Appl. Comput. Math., V.xx, N.xx, 20xx, pp.xx-xx

A SURVEY OF KRYLOV-BASED METHODS FOR MODEL REDUCTION IN LARGE-SCALE MIMO DYNAMICAL SYSTEMS

K. JBILOU[1]

Abstract. In this work, we review different numerical methods for low order model reduction and linear control problems. These methods are essentially based on block, global, extended block or rational block Krylov subspace methods when we are dealing with large-scale dynamical systems. In this paper, the main schemes for model reduction are outlined, and personal contribution to some theoretical aspects is also given. Tools and approaches in model simplification, developed by some research groups are discussed, and the main results obtained during the last few years are summarized.

Keywords: Dynamical systems, Krylov subspaces, Large-scale, Model reduction.

AMS Subject Classification: 65F10, 65F30

1. Introduction

A dynamical system is an interconnected set of mechanical, electrical or other devices which has to provide a desired function. This system has input and output variables. The output variables can be measured while the input variables can influence the outputs of the system and could be controlled to give more interesting properties to the dynamical system. Controlling a system according to measurements of the output variables is called feedback. It needs the knowledge of the state variables, assumed to be known. They can be estimated by a special system called the observer. Generally, a dynamical system comes from the discretization (in the space) of partial differential equations (PDE's). Attention is often devoted to the classical approximation of complex dynamic systems, and the first type of approximation is devoted to obtaining linearized time-invariant models from nonlinear, distributed, or time-variant systems. Indeed, even a linear time-invariant system derived from this type of approximation is often too complicated to be investigated due to the large number of state variables that are included.
Linear systems have already been studied for a long time and from many different points of view: in physics, mathematics, engineering, and so on. In engineering contexts, linear systems have been extensively studied since the 1930s [9, 85, 86, 87, 94]. Moreover, a model is often too complicated to be used in real problems, so approximation procedures based on physical considerations or using mathematical approaches must be used to achieve simpler models than the original ones. One model reduction scheme that is well grounded in theory and most commonly used is the so-called Balanced Model Reduction first introduced by Mullis and Roberts [87] and later in the systems and control literature by Moore [86]. To apply balanced reduction, first the system is transformed to a basis where the states which are difficult to reach are simultaneously difficult to observe. This is achieved by simultaneously diagonalizing the reachability and the observability Gramians, which are solutions to the reachability and the observability Lyapunov equations. Then, the reduced model is obtained by truncating the states which have this property. This is called the Lyapunov balancing method. When applied to stable systems, Lyapunov balanced reduction preserves stability and provides a bound on the approximation

[1] L.M.P.A. University Littoral Cote d'Opale, 50 rue F. Buisson, BP 699, F-62228 Calais Cedex France.
e-mail: jbilou@lmpa.univ-littoral.fr
Manuscript received xx.

error. For small-to-medium scale problems, Lyapunov balancing can be implemented efficiently. However, for large-scale settings, exact balancing is expensive to implement because it requires dense matrix factorizations and results in a computational complexity of $O(n^3)$ and a storage requirement of $O(n^2)$, see [9, 23, 54]. For large problems, direct methods could not be applied and then Krylov-based [41, 62, 63, 64, 66] or ADI-based methods [19, 21, 91] are required to compute these Gramians that are given in factored forms which allows to save memory. Besides the Lyapunov balancing method, other types of balancing exist such as stochastic balancing, bounded real balancing, positive real balancing, LQG balancing and frequency weighted balancing requiring the solution of continuous time algebraic Riccati equations see [9, 56]. A closely related balancing method is positive real balancing which is applied for model reduction of positive real (passive) systems, an important subclass of dynamical systems. All the balancing techniques mentioned above try to approximate the full-order model over all frequencies. However, in many applications one is only interested in a given frequency interval. In these cases the frequency weighted balanced reduction is used which tries to reduce the error between and over the specified frequency range, i.e. the weighted error. Several ways of weighted balancing have been introduced in the literature [9, 46].
Another important class of model reduction methods is called the moment (or Markov) matching methods. These methods use directly the computed orthonormal bases of some Krylov-type subspaces and then use these bases to construct directly the reduced model order. Among these methods are the Arnoldi and the Lanczos methods [9, 62, 63]. The standard versions of these algorithms tend to create reduced order models that poorly approximate some frequency dynamics. To ameliorate this problem, rational Arnoldi and Lanczos algorithms [1, 34, 16, 17, 36, 37, 51, 42] have been developed which produce reduced models that match the moments of the transfer function at different frequencies. The Krylov-subspace based methods have the advantages for large problems. This is due to the fact that they need only matrix-vector operations and no decomposition of large matrices is required.. They require $\mathcal{O}(nm^2)$ operations for sparse problems and the requirement of memory is about $\mathcal{O}(mn)$ where n is the size of the original problem and m is the size of the reduced problem. The main drawbacks of these methods is that although the original system is stable, the reduced order model may not be stable. To remedy this drawback one can apply implicit restart techniques [52, 89]. It is also known that the Lanczos method is related to the Padé approximation [30] which is also a useful method for generating reduced order models [29, 31, 39]. In the recent years a great attention was given to the Padé approximation in simulation of large electronic VLSI circuits. Due to the instability of generating Padé approximants directly from Taylor series it is preferable to use the Lanczos-type methods. The main disadvantage is the fact that reduced order modeling techniques based on Padé approximation do not preserve the stability of the original system. In [14] a partial-Padé type approximant was used to preserve stability and passivity of the original system.

Consider the following linear time invariant (LTI) dynamical system

$$\mathbf{\Sigma}\begin{cases} x'(t) = Ax(t) + Bu(t);\ x(t_0) = x_0 \\ \\ y(t) = Cx(t) \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^p$, $y(t) \in \mathbb{R}^s$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$ and $C \in \mathbb{R}^{s\times n}$ with $p, s \ll n$. The vector x is called the state vector and it belongs to the state space. The vector u is the input (or the control) vector and $y(t)$ is the output (to be measured). If $s = p = 1$, then the LTI dynamical system (1) is called Single-Input Single-Output (SISO) and is called Multiple-Input Multiple-Output (MIMO) otherwise. The control problem consists of acting on the input vector $u(t)$ so that the output vector $y(t)$ has a desirable time trajectory and modifying the input $u(t)$ according to the output $y(t)$ which is observed or to the state $x(t)$ is called feedback. The LTI

dynamical system (1) can also be denoted as

$$\mathbf{\Sigma} \equiv \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]. \tag{2}$$

In many applications, such as circuit simulation, or time dependent PDE control problems, the dimension n of $\mathbf{\Sigma}$ is quite large, while the number of inputs and outputs is small $p, s \ll n$. In these large-scale settings, the system dimension makes the computation infeasible due to memory, time limitations and ill-conditioning. The goal is to produce a low dimensional system that has similar response characteristics as the original system with lower storage requirements and evaluation times.
The reduced order dynamical system can be stated as follows

$$\mathbf{\Sigma}_m \left\{ \begin{array}{l} x'_m(t) = A_m x_m(t) + B_m u(t) \\ \\ y_m(t) = C_m x_m(t) \end{array} \right. \tag{3}$$

where $x_m(t) \in \mathbb{R}^m$, $y_m(t) \in \mathbb{R}^s$, $A_m \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{m\times p}$ and $C_m \in \mathbb{R}^{s\times m}$ with $m \ll n$. The system (3) is also represented as

$$\mathbf{\Sigma}_m \equiv \left[\begin{array}{c|c} A_m & B_m \\ \hline C_m & 0 \end{array}\right].$$

The reduced order dynamical system (3) should be constructed such that

- The output $y_m(t)$ of the reduced system approaches the output $y(t)$ of the original system.
- Some properties of the original system such as passivity and stability (if possible) are preserved.
- The computation methods are steady and efficient.

Some Notations. For two matrices Y and Z in $\mathbb{R}^{n\times p}$, we define the inner product $\langle Y, Z\rangle_F = \mathrm{tr}(Y^T Z)$ where $\mathrm{tr}(Y^T Z)$ denotes the trace of the matrix $Y^T Z$. The associated norm is the Frobenius norm denoted by $\|.\|_F$. A system of vectors (matrices) of $\mathbb{R}^{n\times p}$ is said to be F-orthonormal if it is orthonormal with respect to $\langle ., .\rangle_F$. For $Y = [y_{i,j}] \in \mathbb{R}^{n\times p}$, we denote by $\mathrm{vec}(Y)$ the vector of $\mathbb{R}^{np}$ defined by $\mathrm{vec}(Y) = [y(.,1)^T,\ y(.,2)^T,\ \ldots\ , y(.,p)^T]^T$ where $y(.,j)$, $j = 1, \ldots, p$, is the j-th column of Y. The matrix $A \otimes B = [a_{i,j}B]$ is the Kronecker product of the matrices A and B. For this product, we have the following properties [76]

(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 nonsingular matrices of dimension $n \times n$ and $p \times p$ respectively, then $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
(4) If A and B are $n \times n$ and $p \times p$ matrices, then $det(A \otimes B) = det(A)^p det(B)^n$ and $tr(A \otimes B) = tr(A)tr(B)$.
(5) $\mathrm{vec}(AXB) = (B^T \otimes A)\,\mathrm{vec}(X)$,
(6) $\mathrm{vec}(A)^T\,\mathrm{vec}(B) = \mathrm{trace}(A^T B)$.

We will also use the matrix product denoted by $\diamond$ that we already introduced in [28] and defined as follows:

Definition 1.0.1. Let $A = [A_1, A_2, \ldots, A_p]$ and $B = [B_1, B_2, \ldots, B_l]$ be matrices of dimension $n \times ps$ and $n \times ls$ respectively where A_i and B_j $(i = 1, \ldots, p;\ j = 1, \ldots, l)$ are $n \times s$ matrices. Then the $p \times l$ matrix $A^T \diamond B$ is defined by:

$$A^T \diamond B = \begin{pmatrix} \langle A_1, B_1\rangle_F & \langle A_1, B_2\rangle_F & \ldots & \langle A_1, B_l\rangle_F \\ \langle A_2, B_1 >_F & \langle A_2, B_2\rangle_F & \ldots & \langle A_2, B_l\rangle_F \\ \vdots & \vdots & \vdots & \vdots \\ \langle A_p, B_1\rangle_F & \langle A_p, B_2\rangle_F & \ldots & \langle A_p, B_l\rangle_F \end{pmatrix}.$$

It is not difficult to show the following properties satisfied by the product $\diamond$.

Proposition 1.0.2. *Let $A, B, C \in \mathbb{R}^{n\times ps}$, $D \in \mathbb{R}^{n\times n}$, and $L \in \mathbb{R}^{p\times p}$. Then we have*

(1) $(A+B)^T \diamond C = A^T \diamond C + B^T \diamond C$.
(2) $A^T \diamond (B+C) = A^T \diamond B + A^T \diamond C$.
(3) $(A^T \diamond B)^T = B^T \diamond A$.
(4) $(DA)^T \diamond B = A^T \diamond (D^T B)$.
(5) $A^T \diamond (B(L \otimes I_s)) = (A^T \diamond B)L$.
(6) $||A^T \diamond B||_F \leq ||A||_F \, ||B||_F$.

2. Stability, controllability and observability of a dynamical system

2.1. **Stability-Passivity.**

We first give the definition of stable dynamical systems.

Definition 2.1.1. The LTI dynamical system (1) is

- Asymptotically stable, if and only if A is stable $()\Lambda(A) \subset \mathbb{C}^-$).
- Stable, if and only if all eigenvalues of A have nonpositive real parts, and in addition, all pure imaginary eigenvalues have multiplicity one.

