|\tilde{W}_{m+1}\,\tilde{Y}_m\,\mathcal{W}_m^T\|_F \end{aligned} \tag{3.7}$$

where $\tilde{X}_m$, $\tilde{Y}_m$ are the $s \times n$ matrices corresponding to the last s rows of X_m and Y_m respectively.

Proof. From (3.1) and (3.3), we have

$$R(P_m) = A\mathcal{V}_m\,X_m\mathcal{V}_m^T + \mathcal{V}_m\,X_m\mathcal{V}_m^T\,A^T + B\,B^T,$$

then using (2.3) and the fact that $B = V_1\,\beta$, we get

$$\begin{aligned} R(P_m) &= \left(\mathcal{V}_m\,T_m + V_{m+1}\,\beta_{m+1}\,E_m^T\right)X_m\,\mathcal{V}_m^T \\ &\quad + \mathcal{V}_m\,X_m\left(T_m^T\,\mathcal{V}_m^T + E_m\,\beta_{m+1}^T\,V_{m+1}^T\right) + V_1\,\beta\,\beta^T\,V_1^T \\ &= \mathcal{V}_{m+1}\begin{pmatrix} T_m\,X_m \\ \beta_{m+1}\,E_m^T\,X_m \end{pmatrix}\mathcal{V}_m^T \end{aligned}$$

$$+ \mathcal{V}_m \begin{pmatrix} X_m\, T_m^T & X_m\, E_m\, \beta_{m+1}^T \end{pmatrix} \mathcal{V}_{m+1}^T + V_1\, \beta\, \beta^T\, V_1^T$$

$$= \mathcal{V}_{m+1} \begin{pmatrix} T_m\, X_m + X_m\, T_m^T + E_1\, \beta\, \beta^T\, E_1^T & X_m\, E_m\, \beta_{m+1}^T \\ \beta_{m+1}\, E_m^T\, X_m & 0 \end{pmatrix} \mathcal{V}_{m+1}^T.$$

Since X_m is the solution of the reduced Lyapunov equation (3.3) and $\tilde{V}_{m+1} = V_{m+1}\, \beta_{m+1}$, then

$$\begin{aligned} R(P_m) &= \mathcal{V}_{m+1} \begin{pmatrix} 0 & X_m\, E_m\, \beta_{m+1}^T \\ \beta_{m+1}\, E_m^T\, X_m & 0 \end{pmatrix} \mathcal{V}_{m+1}^T \\ &= V_{m+1}\, \beta_{m+1}\, E_m^T\, X_m\, \mathcal{V}_m^T + \mathcal{V}_m\, X_m\, E_m\, \beta_{m+1}^T\, V_{m+1}^T \\ &= \tilde{V}_{m+1}\, E_m^T\, X_m\, \mathcal{V}_m^T + \mathcal{V}_m\, X_m\, E_m\, \tilde{V}_{m+1}^T. \end{aligned}$$

As X_m is a symmetric matrix, it follows that

$$\| R(P_m) \|_F \leq 2\, \| \tilde{V}_{m+1}\, E_m^T\, X_m\, \mathcal{V}_m^T \|_F \leq 2\, \left\| \tilde{V}_{m+1}\, \tilde{X}_m\, \mathcal{V}_m^T \right\|_F,$$

where $\tilde{X}_m = E_m^T\, X_m$ represents the s last rows of X_m.

Similarly, from (2.4) and as $C^T = W_1\, \delta^T$, $\tilde{W}_{m+1} = W_{m+1}\, \delta_{m+1}^T$ and the fact that Y_m is symmetric and is the solution of the reduced Lyapunov equation (3.4), we obtain the second inequality of (3.7). □

To reduce the cost in the coupled Lyapunov block Lanczos method, the solution of the low-order Lyapunov equations are computed every k_0 iterations where k_0 is a chosen parameter. Note also that the approximate solutions are computed only when

$$r_m := 2\, \left\| \tilde{V}_{m+1}\, \tilde{X}_m\, \mathcal{V}_m^T \right\|_F \leq \epsilon \quad \text{and} \quad s_m = 2\, \left\| \tilde{W}_{m+1}\, \tilde{Y}_m\, \mathcal{W}_m^T \right\|_F \leq \epsilon,$$

where ϵ is a chosen tolerance. Summarizing the previous results, we get the following algorithm

Algorithm 2. The coupled Lyapunov block Lanczos algorithm (CLBL)

- *Inputs :* A an $n \times n$ stable matrix, B an $n \times s$ matrix and C an $s \times n$ matrix.
- *Step 0*. Choose a tolerance $\epsilon > 0$, an integer parameter k_0 and set $k = 0$; $m = k_0$;

- *Step 1*. For $j = k+1, k+2, \ldots, m$;
  construct the block tridiagonal matrix $\mathcal{T}_m$, the biorthonormal bases $\{V_{k+1}, \ldots, V_m\}$, $\{W_{k+1}, \ldots, W_m\}$, $\tilde{V}_{m+1}$ and $\tilde{W}_{m+1}$ by Algorithm 1 applied to the triplet (A, B, C^T);
  end For

- *Step 2*. Solve the low-dimensional Lyapunov equations:

$$\mathcal{T}_m X_m + X_m \mathcal{T}_m^T + E_1 \beta \beta^T E_1^T = 0;$$
$$\mathcal{T}_m^T Y_m + Y_m \mathcal{T}_m + E_1 E_1^T = 0;$$

- *Step 3*. Compute the upper bounds for the residual norms:

$$r_m = 2 \left\| \tilde{V}_{m+1} \tilde{X}_m \mathcal{V}_m^T \right\|_F \quad \text{and} \quad s_m = 2 \left\| \tilde{W}_{m+1} \tilde{Y}_m \mathcal{W}_m^T \right\|_F;$$

- *Step 4*. If $r_m > \epsilon$ or $s_m > \epsilon$, set $k = k + k_0$; $m = k + k_0$ and go to step 1.

- *Step 5*. The approximate solutions are represented by the matrix product:

$$P_m = \mathcal{V}_m X_m \mathcal{V}_m^T \quad \text{and} \quad Q_m = \mathcal{W}_m Y_m \mathcal{W}_m^T.$$

We end this section by the following result which shows that the approximate solutions P_m and Q_m are the exact solutions of two perturbed Lyapunov matrix equations.

Theorem 3.2. *Suppose that m steps of Algorithm* 2 *have been run. Let* P_m*,* Q_m *be the approximate solutions defined by* (3.1)*,* (3.3) *and* (3.2)*,* (3.4) *respectively. Then* P_m *and* Q_m *are the exact solutions of the perturbed Lyapunov matrix equations*

$$(A - \Delta_1) P_m + P_m (A - \Delta_1)^T + B B^T = 0, \tag{3.8}$$

$$(A - \Delta_2)^T Q_m + Q_m (A - \Delta_2) + C^T C = 0, \tag{3.9}$$

where

$$\Delta_1 = \tilde{V}_{m+1} W_m^T \quad \text{and} \quad \Delta_2 = V_m \tilde{W}_{m+1}^T. \tag{3.10}$$

Proof. Multiplying (3.3) on the left by $\mathcal{V}_m$, on the right by $\mathcal{V}_m^T$, we obtain

$$\mathcal{V}_m \mathcal{T}_m X_m \mathcal{V}_m^T + \mathcal{V}_m X_m T_m^T \mathcal{V}_m^T + \mathcal{V}_m E_1 \beta \beta^T E_1^T \mathcal{V}_m^T = 0.$$

Using (2.3), we have

$$\begin{aligned} &\left[A \mathcal{V}_m - V_{m+1} \beta_{m+1} E_m^T\right] X_m \mathcal{V}_m^T \\ &\qquad + \mathcal{V}_m X_m \left[\mathcal{V}_m^T A^T - E_m \beta_{m+1}^T V_{m+1}^T\right] + B B^T = 0, \end{aligned}$$

and since $\mathcal{W}_m^T \mathcal{V}_m = I_{ms}$, then

$$\begin{aligned} &\left[A - V_{m+1} \beta_{m+1} E_m^T \mathcal{W}_m^T\right] \mathcal{V}_m X_m \mathcal{V}_m^T \\ &\qquad + \mathcal{V}_m X_m \mathcal{V}_m^T \left[A^T - \mathcal{W}_m E_m \beta_{m+1}^T V_{m+1}^T\right] + B B^T = 0. \end{aligned}$$

Hence,

$$\left(A - \Delta_1\right) P_m + P_m \left(A - \Delta_1\right)^T + B B^T = 0,$$

where

$$\Delta_1 = V_{m+1} \beta_{m+1} E_m^T \mathcal{W}_m^T = V_{m+1} \beta_{m+1} W_m^T = \tilde{V}_{m+1} W_m^T.$$

We use the same arguments to show (3.9). □

4 Reduced order models via the nonsymmetric block Lanczos process

In this section, we consider the following state formulation of a multi-input and multi-output linear time invariant system (LTI)

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t), \end{cases} \tag{4.1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^s$ is the input, $y(t) \in \mathbb{R}^s$ is the output of interest and $A \in \mathbb{R}^{n \times n}$, $B, C^T \in \mathbb{R}^{n \times s}$. Applying the Laplace transform to (4.1), we obtain

$$\begin{cases} z X(z) &= A X(z) + B U(z) \\ Y(z) &= C X(z), \end{cases}$$

where $X(z)$, $Y(z)$ and $U(z)$ are the Laplace transforms of $x(t)$, $y(t)$ and $u(t)$ respectively.

The standard way of relating the input and output vectors of (4.1) is to use the associated transfer function $F(z)$ such that $Y(z) = F(z)\,U(z)$. Hence, if we eliminate $X(z)$ in the previous two equations, we get:

$$F(z) = C\left(z\,I_n - A\right)^{-1} B. \tag{4.2}$$

We recall that most of model reduction techniques, like the moment-matching approaches, are based on this transfer function [2, 13, 15, 20, 22]. Moreover, if the number of state variables of the previous LTI system is very high, (i.e., if n the order of A is large), direct computations of $F(z)$ becomes inefficient or even prohibitive. Hence, it is reasonable to look for a model of low order that approximates the behavior of the original model (4.1). This low-order model can be expressed as follows

$$\left\{ \begin{array}{rcl} \dot{x}_m(t) & = & A_m\,x_m(t) + B_m\,u(t) \\ y_m(t) & = & C_m\,x_m(t), \end{array} \right. \tag{4.3}$$

where $A_m \in \mathbb{R}^{m\times m}$, B_m and $C_m^T \in \mathbb{R}^{m\times s}$ with $m \ll n$.

In [22] and for the single-input single-output case (i.e., $s = 1$), the authors proposed a method based on the classical Lanczos process to construct an approximate reduced order model to (4.1). The aim of this section is to generalize some of the results given in [22] to the multi-input multi-output case.

More precisely, let us see how to obtain an efficient reduced model to (4.1) by using the nonsymmetric block Lanczos process. This is done by computing an approximate transfer function $F_m(z)$ to the original one $F(z)$. In fact, writing $F(z) = C\,X$ where $X = (z\,I_n - A)^{-1}\,B \in \mathbb{R}^{n\times s}$ and considering the block linear system

$$\left(z\,I_n - A\right) X \quad = \quad B, \tag{4.4}$$

we see that, approximating $F(z)$ can be achieved by computing an approximate solution X_m to X by using the block Lanczos method for solving linear systems with multiple right-hand sides [19].