It can be shown that the solution of the state equation (1), for a given initial vector $x_0 = x(t_0)$, is given as follows

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau. \tag{4}$$

We notice that this solution is a sum of two terms: the first term $e^{A(t-t_0)}x(t_0)$ represents the state evolution of the autonomous system ($u = 0$) while the second term $\displaystyle\int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau$ corresponds to the state evolution for the zero initial condition. This last term written as a convolution product of $e^{At}B$ with $u(t)$ is called the "input-to-state impulse matrix. From (1) and (4), the output response $y(t)$ is given by

$$y(t) = Ce^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} Ce^{A(t-\tau)}Bu(\tau)d\tau.$$

An important property of a dynamical system is its stability which means the ability of the autonomous system ($u = 0$) to recover its equilibrium point after being disturbed from it. Roughly speaking, stability means that for bounded inputs $u(t)$, the state-variable vector $x(t)$ will remain bounded for all times t. A stable system tends to return to its equilibrium state when perturbed from it. Conversely, perturbations are increased by an unstable system. Formally, the dynamical system (1) is asymptotically stable if for any initial condition $x(t_0)$, we have

$$\lim_{t\leftarrow\infty} x(t) = 0. \tag{5}$$

Using the expression (4) with $u = 0$, the limit condition (5) holds if and only if the matrix A has all its eigenvalues in the open left-half plane $\mathbb{C}^-$ and in this case the matrix A is called stable or Hurwitz.

There are several properties associated with stability. Clearly, if A is stable, then also A^{-1} and A^T are stable. Finally, if the product of matrices AB is stable, then also BA can be shown to be stable. It is also clear that, due to the relation between eigenvalues of A and poles of the transfer function, stability can also be formulated in terms of the poles of the transfer function $F(s)$.

Stability is a very natural and important property of physical structures but not strong for structures that contain no sources. A stable structure can become unstable if nonlinear components are connected to it. Another property called passivity is more stronger than stability. An LTI system is passive it is incapable of generating energy. The passivity of the transfer function $F(s)$ is defined as follows:

Definition 2.1.2. The function F is passive if F has no pole in $\mathbb{C}^+$. and $Re(F(s)) \geq 0$ for all $s \in \mathbb{C}^+$. In this case F is also called positive-real.

2.2. **Controllability, observability.** The controllability of a dynamical system is related to the ability of the system to attain a given state under the action of an appropriate control signal. If a state is not controllable, then it is not possible to move this state to another one by acting on the control input. If the matrix representing the dynamics of a non controllable state is stable, then the state is said to be stabilizable.
The observability is related to the possibility of evaluating the state of a system through output measurements. The notions of controllability and observability are due to Kalman [72, 73].

Definition 2.2.1. (Controllability) An LTI dynamical system is controllable if starting from zero initial state, any state can be reached via a suitable control.

Proposition 2.2.2. *The LTI dynamical system (1) is controllable if and only if the controllability matrix*

$$\mathcal{C} = [B, AB, A^2B, \ldots, A^{n-1}B]$$

is of full rank, i.e; $rank(\mathcal{C}) = n$. In this case the pair (A, B) is said to be controllable.

If $rank(\mathcal{C}) = k < n$, then $n - k$ is the number of the uncontrollable modes (the eigenvalues of the matrix A satisfying $rank([\lambda I - A \ \ B] < n$). If all uncontrollable modes are stable then the system is said to be stabilizable.

Definition 2.2.3. (Observability) An LTI dynamical system is observable if without control, different initial states lead to different outputs.

Proposition 2.2.4. *The LTI dynamical system (1) is observable if and only if the observability matrix*

$$\mathcal{O} = [C^T, A^TC^T, (A^2)^TC^T, \ldots, (A^{n-1})^TC^T]^T$$

is of full rank, i.e; $rank(\mathcal{C}) = n$. In this case the pair (A, C) is said to be observable.

If $rank(\mathcal{O}) = l < n$, then $n - l$ is the number of the unobservable modes (the eigenvalues of the matrix A satisfying $rank([\lambda I - A^T \ \ C^T]^T < n$).
The observability is linked to the possibility of evaluating the state of a system through output measurements. If a state is not observable there is no way to determine its evolution. If the dynamics of a non observable state is stable, then the state is said to be detectable.

Proposition 2.2.5. *A stable LTI system (1) is controllable if and only if the controllability Gramian given by (6) is positive definite and it is observable if and only if the observability Gramian (7) is positive definite.*

Definition 2.2.6. (Causality) A dynamical system is called causal if the output $y(t)$ depends only on the pasts inputs: $u(\tau), \tau \leq t$.

3. The transfer function

The state space representation is usually referred as an internal representation of a dynamical system because it involves the state variables x which are internal variables of the system. The input/output representation, also called external representation, is obtained by eliminating the state vector, between the state equation and the output equation with zero initial conditions.
To get the frequency domain description we apply the Laplace transform

$$\mathcal{L}(f(t) = \int_0^\infty f(t)e^{-st}dt$$

to the state equation (1), and we get

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

where $X(s) = \mathcal{L}(x(t))$ and $U(s) = \mathcal{L}(u(t))$. Therefore

$$X(s) = (sI - A)^{-1}BU(s)$$

and by substituting $X(s)$ in the output equation of (1), we get

$$Y(s) = F(s)\,U(s),$$

with

$$F(s) = C(sI - A)^{-1}B.$$

The rational function $F(s)$ is called the `transfer function` related to the dynamical system (1). The elements of this function are real rational functions. The transfer function $F(.)$ is stable if its poles lie in the open left-half plane $\mathbb{C}^-$.

We notice that two LTI systems $\left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ and $\left[\begin{array}{c|c} \widetilde{A} & \widetilde{B} \\ \hline \widetilde{C} & 0 \end{array}\right]$ are called equivalent if they have the same transfer function. It is easy to verify that for any nonsingular $n \times n$ matrix T, the LTI system

$$\left[\begin{array}{c|c} T^{-1}AT & T^{-1}B \\ \hline CT & 0 \end{array}\right]$$

is equivalent to the LTI system $\left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$. Therefore if the main concern is the output under some specific inputs, we have many choices of the state-space description. The choice of the matrix T is very important and the states are connected by the relation $x(t) = T\widetilde{x}(t)$.

We notice that the state-space approach has many advantages: it can be obtained directly from a real problem as in PDE control problems where the matrix A comes from the spatial discretization of the PDE, the control gives the input $u(t)$, the boundary conditions provide the matrix B while C is obtained from measurements of the output.

The transfer function F relates the Laplace transform of the output vector to that of the input vector. Each entry $F_{ij}(s)$ is a rational function representing the transfer function between the i-th input and the j-th output, all other inputs being set equal to zero. The rational function $F(s)$ can be expressed as a sum of a Taylor series around $(s = \infty)$ in the following form

$$F(s) = \frac{1}{s}C(I - \frac{A}{s})^{-1}B = \frac{1}{s}\sum_{i=0}^{\infty} M_i s^{-i},\ M_i = CA^iB,$$

where the matrix coefficients M_i are called the Markov parameters of F.

If A is nonsingular, the development of Neumann expansion around $s = 0$ gives the following expression

$$F(s) = \sum_{i=0}^{\infty} \widetilde{M}_i s^i,\ \tilde{M}_i = -CA^{-i-1}B.$$

The coefficients $\widetilde{M}_i$ are called the moments of F. Notice that the transfer function F is also the Laplace transform of the impulse response h defined by

$$h(t) = Cx(t) = Ce^{-tA}B,\ x(t_0) = 0.$$

A system defined by the triplet (A, B, C) is called a realization of the transfer function F and as we have seen earlier, this realization is not unique. A realization with the smallest dimension is called minimal. It means that the dynamical system is as economical as possible in terms of the number of independent parameters needed to describe the state of the system. A realization

is minimal if and only if it is controllable and observable.
Model reduction techniques consist in looking for a realization with a dimension m smaller than the dimension n of the system. In that case, the transfer function F is approximated by another transfer function F_m of small dimensions. This subject is connected to partial realization and Padé approximation; see [14, 15, 29, 30, 100].

4. Controllability and Observability Gramians

We assume that the LTI dynamical system is stable.

Definition 4.0.7. The controllability Gramian associated to the LTI system (1) is defined as

$$P = \int_0^\infty e^{tA} B B^T e^{tA^T} dt, \tag{6}$$

and the observability Gramian is defined by

$$Q = \int_0^\infty e^{tA^T} C^T C e^{tA} dt. \tag{7}$$

By using the Parseval relation, we obtain the following expressions of the Gramians

$$\begin{aligned} P &= \int_{-\infty}^{+\infty} (j\omega I - A)^{-1} B B^T (j\omega I - A^T)^{-1} d\omega, \\ Q &= \int_{-\infty}^{+\infty} (j\omega I - A^T)^{-1} C^T C (j\omega I - A)^{-1} d\omega. \end{aligned}$$

The two Gramians are the uniques solutions of the following coupled Lyapunov matrix equations

$$\begin{aligned} AP + PA^T + BB^T &= 0, \\ A^T Q + QA + C^T C &= 0. \end{aligned} \tag{8}$$

As can be seen from the expressions (6) and (7), the Gramians P and Q are at least positive semidefinite. It can be shown that the LTI system (1) is controllable iff the controllability Gramian P is positive definite and it is observable iff the observability Gramian is positive definite. We will see later that the product PQ plays an important role in model reduction.
Consider the the new equivalent LTI dynamical system

$$\widetilde{\mathbf{\Sigma}} \equiv \left[\begin{array}{c|c} T^{-1}AT & T^{-1}B \\ \hline CT & 0 \end{array}\right]$$

where T is a nonsingular matrix. Then the associated controllability and observability Gramians $\widetilde{P}$, and $\widetilde{Q}$ are expressed as

$$\begin{aligned} \widetilde{P} &= \int_0^\infty e^{t\widetilde{A}} \widetilde{B}\widetilde{B}^T e^{t\widetilde{A}^T} dt, \\ \widetilde{Q} &= \int_0^\infty e^{t\widetilde{A}^T} \widetilde{C}^T \widetilde{C}^T, e^{t\widetilde{A}} dt \end{aligned}$$

where $\widetilde{A} = T^{-1}AT$, $\widetilde{B} = T^{-1}B$ and $\widetilde{C} = CT$. Hence, we obtain

$$\widetilde{P} = T^{-1} P T^{-T}, \text{ and } \widetilde{Q} = T^T Q T, \tag{9}$$

and these last relations show that the Gramians of two equivalent LTI systems are not similar. However, the similarity is preserved for the product of the controllability and observability Gramians and we have

$$\widetilde{P}\widetilde{Q} = T^{-1} P Q T.$$

Remark 4.0.8. Discrete LTI systems are given as follows

$$\mathbf{\Sigma}_d \begin{cases} x(k+1) = Ax(k) + Bu(k);\ x(0) = x_0 \\ \\ y(k) = Cx(k) \end{cases}$$

If the $\mathbf{\Sigma}_d$ is d-stable (the eigenvalues of A are inside the unit disk), the associated controllability P_d and observability Q_d Gramians satisfy the following symmetric Stein matrix equations

$$AP_dA^T - P_d + BB^T = 0 \ \text{ and } \ A^TQ_dA - Q_d + C^TC = 0.$$

Also given as

$$P_d = \sum_{i=0}^{\infty} A^iBB^T(A^T)^i \ \text{ and } \ Q_d = \sum_{i=0}^{\infty}(A^T)^iC^TCA^i.$$

5. Different dynamical system norms

5.1. **The $\mathcal{H}_2$ norm.** In this subsection, we give the definition of the $\mathcal{H}_2$-norm of the transfer function $F(s)$ associated to the dynamical system $\mathbf{\Sigma}$ represented by (2).

Definition 5.1.1. The $\mathcal{H}_2$-norm of $F(s)$ is defined as

$$\|F(.)\|^2_{\mathcal{H}_2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} trace[F(j\omega)^TF(j\omega)]d\omega,$$

where j is the complex number $j^2 = -1$.
Consider the impulse response $g(t) = \mathcal{L}^{-1}[F(s)] = Ce^{tA}B$ where $\mathcal{L}$ is the Laplace transform

$$\mathcal{L}(g)(s) = \int_0^{\infty} g(t)e^{-st}dt = F(s).$$

Then using the Parseval relation

$$\int_0^{\infty} trace[g(t)^Tg(t)]dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} trace[F(j\omega)^TF(j\omega)]d\omega$$

the $\mathcal{H}_2$ norm can also be expressed as

$$\|F(.)\|^2_{\mathcal{H}_2} = \int_0^{\infty} trace[g(t)^Tg(t)]dt.$$

Therefore, the $\mathcal{H}_2$ norm could be calculated as follows

$$\|F(.)\|^2_{\mathcal{H}_2} = trace\left[B^T\ \left(\int_0^{\infty} e^{tA^T}C^TCe^{tA}dt\right)B\right].$$

Setting

$$Q = \int_0^{\infty} e^{tA^T}C^TCe^{tA}dt,$$

we get

$$\|F(.)\|^2_{\mathcal{H}_2} = trace(B^TQB).$$

Assuming that A is a stable matrix, the observability Gramian Q can be computed by solving the Lyapunov matrix equation (8). We notice that in a similar way, the $\mathcal{H}_2$ norm can be computed by using the controllability Gramian defined by (6). Therefore, $\mathcal{H}_2$ norm can be expressed as

$$\|F(.)\|^2_{\mathcal{H}_2} = trace(CPC^T).$$

The aim of a model reduction is to produce a low order m such that the error between the outputs $y(t) - y_m(t)$ is small over a large class of inputs. Different measures of approximation and different choices of input classes lead to different model reduction techniques.
Assume for example that we want to have

$$\max_{t>0} |y(t) - y_m(t)|$$

small uniformly over all inputs u with a bounded energy, that is

$$\int_0^\infty |u(t)|^2 dt \leq 1.$$

On the other hand

$$\begin{aligned}\max_{t>0} |y(t) - y_m(t)| &= \max_{t>0} \left| \frac{1}{2\pi} \int_{-\infty}^{+\infty} (Y(j\omega - Y_m(j\omega)\, e^{j\omega t} \right| \\ &\leq \frac{1}{2\pi} \int_{-\infty}^{+\infty} |Y(j\omega - Y_m(j\omega)| d\omega.\end{aligned}$$

Now since $Y(s) - Y_m(s) = (F(s) - F_m(s))U(s)$ it follows that

$$\begin{aligned}\frac{1}{2\pi} \int_{-\infty}^{+\infty} |Y(j\omega) - Y_m(j\omega)| d\omega &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} |F(j\omega) - F_m(j\omega)||U(j\omega)| d\omega \\ &\leq \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} |F(j\omega) - F_m(j\omega)|^2 \right)^{1/2} \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} |u(t)|^2 \right)^{1/2} \\ &\leq \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} |F(j\omega) - F_m(j\omega)|^2 \right)^{1/2} = \|F - F_m\|_{\mathcal{H}_2}.\end{aligned}$$

5.2. **The $\mathcal{H}_\infty$-norm.** In this subsection, we recall the well known $\mathcal{H}_\infty$-norm of a transfer function.

Definition 5.2.1. The $\mathcal{H}_\infty$ norm of the transfer function $F(.)$ is defined as

$$\|F(.)\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} \sigma_{max}(F(j\omega))$$

where σ_{max} denotes the largest singular value.

To approximate the $\mathcal{H}_\infty$-norm, we choose a set of frequencies $\Omega_N = \{\omega_1, \omega_2, \ldots, \omega_N\}$ and search for

$$\sup_{1 \leq k \leq N} \sigma_{max}(F(j\omega_k)) \approx \|F(.)\|_{\mathcal{H}_\infty}.$$

5.3. **The Hankel norm.** The Hankel singular values of a stable LTI system are the square roots of the product of the controllability and observability Gramians:

$$\sigma_i(F) = \sigma_i(\mathbf{\Sigma}) = \sqrt{\lambda_i(PQ)}$$

where P and Q are the Gramians associated the LTI dynamical system (1).

Definition 5.3.1. The Hankel norm of a stable LTI dynamical system is given by

$$\| F \|_{\mathcal{H}} = \max_i \sigma_i(F).$$

6. Lyapunov balanced truncation

A well known model reduction scheme is called Balanced Truncation and was first introduced by Mullis and Roberts [87] and later in the systems and control by Moore and Glover [47, 48, 86]; see also [9, 55, 95]. We assume here that the LTI system is stable, controllable and observable (in this case we call it also stable and minimal). Then the controllability and observability Gramians are unique positive definite. The concept of balanced truncation is to transform the original LTI system to an equivalent one in which the states that are difficult to reach are also difficult to observe. This reduces to finding a nonsingular matrix T such that the new Gramians $\widetilde{P}$ and $\widetilde{Q}$ given by (9) are such that

$$\widetilde{P} = \widetilde{Q} = diag(\sigma_1, \ldots, \sigma_n)$$

where σ_i is the i-th Hankel singular value of the LTI system; i.e.

$$\sigma_i = \sqrt{\lambda_i(PQ)}.$$

Let us see how to obtain the matrix T. Consider the Cholesky decompositions of the Gramians P and Q:

$$P = L_c L_c^T, \;\; Q = L_o L_o^T,$$

and consider also the singular value decomposition of $L_c^T L_o$ as

$$L_c^T L_o = Z \Sigma Y^T,$$

where Z and Y are unitary $n \times n$ matrices and Σ is a diagonal matrix containing the singular values. Let T be the matrix T defined by

$$T = L_c Z \Sigma^{1/2}.$$

Then it can be easily verified that

$$\widetilde{P} = \widetilde{Q} = \Sigma,$$

where Σ is also the diagonal matrix whose elements are the Hankel singular values $\sqrt{\lambda_i(PQ)}$ since PQ is similar to $\widetilde{P}\widetilde{Q}$. There are other possible ways for the construction of the matrix T. It was remarked by Glover [47] that the balanced transformation is not unique but unique up to a nonsingular transformation.

As the concept of balancing has the property that the states which are difficult to reach are simultaneously difficult to observe, then, a reduced model is obtained by truncating the states having this property, i.e., those which correspond to small Hankel singular values σ_i. We have the following theorem

Theorem 6.0.2. *Assume that the LTI dynamical system (1) is stable and minimal and having the following balanced realization*

$$\widetilde{\mathbf{\Sigma}} \equiv \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & 0 \end{array}\right],$$

with $\widetilde{P} = \widetilde{Q} = diag(\Sigma_m, \widetilde{\Sigma}_m)$, $\Sigma_m = diag(\sigma_1, \ldots, \sigma_m)$ *and* $\widetilde{\Sigma}_m = diag(\sigma_{m+1}, \ldots, \sigma_n)$.
Then, the reduced order model represented by

$$\widetilde{\mathbf{\Sigma}}_m \equiv \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & 0 \end{array}\right].$$

is stable and we have

$$\|F(.) - F_m(.)\|_{\mathcal{H}_\infty} \leq 2(\sigma_{m+1} + \ldots + \sigma_n).$$

The preceding theorem shows that if the neglected singular values $\sigma_{m+1}, \ldots, \sigma_n$ are small, then the reduced order LTI system is close to the original one.

Let us see now how to construct the low order model $\mathbf{\Sigma}_m$. We set

$$W_m = L_o Y_m \Sigma_m^{-1/2} \text{ and } V_m = L_c Z_m \Sigma_m^{-1/2},$$

where $\Sigma_m = diag(\sigma_1, \ldots, \sigma_m)$ and Z_m and Y_m correspond to the leading m columns of the matrices Z and Y given by the singular value decomposition (6).

The matrices of the reduced LTI system

$$\widetilde{\mathbf{\Sigma}}_m \equiv \left[\begin{array}{c|c} A_m & B_m \\ \hline C_m & 0 \end{array}\right],$$

are given by

$$A_m = W_m^T A V_m, \; B_m = W_m^T B \text{ and } C_m = C V_m.$$

Notice that $V_m W_m^T$ is an oblique projector, $\widetilde{P} W_m = V_m \Sigma_m$ and $\widetilde{Q} V_m = W_m \Sigma_m$.

The use of Cholesky factors in the Gramians P and Q is not applicable for large-scale problems.

Instead, and as we will see later, one can compute low rank approximations of P an Q and use them to construct an approximate balanced truncation model.
Let $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{C}$ be the following matrices

$$\widetilde{A} = \begin{pmatrix} A & 0 \\ 0 & A_m \end{pmatrix}, \ \widetilde{B} = \begin{pmatrix} B \\ B_m \end{pmatrix}, \ \widetilde{C} = \begin{pmatrix} C & C_m \end{pmatrix}.$$

Then, the Graminas corresponding to the error dynamical system

$$\widetilde{\mathbf{\Sigma}} \equiv \left(\begin{array}{c|c} \widetilde{A} & \widetilde{B} \\ \hline \widetilde{C} & 0 \end{array}\right),$$

are the solutions of the following Lyapunov matrix equations

$$\widetilde{A}\widetilde{P} + \widetilde{P}\widetilde{A}^T + \widetilde{B}\widetilde{B}^T = 0,$$

and

$$\widetilde{A}^T\widetilde{Q} + \widetilde{Q}\widetilde{A} + \widetilde{C}^T\widetilde{C} = 0.$$

Therefore, the Hankel norm of the error can be expressed as

$$\|F(s) - F_m(s)\|_{\mathcal{H}} = \sqrt{\lambda_{max}(\widetilde{P}\widetilde{Q})}.$$

Remark 6.0.3. Lyapunov balanced and also other balanced methods try to approximate the full-order model over all frequencies. However, in many applications we are only interested in some given frequency interval. In these cases the frequency weighted balanced reduction is used; see [46, 78, 109, 110].

7. Numerical methods for solving Lyapunov and Sylvester matrix equations

7.1. **Direct methods.** We consider the following Sylvester matrix equation

$$AX + XD = F, \tag{10}$$

where A, D, X and F are $n \times n$, $q \times q$, $n \times q$ and $n \times q$ matrices, respectively. Remark first that equation (10) has a unique solution if and only if $\Gamma(A) \cap \Gamma(D) = \emptyset$ where $\Gamma(A)$ denotes the set of eigenvalues of the matrix A . The case $D = A^T$ leads to the Lyapunov matrix equation. The Sylvester matrix equation (10) can be transformed to the following linear system of equations

$$(I_q \otimes A + D^T \otimes I_n)vec(X) = vec(F),$$

For small and moderate problems, a robust and efficient method for numerically solving Lyapunov or Sylvester equations was proposed by Bartels and Stewart in [18]. This algorithm transforms the matrices A and D into real Schur forms $R_A = Q^T AQ$ and $R_D = U^T DU$ where R_A and R_D are quasi upper triangular, Q and U are orthogonal. This gives the new Sylvester matrix equation

$$R_A Y + Y R_D = C_1$$

in which $C_1 = Q^T FU$ and $X = QYU^T$. This last equation is solved by back substitution.
The Hesseberg-Schur algorithm proposed by Golub and Van Loan [49] is a modification of the Bartels-Stewart algorithm. The matrix A is reduced only to upper Hessenberg form: $A = VH_AV^T$ and the reduced matrix equation

$$H_A Z + Z R_D = C_2,$$

in which $C_2 = V^T FU$ can be solved by solving upper Hessenberg linear systems and then $X = VZU^T$.
One can also use classical Krylov subspace methods such the GMRES [98] to solve shifted linear systems. In fact, consider the real-Schur form of the matrix D: $D = UR_DU^T$ where U is orthogonal and R_D is block-upper triangular (the blocks are of size 1×1 or 2×2). The matrix equation (10) is written as

$$AY + Y R_D = \widetilde{F}, \tag{11}$$

in which $Y = UX$ and $\widetilde{F} = FU$. Assuming for simplicity that the eigenvalues of the matrix D are real then R_D is upper triangular. Solving (11) is equivalent to solve the following p shifted linear systems

$$\begin{aligned}(A - r_{1,1}I)y_1 &= d_1 \\ (A - r_{k,k}I)y_k &= d_k + \sum_{i=1}^{k-1} r_{i,k}y_i,\ k = 2, \ldots, p.\end{aligned}$$

where $R_D = [r_{i,j}]$; $Y = [y_1, \ldots, y_k]$ and $\widetilde{D} = [d_1, \ldots, d_k]$.