Letting $\mathcal{V}_m$, $\mathcal{W}_m$ and $\mathcal{T}_m$ be the biorthonormal bases and the block tridiagonal matrix, respectively, given by the nonsymmetric block Lanczos process applied to the triplet (A, B, C^T) and starting from an initial guess $X_0 = 0$, we can show that

$$X_m = \mathcal{V}_m \left(z\, I_{ms} - \mathcal{T}_m\right)^{-1} E_1\, \beta.$$

Since

$$C = \delta\, W_1^T \quad \text{and} \quad W_1^T\, \mathcal{V}_m = \left(I_s, 0, \ldots, 0\right)^T = E_1^T \in \mathbb{R}^{s \times ms},$$

the transfer function $F(z)$ can then be approximated by

$$F_m(z) = C\, X_m = \delta\, E_1^T \left(z\, I_{ms} - \mathcal{T}_m\right)^{-1} E_1\, \beta. \tag{4.5}$$

The above result allows us to suggest the following reduced order model to (4.1)

$$\begin{cases} \dot{x}_m(t) &= \mathcal{T}_m\, x_m(t) + E_1\, \beta\, u(t) \\ y_m(t) &= \delta\, E_1^T\, x_m(t), \end{cases} \tag{4.6}$$

Next, we show that the reduced order model (4.5) proposed in the above approach approximates the behavior of the original model (4.1). Moreover, we give an upper bound for $\|F_m(z) - F(z)\|$ which enable us to monitor the progress of the iterative process at each step. More precisely, we have the following results

Theorem 4.1. *The matrices $\mathcal{T}_m$, β and δ generated by the block Lanczos process applied to the triplet (A, B, C^T) are such that the first $2m - 1$ Markov parameters of the original and the reduced models are the same, that is,*

$$C\, A^j\, B = \left(\delta\, E_1^T\right) \mathcal{T}_m^k \left(E_1\, \beta\right) \;\text{ for }\; j = 0, 1, \ldots, 2(m-1).$$

Proof. For $j \in \{0, 1, \ldots, 2m - 1\}$, let $j_1, j_2 \in \{0, 1, \ldots, m - 1\}$ such that $j = j_1 + j_2$. Using the results of Lemma 2.1 and the fact that $C = \delta\, W_1^T$, $W_1^T\, \mathcal{V}_m = E_1^T$, we have

$$\begin{aligned} C\, A^j\, B &= \delta\, W_1^T\, A^{j_1+j_2}\, V_1\, \beta = \delta\, \left[(A^T)^{j_2}\, W_1\right]^T \left[A^{j_1}\, V_1\right] \beta \\ &= \delta\, \left[\mathcal{W}_m\, (\mathcal{T}_m^T)^{j_2}\, E_1\right]^T \left[\mathcal{V}_m\, \mathcal{T}_m^{j_1}\, E_1\right] \beta \\ &= \delta\, E_1^T\, \mathcal{T}_m^{j_2} \left[\mathcal{W}_m^T\, \mathcal{V}_m\right] \mathcal{T}_m^{j_1}\, E_1\, \beta = \delta\, E_1^T \mathcal{T}_m^{j_1+j_2}\, E_1\, \beta \\ &= (\delta\, E_1^T)\, \mathcal{T}_m^j\, (E_1\, \beta). \end{aligned}$$

$\square$

Before giving an upper bound for $\|F_m(z) - F(z)\|$, we have to recall the definition of the Schur complement [35] and give the first matrix Sylvester identity [28].

Definition 4.2. *Let $\mathcal{M}$ be a matrix partitioned into four blocks*

$$\mathcal{M} = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix},$$

where the submatrix $\mathcal{D}$ is assumed to be square and nonsingular. The Schur complement of $\mathcal{D}$ in $\mathcal{M}$, denoted by $(\mathcal{M}/\mathcal{D})$, is defined by

$$(\mathcal{M}/\mathcal{D}) = \mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C}.$$

If $\mathcal{D}$ is not a square matrix then a pseudo-Schur complement of $\mathcal{D}$ in $\mathcal{M}$ can still be defined [7, 8]. Now, considering the matrices $\mathcal{K}$ and $\hat{\mathcal{M}}$ partitioned as follows

$$\mathcal{K} = \begin{pmatrix} \mathcal{A} & \mathcal{B} & \mathcal{E} \\ \mathcal{C} & \mathcal{D} & \mathcal{F} \\ \mathcal{G} & \mathcal{H} & \mathcal{L} \end{pmatrix} \quad \hat{\mathcal{M}} = \begin{pmatrix} \mathcal{D} & \mathcal{F} \\ \mathcal{H} & \mathcal{L} \end{pmatrix},$$

we have the following property

Proposition 4.3. *If the matrices $\mathcal{L}$ and $\hat{\mathcal{M}}$ are square and nonsingular, then*

$$\begin{aligned} \left(\mathcal{K}/\hat{\mathcal{M}}\right) &= \left((\mathcal{K}/\mathcal{L})/(\hat{\mathcal{M}}/\mathcal{L})\right) = \left(\begin{pmatrix} \mathcal{A} & \mathcal{E} \\ \mathcal{G} & \mathcal{L} \end{pmatrix}/\mathcal{L}\right) \\ &\quad - \left(\begin{pmatrix} \mathcal{B} & \mathcal{E} \\ \mathcal{H} & \mathcal{L} \end{pmatrix}/\mathcal{L}\right)\left(\hat{\mathcal{M}}/\mathcal{L}\right)^{-1}\left(\begin{pmatrix} \mathcal{C} & \mathcal{F} \\ \mathcal{G} & \mathcal{L} \end{pmatrix}/\mathcal{L}\right). \end{aligned}$$

Theorem 4.4. *Let α_i, β, β_i, δ, δ_i $(1 \le i \le m)$, $\tilde{V}_{m+1}$ and $\tilde{W}_{m+1}$ be the matrices obtained after m steps of the nonsymmetric block Lanczos process applied to the triplet (A, B, C^T). If $(z\, I_{ms} - T_m)$, $(z\, I_n - A)$ are nonsingular and z is such that $|z| > \|A\|_2$, then*

$$\|F(z) - F_m(z)\|_2 \le \frac{\|\delta\|_2\, \|\Gamma_{1,m}(z)\|_2\, \|\tilde{W}_{m+1}\|_2\, \|\tilde{V}_{m+1}\|_2\, \|\Gamma_{m,1}(z)\|_2\, \|\beta\|_2}{|z| - \|A\|_2}, \quad (4.7)$$

where

$$\begin{aligned} \Gamma_{m,1}(z) &= E_m^T\,(z\, I_{ms} - \mathcal{T}_m)^{-1}\, E_1 \\ &= D_m\, \beta_m\, D_{m-1}\, \beta_{m-1}\, \cdots\, D_2\, \beta_2\, D_1, \end{aligned} \quad (4.8)$$

$$\begin{aligned} \Gamma_{1,m}(z) &= E_1^T \, (z I_{ms} - \mathcal{T}_m)^{-1} \, E_m \\ &= D_1 \, \delta_2 \, D_2 \, \cdots \, \delta_{m-1} \, D_{m-1} \, \delta_m \, D_m, \end{aligned} \tag{4.9}$$

and

$$D_m = (z \, I_s - \alpha_m)^{-1}, \, D_{j-1} = (z \, I_s - \alpha_{j-1} - \delta_j \, D_j \, \beta_j)^{-1} \ \text{ for } \ j = m, \ldots, 2.$$

Proof. As $C = \delta \, W_1^T$ and $B = V_1 \, \beta$, we have

$$\begin{aligned} F(z) - F_m(z) &= C \, (z \, I_n - A)^{-1} \, B - \delta \, E_1^T (z \, I_{ms} - \mathcal{T}_m)^{-1} \, E_1 \, \beta \\ &= \delta \, W_1^T \, (z \, I_n - A)^{-1} \, V_1 \, \beta - \delta \, E_1^T (z \, I_{ms} - \mathcal{T}_m)^{-1} \, E_1 \, \beta \\ &= \delta \, G(z) \, \beta, \end{aligned}$$

where $G(z) = W_1^T \, (z \, I_n - A)^{-1} \, V_1 - E_1^T (z \, I_{ms} - \mathcal{T}_m)^{-1} \, E_1$. Now, from (2.3), we have

$$\begin{aligned} \mathcal{V}_m \, (z \, I_{ms} - \mathcal{T}_m)^{-1} &= (z \, I_n - A)^{-1} (z \, I_n - A) \, \mathcal{V}_m \, (z \, I_{ms} - T_m)^{-1} \\ &= (z \, I_n - A)^{-1} \, (z \, \mathcal{V}_m - A \, \mathcal{V}_m) \, (z \, I_{ms} - \mathcal{T}_m)^{-1} \\ &= (z \, I_n - A)^{-1} \, (z \, \mathcal{V}_m - \mathcal{V}_m \, \mathcal{T}_m - \tilde{V}_{m+1} \, E_m^T) \, (z \, I_{ms} - \mathcal{T}_m)^{-1} \\ &= (z \, I_n - A)^{-1} \, \mathcal{V}_m - (z \, I_n - A)^{-1} \, \tilde{V}_{m+1} \, E_m^T \, (z \, I_{ms} - \mathcal{T}_m)^{-1}. \end{aligned}$$

Multiplying the last relation on the left by W_1^T, on the right by E_1 we obtain

$$G(z) = W_1^T \, (z \, I_n - A)^{-1} \, \tilde{V}_{m+1} \, E_m^T \, (z \, I_{ms} - \mathcal{T}_m)^{-1} \, E_1. \tag{4.10}$$

Similarly, by using (2.4), we have

$$\begin{aligned} (z \, I_{ms} - \mathcal{T}_m)^{-1} \, \mathcal{W}_m^T &= (z \, I_{ms} - \mathcal{T}_m)^{-1} \, \mathcal{W}_m^T \, (z \, I_n - A) \, (z \, I_n - A)^{-1} \\ &= (z \, I_{ms} - \mathcal{T}_m)^{-1} \, (z \, \mathcal{W}_m^T - \mathcal{W}_m^T \, A)(z \, I_n - A)^{-1} \\ &= (z \, I_{ms} - \mathcal{T}_m)^{-1} (z \, \mathcal{W}_m^T - \mathcal{T}_m \, \mathcal{W}_m^T - E_m \, \tilde{W}_{m+1}^T) \, (z \, I_n - A)^{-1} \\ &= (\mathcal{W}_m^T \, (z \, I_n - A)^{-1} - (z \, I_{ms} - \mathcal{T}_m)^{-1} \, E_m \, \tilde{W}_{m+1}^T \, (z \, I_n - A)^{-1} \end{aligned}$$

and again, multiplying the last equality on the left by E_1^T, on the right by $\tilde{V}_{m+1}$, we have

$$W_1^T \, (z \, I_n - A)^{-1} \, \tilde{V}_{m+1} - E_1^T \, (z \, I_{ms} - \mathcal{T}_m)^{-1} \, E_m \, \tilde{W}_{m+1}^T \, (z \, I_n - A)^{-1} \, \tilde{V}_{m+1} = 0. \tag{4.11}$$

Combining the formulas (4.10), (4.11) and letting

$$\Gamma_{m,1}(z) = E_m^T\,(z\,I_{ms} - \mathcal{T}_m)^{-1}\,E_1,\ \Gamma_{1,m}(z) = E_1^T\,(z\,I_{ms} - \mathcal{T}_m)^{-1}\,E_m,$$

we get

$$G(z) = \Gamma_{1,m}(z)\,\tilde{W}_{m+1}^T\,(z\,I_n - A)^{-1}\,\tilde{V}_{m+1}\,\Gamma_{m,1}(z),$$

and then

$$F(z) - F_m(z) = \delta\left[\Gamma_{1,m}(z)\,\tilde{W}_{m+1}^T\,(z\,I_n - A)^{-1}\,\tilde{V}_{m+1}\,\Gamma_{m,1}(z)\right]\beta.$$