Since these linear systems cannot be solved simultaneously, this will require a higher time as compared to the block approach. The second drawback is the fact that since the second right-hand side of each linear system depends on the solutions of the preceding linear systems the scheme will not produce generally good approximations. In the case of Lyapunov matrix equations ($D = A^T$) of moderate size and when the right hand side is positive semidefinite ($C = C_1C_1^T$), a method was proposed in [59] by exploiting the structure of the matrix C to get the Cholesky factor of X without computing the solution X; see also [7, 8].

7.2. **Block and matrix Krylov subspace methods.**

7.2.1. *Block Arnoldi-based methods.* In this section, we develop block Krylov subspace methods for computing approximate solutions to large-scale Lyapunov or Sylvester matrix equations. Let V be a matrix of size $n \times p$. Block Krylov subspaces could be divided into three parts:

- The block Krylov subspace $\mathbb{K}_m^b(A, V)$ is the a subspace of $\mathbb{R}^n$ generated by the columns of the matrices $V, AV, \ldots, A^{m-1}V$, see [64, 98]
$$\mathbb{K}_m^b(A, V) = Range([V, AV, \ldots, A^{m-1}V]).$$
- The extended block Krylov subspace $\mathbb{K}_m^{eb}(A, V)$ is the subspace of $\mathbb{R}^n$ generated by the the columns of the matrices $A^{-m}V, \ldots, A^{-1}V, V, AV, \ldots, A^{m-1}V$, see [60, 101]
$$\mathbb{K}_m^{eb}(A, V) = Range([A^{-m}V, \ldots, A^{-1}V, V, AV, \ldots, A^{m-1}V]).$$
Notice that the extended block Krylov subspace $\mathbb{K}_m^{eb}(A, V)$ is the sum of two block Krylov subspaces
$$\mathbb{K}_m^{eb}(A, V) = \mathbb{K}_m^b(A, V) + \mathbb{K}_m^b(A^{-1}, A^{-1}V).$$
- The rational block Krylov $\mathbb{K}_m^{rb}(A, V)$ is the subspace of $\mathbb{R}^n$ generated by the columns of the matrices $(s_1I - A)^{-1}V, \ldots, (s_mI - A)^{-1}V$, see [1, 36, 37, 51, 97]
$$\mathbb{K}_m^{rb}(A, V) = Range([(s_1I - A)^{-1}V, \ldots, (s_mI - A)^{-1}V]),$$
where$s_1, \ldots, s_m$ are some shift-parameters to be selected a priori or a posteriori. Some recent works have been devoted to the good selection of the shift-parameters in the rational-based methods; see [1, 16, 36, 37].

The well known block Arnoldi algorithm allows us to compute an orthonormal basis $\widetilde{V}_m = \{V_1, \ldots, V_m\}$ of the block Krylov subspace $\mathbb{K}_m^b(A, V)$, see [98] while an extended block Arnoldi algorithm was developed in [60, 101] to obtain such a basis for the extended block Krylov subspace $\mathbb{K}_m^{eb}(A, V)$. For the rational block Krylov subspace $\mathbb{K}_m^{rb}(A, V)$, a block version of the classical rational Arnoldi was recently developed in [1].

Next we show how to apply these block methods for solving large-scale Lyapunov matrix equations. Here we consider only the extended block Arnoldi method since it gives generally better results than the classical block Arnoldi algorithm.
The following algorithm allows us to compute an orthonormal basis of the extended Krylov subspace $\mathbb{K}_m^b(A, V)$. This basis contains information on both A and A^{-1} where m is some fixed integer which limits the dimension of the constructed basis.

The extended block Arnoldi process is described as follows:

Algorithm 1 The extended block Arnoldi algorithm (EBA)

- *Inputs:* A an $n \times n$ matrix, V an $n \times p$ matrix and m an integer.
- Compute the QR decomposition of $[V, A^{-1}V]$, i.e., $[V, A^{-1}V] = V_1\Lambda$;
  Set $\mathcal{V}_0 = [\,]$;
- For $j = 1, \ldots, m$
- Set $V_j^{(1)}$: first p columns of V_j; $V_j^{(2)}$: second p columns of V_j
- $\mathcal{V}_j = [\mathcal{V}_{j-1}, V_j]$; $\hat{V}_{j+1} = \left[A\,V_j^{(1)}, A^{-1}\,V_j^{(2)}\right]$.
- Orthogonalize $\hat{V}_{j+1}$ w.r. to $\mathcal{V}_j$ to get V_{j+1}, i.e.,
  for $i = 1, 2, \ldots, j$
  $H_{i,j} = V_i^T\,\hat{V}_{j+1}$;
  $\hat{V}_{j+1} = \hat{V}_{j+1} - V_i\,H_{i,j}$;
  endfor
- Compute the QR decomposition of $\hat{V}_{j+1}$, i.e., $\hat{V}_{j+1} = V_{j+1}\,H_{j+1,j}$.
- endFor.

Since the above algorithm involves implicitly a Gram-Schmidt process, the obtained blocks $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$ ($V_i \in \mathbb{R}^{n\times 2s}$) have their columns mutually orthogonal provided none of the upper triangular matrices $H_{j+1,j}$ are rank deficient.
Hence, after m steps, Algorithm 1 builds an orthonormal basis $\mathcal{V}_m$ of the extended block Krylov subspace $\mathbb{K}_m^{eb}(A, V)$ and a block upper Hessenberg matrix $\mathbb{H}_m$ whose non zeros blocks are the $H_{i,j}$'s. Note that each submatrix $H_{i,j}$ ($1 \le i \le j \le m$) is of order $2p$.

Let $\mathcal{T}_m \in \mathbb{R}^{2ms\times 2ms}$ be the restriction of the matrix A to the extended block Krylov subspace $\mathbb{K}_m^b(A, V)$:

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

It is shown in [101] that $\mathcal{T}_m$ is also block upper Hessenberg with $2p \times 2p$ blocks. Moreover, a recursion is derived to compute $\mathcal{T}_m$ from H_m without requiring matrix-vector products with A. For more details on how to compute $\mathcal{T}_m$ from $\mathcal{H}_m$, we refer to [101]. Other algebraic properties on the extended block algorithm were recently developed in [2]. We note that for large and non-structured 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.
Suppose that m steps of Algorithm 1 have been run, then we have the following relations

$$\mathcal{V}_m^T\,\mathcal{V}_m = I_{2p},$$

and

$$\begin{aligned} A\,\mathcal{V}_m &= \mathcal{V}_{m+1}\,\bar{\mathcal{T}}_m, \\ &= \mathcal{V}_m\,\mathcal{T}_m + V_{m+1}\,T_{m+1,m}\,E_m^T, \end{aligned} \tag{12}$$

where $T_{i,j}$ is the $2p \times 2p$ (i,j)-block of $\mathcal{T}_m$, $\bar{\mathcal{T}}_m = \mathcal{V}_{m+1}^T\,A\,\mathcal{V}_m$ and $E_m = [O_{ps\times 2(m-1)p}, I_{2p}]^T$ is the matrix of the last $2p$ columns of the $2mp \times 2mp$ identity matrix I_{2mp}.

Let see now how to allpy the extended block Arnoldi method for solving large scale Lyapunov or Sylvester matrix equation with a low rank right hand side. Consider the large scale Sylvester matrix equation

$$AX + XD + C_1C_2^T = 0, \tag{13}$$

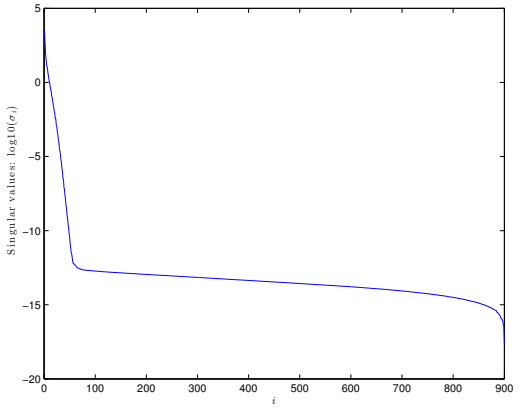

FIGURE 1. The singular values of the exact solution

where A and D are assumed to be stable and large of size $n \times n$ and $q \times q$, respectively. The matrices C_1 and C_2 are of size $n \times p$ and $q \times p$ with $p \ll n$ and $p \ll q$. Since the matrices A and D are assumed to be stable, the exact unique solution of (13) can be expressed as (see [77])

$$X = -\int_0^\infty e^{tA} C_1 C_2^T e^{tD} dt.$$

For general matrices A and D with $\lambda_i(A) + \lambda_j(D) \neq 0$, the solution X can also be represented by the following integral expression

$$X = -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \frac{(\lambda I - A)^{-1} C_1 C_2^T (\mu I - D)^{-1}}{\lambda + \mu} d\lambda d\mu,$$

where Γ_1 and Γ_2 are contours containing and sufficiently close to, the spectra of A and D, respectively.

In general, the solution X of the Sylvester (or Lyapunov) matrix equation (13) has a numerical low rank as shown in Figure 1. In this figure, we plotted the log10 of the singular values of the exact solution obtained by the Bartels-Stewart method [18]. For this figure the matrix A was PDE900 from Harwell Boeing collection, $D = A^T$ and $C_1 = C_2^T$ were matrices whose elements are uniformly distributed in [0 1]. We set $\mathcal{R}(X) = AX + XD + C_1 C_2^T$ the residual corresponding to X. Since the equation (13) has a low rank right-hand side, block Krylov methods allow us to compute low rank approximate solutions. So we will seek for approximate solutions having the form

$$X_m = \mathcal{V}_m Y_m \mathcal{W}_m^T,$$

where the matrix $\mathcal{V}_m$ is obtained by applying the extended block Arnoldi algorithm to the pair (A, C_1) and $\mathcal{W}_m$ is obtained by applying the extended block Arnoldi algorithm to the pair (D, C_2). The matrix Y_m is computed by imposing the following Galerkin condition

$$\mathcal{V}_m^T \mathcal{R}(X_m) \mathcal{W}_m = 0.$$

Therefore, using the relations (12), we obtain the low order Sylvester (or Lyapunov) matrix equation

$$\mathcal{T}_{m,A} Y_m + Y_m \mathcal{T}_{m,D}^T + C_{1,m} C_{2,m}^T = 0,$$

where $C_{1,m} = {\mathcal{V}_m}^T C_1$, $C_{2,m} = {\mathcal{W}_m}^T C_2$, $\mathcal{T}_{m,A} = {\mathcal{V}_m}^T A \mathcal{V}_m$ and $\mathcal{T}_{m,D} = {\mathcal{W}_m}^T D \mathcal{W}_m$.

The solution Y_m of the low order problem is computed by applying the Bartels-Stewart [18] algorithm. When using block Krylov methods, one can compute the norm of the residual without having to compute the approximate solution which will be given in a factored form

when the convergence is achieved and this is very important for saving memory. For that we have the following result, [60, 61, 68, 102]. At step m, we have

Theorem 7.2.1.

$$\| R(X_m) \|_F^2 = \| Y_m E_m (\mathcal{T}_{m+1,m}^D)^T \|_F^2 + \| \mathcal{T}_{m+1,m}^A E_m^T Y_m \|_F^2,$$

where $\mathcal{T}_{m+1,m}^A$ is the $(m+1) \times m$, $2p$-block of the matrix $\bar{\mathcal{T}}_{m,A}$ and E_m is the matrix of the last $2p$ columns of the $2mp \times 2mp$ identity matrix I_{2mp}
To save memory, the approximate solution $X_m = \mathcal{V}_m Y_m \mathcal{W}_m^T$ could be given as a product of two matrices of low ranks. In fact, consider the singular value decomposition of the matrix

$$Y_m = \widetilde{U}\, \widetilde{\Sigma}\, \widetilde{V}^T$$

where $\widetilde{\Sigma}$ is the diagonal matrix of the singular values of Y_m sorted in decreasing order. Let $\widetilde{U}_l$ be the $2ms \times l$ matrix of the first l columns of $\widetilde{U}$ corresponding to the l singular values of magnitude greater than some tolerance *dtol* and the same notation for $\widetilde{V}_l$. We obtain the truncated singular value decomposition $Y_m \approx \widetilde{U}_l\, \widetilde{\Sigma}_l\, \widetilde{V}_l^T$ where $\widetilde{\Sigma}_l = \mathrm{diag}[\sigma_1, \ldots, \sigma_l]$. Setting $\mathcal{Z}_m^{(1)} = \mathcal{V}_m\, \widetilde{U}_l\, \widetilde{\Sigma}_l^{1/2}$, and $\mathcal{Z}_m^{(2)} = \mathcal{W}_m\, \widetilde{V}_l\, \widetilde{\Sigma}_l^{1/2}$ it follows that

$$X_m \approx \mathcal{Z}_m^{(1)} \left(\mathcal{Z}_m^{(2)}\right)^T. \tag{14}$$

Then the storage is reduced from $O(n^2)$ to $O(nl)$. We note that such low rank schemes are the only efficient methods that can effectively solve very large sparse Lyapunov equations. Other numerical low rank methods for solving large Lyapunov or Sylvester matrix equations could be found in [21, 41, 54, 55, 64, 66, 67, 71, 99, 104]. The expression (14) can be used to get approximate solutions to the controllability $P \approx \mathcal{Y}_m \mathcal{Y}_m^T$ and observability $Q \approx \mathcal{Z}_m \mathcal{Z}_m^T$ and then to use these factored forms to construct a reduced order model by an approximate balanced truncation.

7.2.2. *The block Lanczos-based method.* Let V and W be two initial blocs of $\mathbb{R}^{n \times p}$, then the nonsymmetric block Lanczos algorithm [12, 50] applied to the pairs (A, V) and (A^T, W) generates two sequences of bi-orthonormal $n \times p$ matrices $\{V_i\}$ and $\{W_j\}$, such that

$$\mathbb{K}_m(A, V) = \mathit{Range}([V_1, V_2, \ldots, V_m]),$$

and

$$\mathbb{K}_m(A^T, W) = \mathit{Range}([W_1, W_2, \ldots, W_m]).$$

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

$$\begin{cases} W_j^T V_i = 0_p, \ \mathit{if}\ i \neq j, \\ W_i^T V_i = I_p. \end{cases}$$

In [12], a stable version of the nonsymmetric block Lanczos algorithm was proposed. This algorithm, named Adaptive Block Lanczos (ABLE), is summarized as follows.

Setting $\mathbb{V}_m = [V_1, V_2, \ldots, V_m]$ and $\mathbb{W}_m = [W_1, W_2, \ldots, W_m]$, we have the following block Lanczos relations

$$A\mathbb{V}_m = \mathbb{V}_m \mathbb{T}_m + V_{m+1} B_{m+1} \widetilde{E}_m^T, \ \text{ and } \ A^T \mathbb{W}_m = \mathbb{W}_m {\mathbb{T}_m}^T + W_{m+1} C_{m+1}^T \widetilde{E}_m^T,$$

where $\widetilde{E}_m$ is last $mp \times p$ block of the identity matrix I_{mp} and T_m is the block tridiagonal matrix defined by

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

We can also use the block Lanczos method to compute approximate solutions to the large Gramians P and Q related to the dynamical system (1); see [63]. Recently, a new rational block Lanczos method was developed in [16] and applied for model order reduction techniques.

Algorithm 2 The nonsymmetric block Lanczos algorithm (ABLE)

(1) **Inputs:** $A \in \mathbb{R}^{n\times n}, V, W \in \mathbb{R}^{n\times p}$ and $m \in \mathbb{R}$.
(2) **Initialize :** $V_0 = W_0 = 0_p$ and $C_1 = B_1 = 0_p$.
(3) For $j = 1, \ldots, m$
- $S_j = AV_j$ 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 of $W_{j+1}^T V_{j+1} = P_j D_j Q_j^T$,
- $V_{j+1} = V_{j+1} Q_j D_j^{-1/2}$ 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 $C_{j+1} = D_j^{1/2} P_j^T C_{j+1}$,

(4) end For.

7.2.3. *Global Arnoldi type methods.* The global Krylov subspace $\mathbb{K}_m^g(A, V)$ is the subspace of $\mathbb{R}^{n\times p}$ generated by the matrices $V, AV, \ldots, A^{m-1}V$, see [65]

$$\mathbb{K}_m^g(A, V) = span\{V, AV, \ldots, A^{m-1}V\}.$$

The global Arnoldi algorithm [65] computes an F-orthonormal basis $\{V_1, \ldots, V_m\}$ of the matrix Krylov subspace $\mathbb{K}_m^g(A, V)$ satisfying

$$\langle V_i, V_j\rangle_F = \delta_{i,j},\ i, j = 1, \ldots, m \tag{15}$$

where $\delta_{i,j} = 0$ if $i \neq j$ and 1 if $i = j$.

The global Arnoldi algorithm is summarized as follows

Algorithm 3 The modified global Arnoldi algorithm

(1) Compute the $n \times p$ matrix V_1 such that $V_1 = V/ \parallel V \parallel_F$.
(2) For $j = 1, \ldots, m$
- $\tilde{V} = AV_j$,
- for $i = 1, 2, \ldots, j$,
  (a) $h_{i,j} = \text{trace}(V_i^T \tilde{V})$,
  (b) $\tilde{V} = \tilde{V} - h_{i,j} V_i$,
- end
- $h_{j+1,j} = \parallel \tilde{V} \parallel_F$,
- $V_{j+1} = \tilde{V}/h_{j+1,j}$.

(3) End

Setting $\mathcal{V}_m^g = [V_1, \ldots, V_m]$, and using the $\diamond$ product, the F-orthogonality relations (15) could be written as

$$\mathcal{V}_m^{g\,T} \diamond \mathcal{V}_m^g = I.$$

The global Arnoldi process produces an upper Hessenberg matrix $\bar{\mathcal{H}}_m = [h_{i,j}]$, $i = 1, \ldots, m+1$ and $j = 1, \ldots, m$ and an upper Hessenberg matrix $\mathcal{H}_m$ obtained by deleting the last column of $\bar{\mathcal{H}}_m$. We have the following algebraic relations

$$\mathcal{H}_m = \mathcal{V}_m^{g\,T} \diamond (A\mathcal{V}_m^g),$$

and

$$\begin{aligned} A\mathcal{V}_m^g &= \mathcal{V}_{m+1}^g(\bar{\mathcal{H}}_m \otimes I_p) \\ &= \mathcal{V}_{m+1}^g(\mathcal{H}_m \otimes I_p) + h_{m+1,m} V_{m+1}(e_m^T \otimes I_p), \end{aligned}$$

where e_m is the last column of the identity matrix I_m.
When applying the global Arnoldi process to solve the Sylvester (or Lyapunov) low rank right hand side matrix equation (13), we consider approximations of the form

$$X_m = \mathcal{V}_m^g (Y_m \otimes I_p) {\mathcal{W}_m^g}^T,$$

where $\mathcal{V}_m^g$ and $\mathcal{W}_m^g$ are the F-orthonormal matrices (${\mathcal{V}_m^g}^T \diamond \mathcal{V}_m^g = I_p$ and ${\mathcal{W}_m^g}^T \diamond {\mathcal{W}_m}^g = I_p$) obtained by applying simultaneously the global Arnoldi algorithm to the pairs (A, C_1) and (D^T, C_2). Using the Galerkin condition we get the low order Sylvester matrix equation

$$\mathcal{H}_{m,A} Y_m + Y_m \mathcal{H}_{m,B} + \|C_1\|_F \|C_2\|_F e_1 e_1^T = 0,$$

with $e_1 = (1, 0 \ldots, 0)^T$. We can also show that

Theorem 7.2.2. [67] *For the globlal Arnoldi approximations, we have the following results:*

(1) $\|\mathcal{R}(X_m)\|_F^2 \leq (h_{m+1,m}^A)^2 \|e_m^T Y_m\|_2^2 + (h_{m+1,m}^D)^2 \|Y_m e_m\|_2^2$.
(2) *The computed solution X_m could be approximated by the product of two matrices of low rank*

$$X_m \approx Z_m^A Z_m^D.$$

As for the block Arnoldi-based methods, the last expression of the approximate solution could be used to obtain a truncated balanced approximation to get a low order dynamical sytem; see for example [17, 62].

Remark 7.2.3. To get the low order Sylvester or Lyapunov matrix equations, we used the classical Galerkin condition. Recently in [4, 79], new Krylov projection methods were obtained by imposing a minimization property instead of the Galerkin condition. When using the extended block Arnoldi, the problem to solve is given by

$$Y_m = \arg \min_{X_m = \mathcal{V}_m Y \mathcal{V}_m^T} \|\mathcal{R}(X_m)\|_F.$$

This matrix least squares problem leads to a general linear matrix equation $\mathcal{L}(Y) = \mathcal{C}$ that was solved by applying the global conjugate gradient [41] and the global LSQR to the pair $(\mathcal{L}, \mathcal{C})$.

7.2.4. *The rational methods.* In this part, we consider the rational block Arnoldi algorithm [1] for computing an orthonormal basis of the rational block Krylov subspace defined for a given matrix $V \in \mathbb{R}^{n \times p}$ as

$$\mathbf{K}_m^{rb}(A, V) = \text{Range}(\{V, (A - s_1 I)^{-1} V, \ldots, \prod_{i=1}^{m} (A - s_i I)^{-1} B\}).$$

The rational block Arnoldi algorithm generates a sequence of $n \times p$ blocks $\{V_1, \ldots, V_m\}$ whose columns form an orthonormal basis of the rational block Krylov subspace $\mathbf{K}_m^{rb}(A, V)$. The algorithm is described as follows

The shifts $s_1, \ldots, s_m$ are chosen a priori or a posteriori during the process. After m steps, the rational block Arnoldi algorithm generates a block matrix $\mathcal{V}_m = [V_1, \ldots, V_m] \in \mathbb{R}^{n \times mp}$ whose columns form an orthonormal basis of the rational block Krylov subspace $\mathbf{K}_m^{rb}(A, V)$ and an upper $(m+1)p \times mp$ block Hessenberg matrix $\overline{\mathcal{H}}_m$ whose blocks $H_{i,j}$ are defined by Algorithm 4. The $mp \times mp$ upper block Hessenberg matrix $\mathcal{H}_m$ is obtained from $\overline{\mathcal{H}}_m$ by deleting its last p-rows. In the sequel we will also use the restriction matrix $\mathcal{T}_m$ defined by $\mathcal{T}_m := \mathcal{V}_m^T A \mathcal{V}_m$ and give some new algebraic relations generalizing the well known Arnoldi-like relation given for the classical case.

Proposition 7.2.4. *Let $\mathcal{V}_m$, $\overline{\mathcal{H}}_m$ and $\mathcal{H}_m$ be the matrices generated by the rational block Arnoldi algorithm. Let $\mathcal{S}_m$ be the block-diagonal matrix $blkdiag(s_1 I_p, \ldots, s_m I_p)$ where $\{s_1, ..., s_m\}$ denotes the set of shifts used in the algorithm. Then we have the following relation*

$$\mathcal{T}_m := \mathcal{V}_m^T A \mathcal{V}_m = (I_{mp} + \mathcal{H}_m S_m - \mathcal{V}_m^T A V_{m+1} H_{m+1,m} E_m^T) \mathcal{H}_m^{-1},$$

Algorithm 4 The Rational Block Arnoldi Algorithm

- *Input:* $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$ and a fixed integer m.
- Compute $V_1 = QR(B)$, $\mathcal{V}_1 = [V_1]$.
- For $j = 1, \ldots, m-1$
  (1) $\widetilde{V}_{j+1} = (A - s_{j+1}I)^{-1}V_j$.
  (2) Orthogonalization step:
  For $i = 1, 2, \ldots, j$
  $H_{i,j} = V_i^\top \widetilde{V}_{j+1}$;
  $\widetilde{V}_{j+1} = \widetilde{V}_{j+1} - V_i H_{i,j}$;
  End For
  (3) $QR\big(\widetilde{V}_{j+1}\big) = V_{j+1}H_{j+1,j}$.
  (4) $\mathcal{V}_{j+1} = [\mathcal{V}_j,\ V_{j+1}]$.
- End For.

where $E_m^T = [0_p, \ldots, 0_p, I_p] = (e_m^T \otimes I_p)$.