Finally, using the inequality $\|(z\,I_n - A)^{-1}\|_2 \le \frac{1}{|z|-\|A\|_2}$ for $|z| > \|A\|_2$, we have

$$\begin{aligned}
&\|F(z) - F_m(z)\|_2 \\
&\quad\le \|\delta\|_2\,\|\Gamma_{1,m}(z)\|_2\,\|\tilde{W}_{m+1}\|_2\,\|(z\,I_n - A)^{-1}\|_2\,\|\tilde{V}_{m+1}\|_2\,\|\Gamma_{m,1}(z)\|_2\,\|\beta\|_2 \\
&\quad\le \frac{\|\delta\|_2\,\|\Gamma_{1,m}(z)\|_2\ \ |\tilde{W}_{m+1}\|_2\,\|\tilde{V}_{m+1}\|_2\ \ |\Gamma_{m,1}(z)\|_2\,\|\beta\|_2}{|z| - \|A\|_2}.
\end{aligned}$$

Next, set $D_m = (z\,I_s - \alpha_m)^{-1}$ and remark that $\Gamma_{m,1}(z)$ is the Schur complement of

$$(z\,I_{ms} - \mathcal{T}_m) \quad \text{in} \quad \begin{pmatrix} 0 & -E_m^T \\ E_1 & z\,I_{ms} - \mathcal{T}_m \end{pmatrix}.$$

Hence, using the result of Proposition 4.3, we have

$$\begin{aligned}
\Gamma_{m,1}(z) &= \left(\begin{pmatrix} 0 & -E_m^T \\ E_1 & z\,I_{ms} - \mathcal{T}_m \end{pmatrix} / (z\,I_{ms} - \mathcal{T}_m)\right) \\
&= \left(\begin{pmatrix} 0 & 0 & -I_s \\ E_1 & z\,I_{(m-1)s} - \mathcal{T}_{m-1} & -\delta_m\,E_{m-1} \\ 0 & -\beta_m\,E_{m-1}^T & z\,I_s - \alpha_m \end{pmatrix} / (z\,I_{ms} - \mathcal{T}_m)\right) \\
&= \left(\begin{pmatrix} 0 & -I_s \\ 0 & D_m^{-1} \end{pmatrix} / D_m^{-1}\right) - \left(\begin{pmatrix} 0 & -I_s \\ -\beta_m\,E_{m-1}^T & D_m^{-1} \end{pmatrix} / D_m^{-1}\right) \\
&\quad\times \left(\begin{pmatrix} z\,I_{(m-1)s} - \mathcal{T}_{m-1} & -\delta_m\,E_{m-1} \\ -\beta_m\,E_{m-1}^T & D_m^{-1} \end{pmatrix} / D_m^{-1}\right)^{-1} \\
&\quad\times \left(\begin{pmatrix} E_1 & -\delta_m\,E_{m-1} \\ 0 & D_m^{-1} \end{pmatrix} / D_m^{-1}\right) \\
&= D_m\,\beta_m\,E_{m-1}^T\left[z\,I_{(m-1)s} - \mathcal{T}_{m-1} - \delta_m\,E_{m-1}\,D_m\,\beta_m\,E_{m-1}^T\right]^{-1}\,E_1 \\
&= D_m\,\beta_m\,E_{m-1}^T\,(z\,I_{(m-1)s} - \hat{\mathcal{T}}_{m-1})^{-1}\,E_1,
\end{aligned}$$

where

$$E_1^T = \left(I_s, 0_s, \cdots, 0_s\right) \in \mathbb{R}^{s\times(m-1)s},\ E_{m-1}^T = \left(0_s, \cdots, 0_s, I_s\right) \in \mathbb{R}^{s\times(m-1)s}$$

and $\hat{\mathcal{T}}_{m-1}$ is a block tridiagonal matrix having the same elements that $\mathcal{T}_{m-1}$ except the $(m-1, m-1)$ block which is equal to $(\alpha_{m-1} + \delta_m\, D_m\, \beta_m)$.

Again, applying Proposition 4.3 to compute $E_{m-1}^T\,(z\, I_{(m-1)s} - \hat{\mathcal{T}}_{m-1})^{-1}\, E_1$, and so on, we finally obtain (4.8). Similarly, we remark that $\Gamma_{1,m}(z)^T$ is the Schur complement of

$$\left(z\, I_{ms} - \mathcal{T}_m^T\right) \quad \text{in} \quad \begin{pmatrix} 0 & -E_m^T \\ E_1 & z\, I_{ms} - \mathcal{T}_m^T \end{pmatrix}.$$

Then, using the same arguments as for $\Gamma_{m,1}(z)$ we get (4.9). □

Summarizing the previous results, we get the following algorithm

Algorithm 4. Model reduction via the block Lanczos process

- *Inputs :* A the system matrix, B the input matrix, C the output matrix.
- *Step 0 .* Choose a tolerance $\epsilon > 0$, an integer parameter k_0 and set $k = 0$; $m = k_0$.
- *Step 1 .* For $j = k+1, k+2, \ldots, m$
  construct the block tridiagonal matrix $\mathcal{T}_m$, $\tilde{V}_{m+1}$ and $\tilde{W}_{m+1}$ by Algorithm 1 applied to the triplet (A, B, C^T).
  compute the matrices $\Gamma_{1,m}(z)$ and $\Gamma_{m,1}(z)$ using (4.8), (4.9).
  end For
- *Step 2 .* Compute the upper bound for the residual norm:

$$r_m = \frac{\|\delta\|_2\, \|\Gamma_{1,m}(z)\|_2\, \|\tilde{W}_{m+1}\|_2\, \|\tilde{V}_{m+1}\|_2\, \|\Gamma_{m,1}(z)\|_2\, \|\beta\|_2}{|z| - \|A\|_2}.$$

- *Step 3 .* If $r_m > \epsilon$, set $k = k + k_0$; $m = k + k_0$ and go to step 1.
- *Step 4 .* The reduced order model is $A_m = \mathcal{T}_m$, $B_m = E_1\, \beta$ and $C_m = \delta\, E_1^T$.

5 Numerical experiments

In this section, we present some numerical experiments to illustrate the behavior of the block Lanczos process when applied to solve coupled Lyapunov equations. We also applied the block Lanczos process for model reduction in large scale dynamical systems. All the experiments were performed on a computer of Intel Pentium-4 processor at 3.4GHz and 1024 MBytes of RAM. The experiments were done using Matlab 6.5.

Experiment 1. In this first experiment, we compared the performance of the coupled Lyapunov block Lanczos (CLBL) and the coupled Lyapunov block Arnoldi (CLBA) algorithms [21]. Notice that:

- In all the experiments, the parameter k_0 used to compute the solutions of the low-order Lyapunov equations is $k_0 = 5$.
- For the coupled Lyapunov block Arnoldi algorithm, the tests were stopped when the residual given in [21] was less than $\epsilon = 10^{-6}$.
- For the coupled Lyapunov block Lanczos algorithm, the iterations were stopped when

$$\max\left\{\| \tilde{V}_{m+1}\, \tilde{X}_m \|_F,\, \| \tilde{W}_{m+1}\, \tilde{Y}_m \|_F\right\} \leq \epsilon = 10^{-6}. \tag{5.1}$$

We note that $Res_1 = \|\ A\, P_m + P_m\, A^T + B\, B^T\ \|_F$ and $Res_2 = \|\ A^T\, Q_m + Q_m\, A + C^T\, C\ \|_F$ are the exact Frobenius residual norms for the first and the second Lyapunov equations (1.3) respectively. The number $Iter$ of iterations required for CLBA corresponds to the total number of iterations needed for solving separately the two Lyapunov equations (1.3) by the Lyapunov block Arnoldi method.

The matrices A_1 and A_2 tested in this experiment comes from the five-point discretization of the operators [30]

$$\begin{aligned} L_1(u) &= \Delta u - (x-y)\frac{\partial u}{\partial x} - \sin(x+y)\frac{\partial u}{\partial y} - 10^3\, e^{x\,y}\, u, \\ L_2(u) &= \Delta u - \frac{1}{2}\sqrt{x+y}\,\frac{\partial u}{\partial x} - (\cos(x)+\cos(y))\frac{\partial u}{\partial y} - (x+y)\, u, \end{aligned}$$

on the unit square [0, 1] × [0, 1] with homogeneous Dirichlet boundary conditions. The dimension of each matrix is $n = n_0^2$ where n_0 is the number of inner grid points in each direction. The obtained stiffness matrices A_1 and A_2 are sparse and nonsymmetric with a band structure [30]. The entries of the matrices B and C were random values uniformly distributed on [0, 1].

(A_i, n_0, s)	Algorithm	$Iter$	Res_1	Res_2	flops
$(A_1, 60, 3)$	CLBA	45	$6.96\,10^{-8}$	$8.87\,10^{-8}$	$2.89\,10^{8}$
$nnz(A_1) = 17760$	CLBL	55	$7.49\,10^{-7}$	$4.80\,10^{-8}$	$9.57\,10^{7}$
$(A_1, 50, 4)$	CLBA	35	$3.73\,10^{-7}$	$3.94\,10^{-7}$	$2.29\,10^{8}$
$nnz(A_1) = 12300$	CLBL	45	$7.49\,10^{-7}$	$4.80\,10^{-8}$	$9.51\,10^{7}$
$(A_1, 50, 3)$	CLBA	35	$3.08\,10^{-7}$	$2.86\,10^{-7}$	$2.29\,10^{8}$
$nnz(A_1) = 12300$	CLBL	50	$1.05\,10^{-8}$	$1.72\,10^{-9}$	$6.11\,10^{7}$
$(A_2, 60, 3)$	CLBA	60	$1.34\,10^{-7}$	$1.21\,10^{-7}$	$5.25\,10^{8}$
$nnz(A_2) = 17760$	CLBL	75	$1.14\,10^{-6}$	$2.65\,10^{-7}$	$1.97\,10^{8}$
$(A_2, 60, 2)$	CLBA	50	$1.96\,10^{-7}$	$2.72\,10^{-7}$	$2.28\,10^{8}$
$nnz(A_2) = 17760$	CLBL	55	$1.51\,10^{-8}$	$2.88\,10^{-8}$	$7.47\,10^{7}$
$(A_2, 50, 4)$	CLBA	50	$5.84\,10^{-8}$	$9.74\,10^{-8}$	$4.65\,10^{8}$
$nnz(A_2) = 12300$	CLBL	55	$4.03\,10^{-8}$	$3.01\,10^{-9}$	$1.51\,10^{8}$

Table 5.1 – Effectiveness of the coupled Lyapunov block Lanczos and coupled Lyapunov block Arnoldi methods.

Experiment 2. The dynamical system used in this experiment is a non trivial constructed model (**FOM**) from [9, 31]. Originally, the system obtained from the FOM model is SISO and is of order $n = 1006$. So, in order to get a MIMO system, we modified the inputs and outputs. The state matrices are given by

$$A = \begin{pmatrix} \tilde{A}_1 & & & \\ & \tilde{A}_2 & & \\ & & \tilde{A}_3 & \\ & & & \tilde{A}_4 \end{pmatrix}, \quad B = \big[b_1, b_2, \ldots, b_6\big], \quad C^T = \big[c_1, c_2, \ldots, c_6\big],$$

where

$$\tilde{A}_1 = \begin{pmatrix} -1 & 100 \\ -100 & -1 \end{pmatrix}, \quad \tilde{A}_2 = \begin{pmatrix} -1 & 200 \\ -200 & -1 \end{pmatrix}, \quad \tilde{A}_3 = \begin{pmatrix} -1 & 400 \\ -400 & -1 \end{pmatrix}$$

and $\tilde{A}_4 = \mathrm{diag}(-1, \ldots, -1000)$. The columns of B and C are such that

$$b_1^T = c_1 = (\underbrace{10, \ldots, 10}_{6}, \underbrace{1, \ldots, 1}_{1000}), \quad \text{and} \quad b_2, \ldots, b_6; c_2, \ldots, c_6$$

are random column vectors.