The rational block Arnoldi algorithm can be applied for solving Lyapunov or Sylvester matrix equationns in the same way as we explained for the extended block Arnoldi. For more details on how to select the shift-parameters $s_1, \ldots, s_m$, we refer to [1, 36, 37, 51]. We notice that we can also define a rational version of the block Lanczos method [16, 52].

7.3. **ADI-Smith type methods.** In this subsection, we consider here the low rank right hand side Lyapunov matrix equation

$$AX + XA^T + BB^T = 0, \tag{16}$$

where $B \in \mathbb{R}^{n\times p}$, $p \ll n$ and we assume that A is stable.

7.3.1. *The ADI iteration.* The ADI method was originally proposed by Peaceman and Rachford [90] for linear systems derived from the discretization of elliptic boundary value problems. Later, the method was used to solve standard Lyapunov and Sylvester matrix equations; see [19, 33, 82, 83, 91, 107].
Starting from $X_0^A = 0$, the ADI iteration for the solution of (16) proceeds by alternating between the two following linear systems

$$\begin{cases} (A + \mu_i\, I_n)\, X^A_{i-\frac{1}{2}} = -BB^T - X^A_{i-1}\,(A^T - \mu_i\, I_s), \\ X^A_i\,(A + \mu_i^*\, I_s) = -BB^T - (A - \mu_i^*\, I_n)\, X^A_{i-\frac{1}{2}},\ i = 1, 2, \ldots. \end{cases} \tag{17}$$

The scalars μ_j are real or complex parameters in $\mathbb{C}^-$ chosen in order that the approximated solutions X_i converge rapidly to the exact solution X of the Sylvester equation (16) as i increases. The two equations (17) are equivalent to

$$\begin{aligned} X_i^A \;=\; & (A - \mu_i^*\, I)\,(A + \mu_i\, I)^{-1}\, X^A_{i-1}\ (A^T + \mu_i^*\, I)^{-1}(A^T - \mu_i\, I) \\ & -2\mathcal{R}e(\mu_i)\,(A + \mu_i\, I)^{-1}\, BB^T\,(A^T + \mu_i^*\, I)^{-1}. \end{aligned}$$

When l is the number of the used μ_i shifts, the spectral radius $\rho_{ADI}(A)$ is given by

$$\rho_{ADI} = \rho\Big(\prod_{i=1}^{m}(A - \mu_i^*\, I)\,(A + \mu_i\, I)^{-1}\Big),$$

determines the rate of convergence of the ADI method. The minimization of this spectral radius, leads to the following ADI min-max problem:

$$\{\mu_1, \mu_2, \ldots, \mu_m\} = arg \min_{\{\mu_1,\mu_2,\ldots,\mu_m\}\in\mathbb{C}_-} \Big(\max_{\lambda\in\Lambda(A)} \frac{|\,(\lambda - \mu_1^*)\ldots(\lambda - \mu_m^*)\,|}{|\,(\lambda + \mu_1)\ldots(\lambda + \mu_m)\,|}\Big), \tag{18}$$

where $\Lambda(A)$ denotes the spectrum of the matrix A. For symmetric matrices, the minimax problem (18) has been solved [107], but in the general case where the eigenvalues of the matrix A are not real this is still an open problem. The performance of ADI-based methods is related to the choice of the parameters μ_i. The classical approach is to cover the spectrum of the matrix A by a domain $\Omega \subset \mathbb{C}_-$ and then to solve the minimax problem with respect to Ω. In [91, 93] a heuristic procedure was proposed to find sub-optimal parameters. This technique first generates a discrete set which approximates the spectrum $\Lambda(A)$ using a pair of Arnoldi processes. The first one acts on the matrix A and produces k_+ Ritz values which tend to approximate the eigenvalues far from the origin. The second process, acting on the matrix A^{-1}, produces k_- Ritz values close to the origin. The set of the ADI shift parameters is then chosen as a subset of these Ritz values.
If the matrix A is diagonalizable $A = \mathbb{V}D\mathbb{V}^{-1}$, then we have

$$\|X_m^A - X\|_F \leq \kappa(\mathbb{V})^2 \rho_{ADI}^2 \, \|X\|_F,$$

where X is the exact solution of the lyaponov matrix equation (16), $\mathbb{V}$ is the matrix of eigenvectors of A and $\kappa(\mathbb{V})$ is the condition number of $\mathbb{V}$.
We notice that similar results are also obtained when applying the ADI method for solving Sylvester matrix equations; see for example [19, 26].
Each step of the ADI iteration requires a sparse direct factorization of the matrix $(A + \mu_i I)$ and triangular solves from this factorization. If the shifts are complex, the operations should be done in complex arithmetic. However, the numerical tests show that working with real shifts returns generally good results as compared to the complex shifts.

7.3.2. *Smith's method.* Let μ be a scalar such that $\mu < 0$. Then it can be verified that the Lyapunov matrix equation (16) is equivalent to the following symmetric Stein matrix equation

$$\mathcal{A}_\mu X \mathcal{A}_\mu^T - X + B_\mu B_\mu^T = 0, \tag{19}$$

where

$$\mathcal{A}_\mu = (A - \mu I)(A + \mu I)^{-1}, \; B_\mu = \sqrt{-2\mu}(A + \mu I)^{-1} B.$$

Therefore, since A is a stable matrix, $\mathcal{A}_\mu$ is d-stable ($\rho(\mathcal{A}_\mu) < 1$) and then the unique solution of the Stein equation (19) is given by

$$X = \sum_{i=0}^{\infty} \mathcal{A}_\mu^i B_\mu B_\mu^T (\mathcal{A}_\mu^T)^i.$$

The Smith sequence of approximate solutions to (19) is defined from $X_0^S = 0$ by

$$X_i^S = \sum_{k=1}^{i} \mathcal{A}_\mu^{k-1} B_\mu B_\mu^T (\mathcal{A}_\mu^T)^{k-1},$$

which is equivalent to

$$X_0^S = 0, \; X_{i+1}^S = \mathcal{A}_\mu X_i^S \mathcal{A}_\mu^T + B_\mu B_\mu^T.$$

For more details, see [9, 91, 93]. The connection between the ADI method and the rational Krylov subspace method for solving Lyapunov matrix equations was recently developed in [38].
The ADI iteration converges slowly when using only one shift and then one has to consider l shifts in a cyclic manner . This gives the well known Smith-l iteration

$$X_i^{Sl} = \sum_{k=0}^{i-1} \mathcal{A}_d^k X_l^A (\mathcal{A}_d^k)^T,$$

where $\mathcal{A}_d = \prod_{j=1}^{l} (A - \mu_i I)(A + \mu_i)^{-1}$ and X_l^A is the l-th iterate produced by the ADI iteration with the shifts $\mu_1, \ldots, \mu_l$. The main drawback of these Smith-type methods is the storage

requirement. However, for many problems, it was observed that the solution has a numerical low rank and this leads to low rank extensions of the ADI method. These methods are the Low-Rank (LR-ADI) [91], the Cholesky-Factored (CF-ADI) [82], the LR-Smith(l) [91]. The key idea in these low rank versions is to compute the approximate solution in factored forms $X_i = Z_i Z_i^T$, where Z_i has a low rank. As one needs to store only the factor Z_i, this is very important for the storage requirement especially when dealing with large problems. For example, the LR-ADI method produces iterates X_i^A in the factored form

$$X_i^A = Z_i^A (Z_i^A)^T,$$

where

$$Z_i^A = [(A - \mu_i I)(A + \mu_i)^{-1} Z_{i-1}^A, \quad \sqrt{-2\mu_i}(A + \mu_i)^{-1} B],$$

with $Z_1^A = \sqrt{-2\mu_1}(A + \mu_1)^{-1} B]$. For more details on the LR-Smith(l) and CF-ADI, see [9, 82, 91].

To improve the computational cost per step, a new strategy called the modified low-rank (LR-Smith(l)) method was proposed in [54] in which the idea is to compute the singular value decomposition of the iterate at each step and replace this iterate with its best low-rank approximation. Some convergence results on these methods are found in [9, 54].

8. Model reduction via Krylov methods

An important class of numerical methods for model reduction is the Krylov-based order reduction. The aim of this class of methods is to match some of the first coefficients of the Taylor (or Neumann) series expansion of the transfer functions of the original and reduced models. When the series expansion is developed about a finite point, we speak about moment matching and Padé interpolation and when it is about $s_0 = \infty$, we are dealing Markov parameter matching and partial realization. These methods find the reduced order model in a relatively short time with a good numerical accuracy via a projection using bases of particular Krylov subspaces. The main advantage of this approach is that it requires a low computational effort and small memory storage especially when compared to other reduction approaches. However, the stability of the reduced order model cannot be guaranteed and neither an a priori nor a posteriori error bound measuring the accuracy of the approximation exists.
Approximating a dynamical system by a low order one, could be done by matching some of the Markov parameters M_i of the original system corresponding to the terms of the Laurent series of F when expanded around $s_0 = \infty$:.

$$F(s) = \sum_{i=0}^{\infty} M_i s^{-i} \in \mathbb{R}^{p \times p}.$$

This is also called the partial realization. If the dynamical system is represented by the internal description

$$\mathbf{\Sigma} \equiv \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right],$$

then one should find two matrices $\mathcal{V}_m$ and $\mathcal{W}_m$ in $\mathbb{R}^{n \times p}$ and construct the low order model

$$\mathbf{\Sigma}_m \equiv \left[\begin{array}{c|c} A_m & B_m \\ \hline C_m & 0 \end{array}\right],$$

by applying the oblique projector $\mathcal{P}_m = \mathcal{V}_m \mathcal{W}_m^T$ to the original system $\mathbf{\Sigma}$ and then

$$A_m = \mathcal{W}_m^T A \mathcal{V}_m, \; B_m = \mathcal{W}_m^T B, \; C_m = C \mathcal{V}_m.$$

The matrices $\mathcal{V}_m$ and $\mathcal{W}_m$ could be chosen by using either the extended block (or the rational block) Arnoldi process [62, 64, 98] or the block Lanczos process [16, 12, 40, 50, 63].

If we apply the extended block Arnoldi algorithm to the pair (A, B) then we get an orthonormal matrix $\mathcal{V}_m$ and in this case, the first m Markov parameters are matched

$$M_i = CA^{i-1}B = C_m A_m^{i-1} B_m,\ i = 1, \ldots, m. \tag{20}$$

Wen using the extended block Arnoldi algorithm, we also have

$$CA^{-i+1}B = C_m A_m^{-i+1} B_m,\ i = 1, \ldots, m,$$

and this means that the extended block Arnoldi allows to match the first m moments and the first m Markov parameters and this is a special property of this process.
Now, if we apply the block Lanczos process to the pair (A, B) and (A^T, C), we obtain two biorthonormal matrices $\mathcal{V}_m$ and $\mathcal{W}_m$, that is $\mathcal{W}_m^T \mathcal{V}_m = I$ and in this case the reduced order system has the property that it matches $2m$ Markov parameters

$$M_i = CA^{i-1}B = C_m A_m^{i-1} B_m,\ i = 1, \ldots, 2m.$$

The relation (20) is still valid with global Arnoldi algorithm. We notice that when using the block Arnoldi algorithm, the reduced matrix A_m is an upper block Hessenberg matrix while this matrix reduces to a block tridiagonal form when applying the block Lanczos process and to a Hessenberg form when using the global Arnoldi process.
Model reduction via Krylov-based methods guarantees moment and/or Markov parameters matching without the need of explicitly calculating these coefficients. This is an important advantage for large-scale systems for which this calculation is numerically ill-conditioned. For matching any given number of moments about a single expansion point s_0, then only one LU factorization of $(s_0 I - A)$ is required. The remaining operations are then only matrix-vector multiplications. The reduced model is directly calculated through a projection to the lower dimensional subspace.

The advantages of using Krylov methods as remarked in [9] can be stated as follows

- Krylov methods such as Arnoldi or Lanczos, need only matrix-vector operations and there is no need to use decomposition of large matrices.
- They require $\mathcal{O}(nm^2)$ operations for sparse problems and the requirement of memory is about $\mathcal{O}(mn)$
- The derivation of algorithms are more simple.

The drawbacks of these methods are

- Altghouht the original system is stable, the reduced order model may not be stable. To remedy this drawback we can apply implicit restarts techniques [9, 52]
- The obtained reduced-order model tends to approximate high frequencies. This problem can be solved by using Rational Krylov approaches [1, 14, 16, 36, 37, 38, 51].
- For some problems we have a loss of orthogonality and one need to use re-orthogonalisations.

Another way for constructing a Krylov-based reduced order model is as follows:
Set

$$F(s) = CX,\ X = (sI - A)^{-1} B,$$

which is equivalent to solve the multiple linear system of equations with the same matrix

$$(sI - A)X = B.$$

Applying for example the extended block Arnoldi process to the pairs (A, B) we get the matrix $\mathcal{V}_m$ and a un upper block Hessenberg matrix H_m. Considering the approximation $X \approx X_m = \mathcal{V}_m Y_m$, and using the Galerkin condition, we get the projected linear problem

$$\mathcal{V}_m^T (sI - A) \mathcal{V}_m Y_m = \mathcal{V}_m^T B,$$

which is equivalent to

$$(sI - H_m) Y_m = B_m.$$

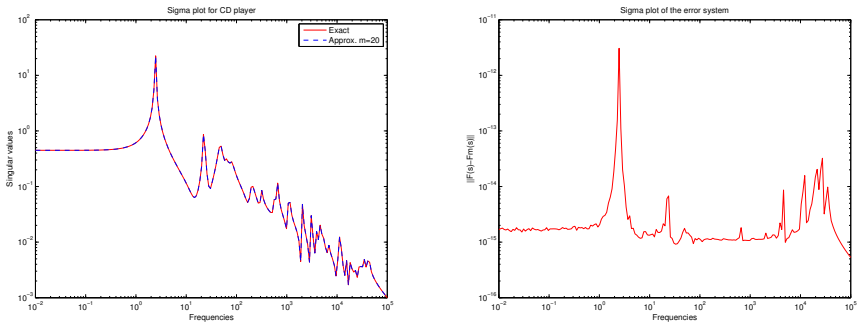

FIGURE 2. Left: The exact and approximated ($m = 20$) singular values. Right: The norm of the errors

Then we get the approximate transfer function

$$F_m(s) = CX_m = C_m(sI - H_m)^{-1}B_m,\ H_m = \mathcal{V}_m^T A\mathcal{V}_m,\ C_m = C\mathcal{V}_m.$$

If the transfer function $F(s)$ is expanded as a Taylor series around a given point s_0 (assumed to be not a pole of F), then

$$F(s) = M_0 + M_1(s - s_0) + M_2(s - s_0)^2 + ...$$

where the matrix coefficients M_j for $j \geq 0$ are known as moments of the system around s_0, it can be shown that

$$M_j(s_0) = -C(A - s_0 I_n)^{-(j+1)}B,\ j = 0, 1, \ldots.$$

These moments are the value of the transfer function of the system (1) and its derivatives evaluated at s_0

$$M_i(s_0) = \frac{d^i}{ds^i}F(s)_{s=s_0}.$$

The model-order reduction using a moment matching method consists in finding a lower order transfer function $F_m(s)$ having a power series expansion at s_0 as follows

$$F_m(s) = \widehat{M}_0 + \widehat{M}_1(s - s_0) + \widehat{M}_2(s - s_0)^2 + ...,$$

such that a certain number of moments are matched.

If $s_0 = 0$, the problem is known as a Padé approximation [9, 29, 53, 106] and in this case we can use Krylov-based methods such as the extended block Arnoldi to the pair $(A^{-1}, A^{-1}B)$ to get a low order model that approximates well $F(s)$ for small frequencies. We can also generate a reduced-order model using a multipoint Padé approximation or a multipoint rational interpolation [1, 16, 36, 37, 42, 44, 45, 51, 81] where the reduced system matches the moments of the original system at multiple interpolation points. For small and large frequencies, the extended block Arnoldi process gives good results as shown in the next numerical test.

We consider the well known CD-player model. This model is usually used for testing model-order reduction methods. It describes the dynamics between the lens actuator and the radial arm position of a CD player and its dimension is $n = 120$ and $p = 2$. The left side of Figure 2 represents the singular values of the exact and the approximated transfer functions while in the right of this figure, we plotted the obtained error $\|F(s) - F_m(s)\|_2$ with $m = 20$.

9. Model-reduction via matrix algebraic matrix Riccati equations

9.1. **The LQG-Riccati method for model reduction.** In this subsection, we assume that the original system is no longer stable and present a reduced order model method called the Linear Quadratic Gaussian (LQG) balanced truncation method; see [23, 39, 87]. Assume that the dynamical system (1) is minimal, then the basic idea of (LQG) balanced truncation is to replace the Lyapunov Gramians P and Q used for the classical balanced truncation for stable

systems by the stabilizing solutions of the dual Kalman Filtering Riccati Equation (FARE) and Control Algebraic Riccati Equation (CARE) defined as follows

$$AP + PA^T - PC^TCP + BB^T = 0, \quad (FARE)$$

and

$$A^TQ + QA - QB^TBQ + C^TC = 0. \quad (CARE)$$

Let P_+ and Q_+ be the stabilizing (and positive semidefinite) solutions of the matrix equations (FARE) and (CARE), respectively which means that the eigenvalues of the closed loops $A - P_+C^TC$ and $A^T - Q_+BB^T$ lie in the open left half plane $\mathbb{C}^-$. It is known [39] that, as for the classical balanced truncation, the eigenvalues of the product P_+Q_+ are invariant quantities under any state coordinate transformation $\tilde{x}(t) = Tx(t)$ where T is a nonsingular $n \times n$ matrix and we have

$$\widetilde{P}_+\widetilde{Q}_+ = TP_+Q_+T^{-1},$$

where $\widetilde{P}_+$ and $\widetilde{Q}_+$ correspond to the Gramians of the new states. The solutions P_+ and Q_+ have usually low numerical rank and can then be approximated by low rank factorizations $P_+ \approx \mathcal{Z}_m\mathcal{Z}_m^T$ and $Q_+ \approx \mathcal{Y}_m\mathcal{Y}_m^T$ where the matrix factors $\mathcal{Y}_m$ and $\mathcal{Z}_m$ have low ranks. As in the classical balanced truncation method, these factor could be used to contsruct the (LQG)-balanced truncation reduced order model.

9.2. **Positive-real balanced truncation.** An important class of LTI dynamical systems is that for which the tranfer function $F(s)$ is positive real (or passive) [9, 14], i.e.

- F has no pole in $\mathbb{C}^+$.
- $Re(F(s)) \geq 0$ for all $s \in \mathbb{C}^+$.

If we want to conserve the passivity of the reduced order model, the corresponding transfer function F_m should be positive real. This could be done by truncating a positive real balanced realization of the system. This balanced realization of the LTI system $\mathbf{\Sigma}$ is obtained by computing the two minimal positive semi-definite Gramians $\mathcal{P}$ and $\mathcal{Q}$ of the following dual CAREs

$$\begin{aligned}(A - BC)\mathcal{P} + \mathcal{P}(A - BC)^T + \mathcal{P}C^TC\mathcal{P} + BB^T &= 0,\\ (A - BC)^T\mathcal{Q} + \mathcal{Q}(A - BC) + \mathcal{Q}BB^T\mathcal{Q} + C^TC &= 0.\end{aligned}$$

For more details, see [55, 88]. Therefore, the Gramians $\mathcal{P}$ and $\mathcal{Q}$ could be computed in a factored form in a similar way as we explained for LQG problem and then those factors are used to construct a new reduced order positive-real balanced truncated model $\mathbf{\Sigma_m}$ where m is much smaller than the size n of the original system. It shown in [9, 23, 55] that in this case $\mathbf{\Sigma_m}$ is asymptotically stable positive real and we have the following upper bound

$$\|F(s) - F_m(s)\|_{\mathcal{H}_\infty} \leq 2\|I + F(s)\|_{\mathcal{H}_\infty}\|I + F_m(s)\|_{\mathcal{H}_\infty} \sum_{i=m+1}^{n} \tilde{\sigma}_i,$$

where $\tilde{\sigma}_i = \sqrt{\lambda_i(\mathcal{P}\mathcal{Q})}$ and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$.
Next, we give some Krylov based methods for computing approximate solution of large-scale algebraic Riccati equations.

9.3. **Numerical methods for large-scale algebraic matrix Riccati equations.**

9.3.1. *Krylov-based methods.* Consider the continuous-time algebraic Riccati equation

$$A^T X + X A - X B B^T X + C^T C = 0 \quad (CARE), \tag{21}$$

where $A \in \mathbb{R}^{n\times n}$ is nonsingular, $B \in \mathbb{R}^{n\times p}$ and $C \in \mathbb{R}^{p\times n}$. The matrices B and C are assumed to be of full rank with $p \ll n$. Matrix algebraic Riccati equations play a fundamental role in many areas such as control, filter design theory, model reduction problems, differential equations and robust control problems [10, 24, 77, 85]. For historical developments, applications and

importance of algebraic Riccati equations, we refer to [3, 24, 35] and the references therein.
The desired solution X is symmetric positive semidefinite and stabilizing (satisfies the fact that the eigenvalues of the resulting closed-loop matrix $A - B\,B^T\,X$ are in the open complex left-half plane). This stabilizing solution exists and is unique under certain assumptions on the problem [77].
Let $x(t)$ be the state vector of dimension n, $u(t)$ the control vector of $\mathbb{R}^p$ and $y(t)$ the output vector of length p. We consider the following Linear Quadratic Regulator (LQR) problem:

Minimize

$$J(x_0, u) = \frac{1}{2}\int_0^{+\infty} \left(y(t)^T\,y(t) + u(t)^T\,u(t)\right)\,dt,$$

under the dynamic constrains

$$\left\{\begin{array}{rcl} \dot{x}(t) & = & A\,x(t) + B\,u(t) \text{ with } x(0) = x_0. \\ y(t) & = & C\,x(t). \end{array}\right.$$

Under the hypotheses : the pair (A, B) is stabilizable (i.e., $\exists$ a matrix S such that $A - B\,S$ is stable) and the pair (C, A) is detectable (i.e., (A^T, C^T) stabilizable), a unique optimal solution $\bar{u}$ that minimize the functional $J(x_0, u)$ exists [108], and can be determined through a feedback operator K such that $\bar{u}(t) = K\,x(t)$, where $K = -B^T\,X$ and $X \in \mathbb{R}^{n\times n}$ is the unique symmetric positive semidefinite and stabilizing solution of the matrix equation (21).
We notice that the unique stabilizing solution of (21) can be obtained by considering the n dimensional stable invariant subspace $\mathbb{S}_{\mathcal{H}}$ (i.e., the subspace corresponding to the eigenvalues of the Hamiltonian matrix $\mathcal{H}$ (see [84, 85])

$$\mathcal{H} = \left(\begin{array}{cc} A & B\,B^T \\ C^T\,C & -A^T \end{array}\right) \in \mathbb{R}^{2n\times 2n}.$$