The response plot and error plot given below show the singular values $\sigma_{\max}(F(\jmath\,\omega))$ and $\sigma_{\max}(F_m(\jmath\,\omega) - F(\jmath\,\omega))$ as a function of the frequency ω respectively, where $\sigma_{\max}(.)$ denotes the largest singular value and $\omega \in [10^{-1}\ 10^5]$. As a stopping criterion, we used the upper bound (4.7). More precisely, we stopped the computation when

$$\max_{w} \left(\frac{\|\delta\|_2 \, \|\Gamma_{1,m}(z)\|_2 \, \|\tilde{W}_{m+1}\|_2 \, \|\tilde{V}_{m+1}\|_2 \, \|\Gamma_{m,1}(z)\|_2 \, \|\beta\|_2}{|z| - \|A\|_2} \right) \leq \eta = 10^{-7}.$$

The frequency response (solid line) of the modified FOM model is given in the top of Figure 5.1 and is compared to the frequencies responses of order $m = 12$ when using the block Arnoldi process (dash-dotted line) and the block Lanczos process (dashed line). The exact errors $\|F(z) - F_{12}(z)\|_2$ produced by the two processes are shown in the bottom of Figure 5.1.

6 Conclusion

In this paper, we applied the block Lanczos process for solving coupled Lyapunov matrix equations and also for model reduction. We gave some new theoretical results and showed the effectiveness of this process with some numerical examples.

Acknowledgments. We would like to thank the referees for their recommendations and helpful suggestions.

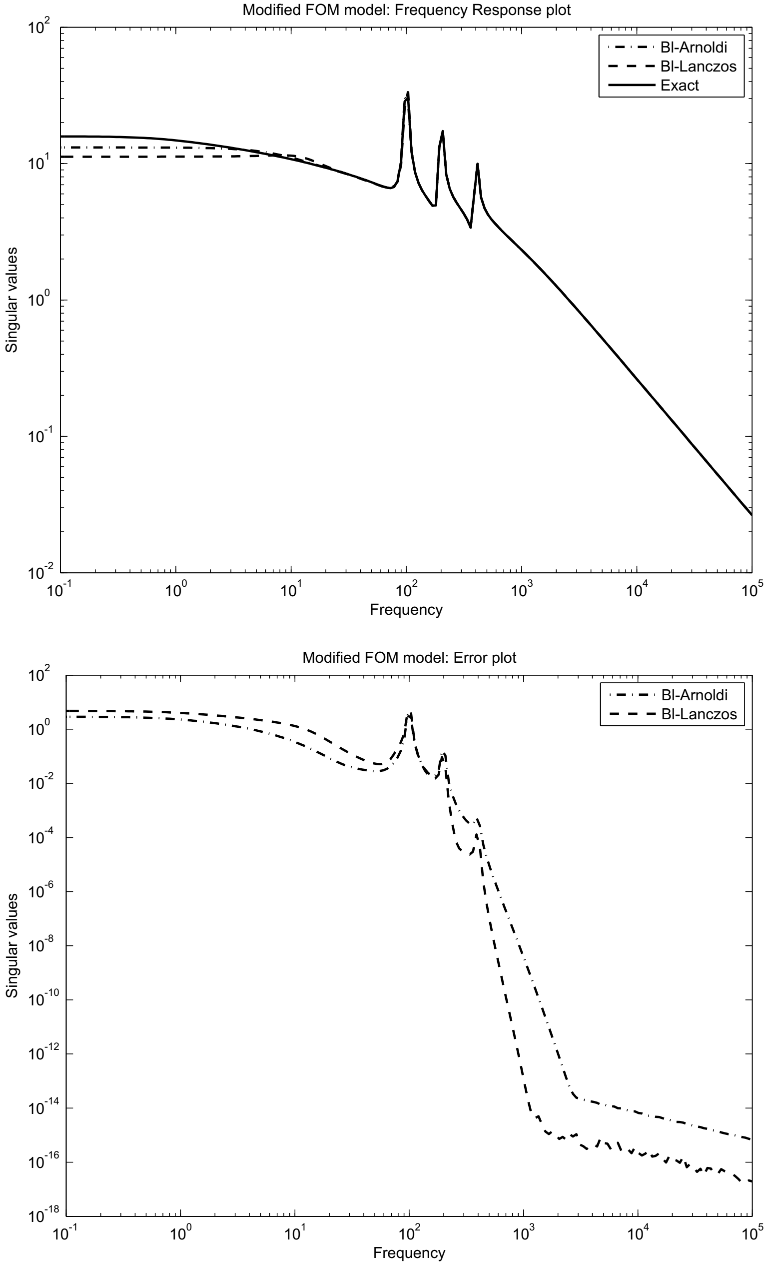

Figure 5.1 – Top: $\|F(\jmath\omega)\|_2$ (solid line) and it's approximations $\|F_{12}^{BA}(\jmath\omega)\|_2$ and $\|F_{12}^{BL}(\jmath\omega)\|_2$. Bottom: Exact errors $\|F(\jmath\omega) - F_{12}^{BA}(\jmath\omega)\|_2$ (dashed-dotted line) and $\|F(\jmath\omega) - F_{12}^{BL}(\jmath\omega)\|_2$ (dashed line).

REFERENCES

[1] A.C. Antoulas and D.C. Sorensen, *Projection methods for balanced model reduction*, Technical Report, Rice University, Houston, Tx, (2001).

[2] Z. Bai and Q. Ye, *Error estimation of the Padé approximation of transfer functions via the Lanczos process*. Elect. Trans. Numer. Anal., **7** (1998), 1–17.

[3] Z. Bai, *Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems*. App. Num. Math., **43** (2002), 9–44.

[4] Z. Bai, D. Day and Q. Ye, *ABLE: An adaptive block Lanczos method for non-Hermitian eigenvalue problems*. SIAM J. Mat. Anal. Appl., **20**(4) (1999), 1060–1082.

[5] R.H. Bartels and G.W. Stewart, *Solution of the matrix equation* $AX + XB = C$. Comm. ACM, **15** (1972), 820–826.

[6] D.L. Boley and B.N. Datta, *Numerical methods for linear control systems*, in: C. Byrnes, B. Datta, D. Gilliam, C. Martin (Eds.) Systems and Control in the twenty-First Century, Birkhauser, pp. 51–74, (1996).

[7] C. Brezinski and M.R. Zaglia, *A Schur complement approach to a general extrapolation algorithm*. Linear Algebra Appl., **368** (2003), 279–301.

[8] D. Carlson, *What are Schur complements, anyway?* Linear Algebra Appl., **74** (1986), 257–275.

[9] Y. Chahlaoui and P. Van Dooren, *A collection on Benchmark examples for model reduction of linear time invariant dynamical systems*. SILICOT Working Note 2002-2. http://www.win.tue.nl/niconet/NIC2/benchmodred.html.

[10] B.N. Datta, *Linear and numerical linear algebra in control theory: Some research problems*. Lin. Alg. Appl., **197-198** (1994), 755–790.

[11] B.N. Datta, *Large-Scale Matrix computations in Control*. Applied Numerical Mathematics, **30** (1999), 53–63.

[12] B.N. Datta, *Krylov Subspace Methods for Large-Scale Matrix Problems in Control*. Future Gener. Comput. Syst., **19**(7) (2003), 1253–1263.

[13] B.N. Datta, *Numerical Methods for Linear Control Systems Design and Analysis*. Elsevier Academic Press, (2003).

[14] A. El Guennouni, K. Jbilou and A.J. Riquet, *Block Krylov subspace methods for solving large Sylvester equations*. Numer. Alg., **29** (2002), 75–96.

[15] P. Feldmann and R.W. Freund, *Efficient Linear Circuit Analysis by Padé Approximation via The Lanczos process*. IEEE Trans. on CAD of Integrated Circuits and Systems, **14** (1995), 639–649.

[16] K. Glover, *All optimal Hankel-norm approximations of linear multivariable systems and their L-infinity error bounds*. International Journal of Control, **39** (1984), 1115–1193.

[17] K. Glover, D.J.N. Limebeer, J.C. Doyle, E.M. Kasenally and M.G. Safonov, *A characterisation of all solutions to the four block general distance problem*. SIAM J. Control Optim., **29** (1991), 283–324.

[18] G.H. Golub, S. Nash and C. Van Loan, *A Hessenberg-Schur method for the problem* $AX + XB = C$. IEEE Trans. Autom. Control, **AC-24** (1979), 909–913.

[19] G.H. Golub and R. Underwood, *The block Lanczos method for computing eigenvalues*. Mathematical Software III, J.R. Rice, ed., Academic Press, New York, pp. 361–377, (1977).

[20] E.J. Grimme, D.C. Sorensen and P. Van Dooren, *Model reduction of state space systems via an implicitly restarted Lanczos method*. Numer. Alg., **12** (1996), 1–32.

[21] I.M. Jaimoukha and E.M. Kasenally, *Krylov subspace methods for solving large Lyapunov equations*. SIAM J. Matrix Anal. Appl., **31**(1) (1994), 227–251.

[22] I.M. Jaimoukha and E.M. Kasenally, *Oblique projection methods for large scale model reduction*. SIAM J. Matrix Anal. Appl., **16**(2) (1995), 602–627.

[23] K. Jbilou and A.J. Riquet, *Projection methods for large Lyapunov matrix equations*. Linear Alg. Appl., **415**(2-2) (2006), 344–358.

[24] S.J. Hammarling, *Numerical solution of the stable, nonnegative definite Lyapunov equation*. IMA J. Numer. Anal., **2** (1982), 303–323.

[25] A.S. Hodel, *Recent applications of the Lyapunov equation in control theory*, in Iterative Methods in Linear Algebra, R. Beauwens and P. de Groen, eds., Elsevier (North Holland), pp. 217–227, (1992).

[26] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, (1991).

[27] J. Lasalle and S. Lefschetz, *Stability of Lyapunov direct methods*. Academic Press, New York, (1961).

[28] A. Messaoudi, *Recurssive interpolation algorithm: a formalism for solving systems of linear equations, I. direct methods*. Comput. Appl., **76** (1996), 31–53.

[29] B.C. Moore, *Principal component analysis in linear systems: controllability, observability and model reduction*. IEEE Trans. Automatic Contr., **AC-26** (1981), 17–32.

[30] T. Penzl, *LYAPACK A MATLAB toolbox for Large Lyapunov and Riccati Equations, Model Reduction Problem, and Linear-quadratic Optimal Control Problems*. http://www.tu-chemintz.de/sfb393/lyapack.

[31] T. Penzl, *Algorithms for model reduction of large dynamical systems*. Technical Report, SFB393/99-40. TU Chemintz, 1999. http://www.tu-chemintz.de/sfb393/sfb99pr.html.

[32] Y. Saad, *Numerical solution of large Lyapunov equations*, in Signal Processing, Scattering, Operator Theory and Numerical Methods, M.A. Kaashoek, J.H. Van Shuppen and A.C. Ran, eds., Birkhauser, Boston, pp. 503–511, (1990).

[33] Y. Shamash, *Stable reduced-order models using Padé type approximations*. IEEE. Trans. Automatic Control, **AC-19** (1974), 615–616.

[34] Y. Shamash, *Model reduction using the Routh stability criterion and the Padé approximation technique*. Internat. J. Control, **21** (1975), 475–484.

[35] I. Schur, *Potenzreihn im Innern des Einheitskreises*. J. reine Angew. Math, **147** (1917), 205–232.

[36] Q. Ye, *An Adaptive Block Lanczos Algorithm*. Num. Alg., **12** (1996), 97–110.

[37] K. Zhou, J.C. Doyle and K. Glover, *Robust and Optimal Control*. New Jersey: Prentice Hall, (1996).