If $\mathbb{S}_{\mathcal{H}}$ is spanned by the columns of $\left[X_1^T, X_2^T\right]^T \in \mathbb{R}^{2n\times n}$ and X_1 is nonsingular, then $X = -X_2\,X_1^{-1}$ is the stabilizing solution of (21).
During the last decades, many numerical methods for solving the CARE (21) with small and dense matrices have been developed. The standard computational methods are based on the Schur and structure-preserving Schur methods [84, 106, 32, 85], matrix Sign function methods [13, 96], Newton-type methods [11, 22, 57, 74, 77] and the Symplectic Lanczos method [22]. Other methods are also found in [5, 6]. Generally, the matrices A, B and C are obtained from the discretization of operators defined on infinite dimensional subspaces. Moreover, the matrix A is in general sparse, banded and very large and such problems few attempts have been made to solve (21). Among of them is theLow Rank Cholesky Factorized Newton (LRCF-Newton) method [21]. At each step of the outer Newton iteration, one needs to solve a large Lyapunov matrix equation. In the LRCF-Newton method, these Lyapunov matrix equations are solved by the Low Rank Cholesky Factorized method [92] which is based on the solution of linear systems with shifted matrices $A - \mu_i\,I$ where the μ_i's are the ADI parameters.
Another point of view is to consider Krylov-based methods described as follows. As for Lyapunov or Sylvester matrix equations, and since the right-hand side of CARE (21) has a low rank, we can consider low-rank approximate solutions given by

$$X_m = \mathcal{V}_m Y_m \mathcal{V}_m^T,$$

where the matrix $\mathcal{V}_m = [V_1, \ldots, V_m]$ of size $n \times 2mp$ is obtained by applying the extended block Arnoldi algorithm to the pair (A^T, C^T). Therefore, applying the Galerkin condition to the residual $\mathcal{R}(X_m) = A^T\,X_m + X_m\,A - X_m\,B\,B^T\,X_m + C^T\,C$, we get

$$\mathcal{V}_m^T \mathcal{R}(X_m) \mathcal{V}_m = 0.$$

Therefore, we obtain the low order matrix Riccati equation

$$\mathcal{T}_m Y_m + \mathcal{T}_m^T Y_m - Y_m \widetilde{B}_m \widetilde{B}_m^T Y_m + \widetilde{C}_m^T \widetilde{C}_m = 0, \tag{22}$$

where $\mathcal{T}_m = \mathcal{V}_m^T A^T \mathcal{V}_m$, $\widetilde{B}_m = \mathcal{V}_m^T B$ and $\widetilde{C}_m^T = \mathcal{V}_m^T C^T$. We assume that the projected algebraic Riccati equation (22) has a unique symmetric positive semidefinite and stabilizing solution Y_m. This solution can be obtained by a standard direct method such as the Schur method [84]. As m increases, the computation of X_m becomes expensive and so, in order to stop the iterations ($\| \mathcal{R}(X_m) \| < \epsilon$) without having to compute extra products involving the matrix A, we can use the following result

Theorem 9.3.1. [60] *Let $X_m = \mathcal{V}_m Y_m \mathcal{V}_m^T$ be the approximation obtained at step m by the extended block Arnoldi-CARE method and let Y_m be the symmetric positive semi-definite stabilizing solution of the low-dimensional CARE (22), then we have*

$$\| \mathcal{R}(X_m) \|_2 = \| T_{m+1,m} \tilde{Y}_m \|_2,$$

where $\tilde{Y}_m$ is the $2p \times 2mp$ matrix corresponding to the last $2p$ rows of Y_m.

We notice that the computed solution X_m could be given as a product of two matrices of low rank. Consider the singular value decomposition of the $2m \times 2m$ matrix $Y_m = U \, \Sigma \, U^T$ where Σ is the diagonal matrix of the singular values of Y_m sorted in decreasing order. Let U_l be the $2m \times l$ matrix of the first l columns of U corresponding to the l singular values of magnitude greater than some tolerance *dtol*. We obtain the truncated singular value decomposition $Y_m \approx U_l \, \Sigma_l \, U_l^T$ where $\Sigma_l = \mathrm{diag}[\sigma_1, \ldots, \sigma_l]$. Setting $Z_m = \mathcal{V}_m \, U_l \, \Sigma_l^{1/2}$, it follows that

$$X_m \approx Z_m \, Z_m^T.$$

Next, we give an upper bound of the norm of the error $X - X_m$ where X is the stabilizing solution of the CARE (21).

Theorem 9.3.2. [60] *Let X_m, Y_m be the m-th approximate solution obtained with the extended block Arnoldi-CARE algorithm and the solution of the projected problem (22) respectively. Let $\tilde{Y}_m$ be the $2p \times 2mp$ matrix corresponding to the last $2p$ rows of Y_m. We set $\gamma_m = \| T_{m+1,m} \tilde{Y}_m \|$, $\eta = \| BB^T \|$, $A_m = A - B \, B^T X_m$ and assume that $\delta_m = \mathrm{sep}(A_m, -A_m^T) > 0$. Then if $\frac{4\gamma_m \, \eta}{\delta_m^2} < 1$, we have*

$$\| X - X_m \| \leq \frac{2\gamma_m}{\delta_m + \sqrt{\delta_m^2 - 4\gamma_m \eta}}.$$

From the numerical point of view, Krylov-based methods require a relatively low computational effort and memory storage and are very effective for large problems. The unique drawback is the fact that there is no proof (under which conditions) that the approximate solutions are stabilizing. However this was true for all the numerical numerical experiments we tested. Other Krylov based methods for solving the CARE (21) could be found in [64, 69, 70, 80, 103].

9.3.2. *Newton-based methods for large matrix Riccati equations.* Let $\mathcal{R}(X)$ be the residual corresponding to the matrix Riccati equation (21):

$$\mathcal{R}(X) = A^T X + XA - XBB^T X + C^T C.$$

The Fréchet derivative of $\mathcal{R}$ at X_k is given by

$$\forall Z, \quad \mathcal{R}'_{X_k}(Z) = (A - BB^T X_k)^T \, Z + Z \, (A - BB^T X_k)$$

The Kleinman-Newton method is defined from X_0 as follows: At step $k+1$, we compute the new iterate X_{k+1} by solving the linear matrix equation

$$(\mathcal{R}'_{X_k})(X_{k+1} - X_k) = -(\mathcal{R}(X_k)).$$

The approximation X_{k+1} is obtained by solving the Lyapunov matrix equation:

$$F_k^T\, X_{k+1} + X_{k+1}\, F_k + W_k W_k^T = 0,$$

with $F_k = A - B\, K_k$, $K_k = B^T\, X_k$ and $W_k = [K_k^T\ C^T]$.

Algorithm 5 The Kleinman-Newton method for CAREs

(1) Choose an initial guess $X_0 \in \mathbb{R}^{n\times n}$ such that $A - BB^T X_0$ is stable and set $K_0 = B^T X_0$.
(2) For k=0,. . . ,itermax
  (a) Solve, for X_{k+1}, the Lyapunov matrix equation

$$F_k^T X_{k+1} + X_{k+1} F_k + W_k W_k^T = 0, \tag{23}$$

  (b) Compute the feedback $K_{k+1} = B^T X_{k+1}$.

We notice here that the Kleinman-Newton formulation summarized in Algorithm 5 is very suitable for large problems since the derived Lyapunov equations (23) have low rank. We have the following well known theorem.

Theorem 9.3.3. [77] *Let X_+ be the stabilizing solution and assume that X_0 is stabilizing, then for the symmetric iterates X_k generated by Algorithm 5, we have*

(1) *X_k is stabilizing for all $k = 0, 1, \ldots$.*
(2) *$X_1 \geq X_2 \geq \ldots \geq X_k \geq \cdots \geq X_+$.*
(3) *$\lim_{k\to\infty} (X_k) = X_+$.*
(4) *There exists a positive constant c such that*

$$\|X_{k+1} - X_+\| \leq c\|X_k - X_+\|^2,\ k = 1, 2, \ldots$$

where $\|\ \|$ is any given matrix norm.

We notice here that the computed approximation X_{k+1} should not be stored explicitly but could be given as a product of two low rank matrices $X_{k+1} = ZZ^T$, so that only Z is used (see [19, 25] for more details). The Lyapunov equations appearing in Algorithm 5 could be solved approximatively by using a Krylov based or ADI-based methods with a high accuracy to guarantee that the Newton iteration converges to the stabilizing solution [21]. As these Lyapunov matrix equations are not solved exactly, one should consider an inexact scheme [43]. This consists in approximating the solution of the Lyapunov equation (23) in the Newton-Kleinman step to a certain tolerance. In the inexact-Newton method on solves, at each step, the problem

$$(\mathcal{R}'_{X_k})(X_{k+1} - X_k) = -(\mathcal{R}(X_k)) - R_{k+1}$$

where R_{k+1} corresponds to the residual associated with the iterative method used for solving the derived Lyapunov equation. We have the following result

Theorem 9.3.4. [75] *Let X_+ be a the stabilizing solution of (21), then there exist $\eta > 0$ and $\delta > 0$ such that for an initial guess X_0 with $\|X_+ - X_0\| \leq \eta$, the iterates produced by the inexact Newton-Kleinman algorithm converge to X_+ if*

$$\|R_{k+1}\| \leq \delta_k \|\mathcal{R}(X_k)\|,$$

where δ_k are positive forcing terms, with $\delta_k \in]0, \delta]$. Furthermore,

(1) *The convergence is linear.*
(2) *The convergence is superlinear if $\lim_{k\to\infty} (\delta_k) = 0$.*
(3) *If $\delta_k \leq M_\delta \|\mathcal{F}(X_k)\|$ for some $M_\delta > 0$, then the convergence is quadratic.*

In [43], a modification of Theorem 9.3.4 is given. Inexact Newton methods for discrete-time matrix Riccati equations have been developed in [27]. Perturbation results and error bounds are given in [105]. Other Newton-based methods are presented in [20, 21, 23, 25, 32, 58].

10. Conclusion

The applications of Krylov subspace based methods and ADI-based methods to model order reduction and control problems have received a great attention these last years. The success of these methods is the fact that they are effective for large problems which is the case in many applications. This work presents the most well known methods for model reduction in large scale MIMO dynamical systems and some other related control problems. We presented two classes of model reduced order techniques for large problems. The first one is called the balanced truncation and is related to the solution of some Gramians obtained by solving large Lyapunov, Stein or matrix algebraic Riccati equations. These matrix equations are solved by Krylov-based methods or by ADI-type iterations. The second class of model order reduction is called moment matching methods and is obtained by deriving projectors (orthogonal or obliques) from some Krylov (SISO) or block Krylov (MIMO) subspaces. The advantage of the latter methods is that the operation requirement need only matrix-vector products and the stockage of average $O(n^2)$ where n is the dimension of the dynamical system.

References

[1] O. Abidi, M. Hached and K. Jbilou, Adaptive rational block Arnoldi methods for model reductions in large-scale MIMO dynamical systems, New Trends in Mathematical, V.4, N.2, 2016, pp.227–239.

[2] O. Abidi and K. Jbilou, On some properties of the extended block and global Arnoldi methods with applications to model reduction, Numerical Algorithms, to appear, 2016.

[3] H. Abou-Kandil, G. Freiling, V. Ionescu, G. Jank, Matrix Riccati Equations in Control and Sytems Theory, in Systems & Control Foundations & Applications, Birkhauser, 2003, 575p.

[4] S. Agoujil, A. Bentbib, K. Jbilou and M. Sadek, A minimization method for large sylvester matrix problems, Electronic Transaction Numerical Analysis, 43, 2014, pp.45–59.

[5] F.A. Aliev, B.A. Bordyug and V.B. Larin, Comments on "On computing the stabilizing solution of a class of discrete-time periodic Riccati equations" by Vasile Dragan, Samir Aberkane and Ivan G. Ivanov, Appl. Comput. Math., V.13, N.2, 2014, pp.266–267.

[6] F.A. Aliev, B.A. Bordyug and V.B. Larin, Comments on "A stability-enhancing scaling procedure for Schur-Riccati solvers", Systems & Control Letters, V.14, N.5, 1990, pp.453–453.

[7] F.A. Aliev and V.B. Larin, Construction of square root factor for solution of the Lyapunov matrix equation, Systems & Control Letters, V.20, N.2, 1993, pp.109–112.

[8] F.A. Aliev and V.B. Larin, Generalized Lyapunov equation and factorization of matrix polynomials, Systems & Control Letters, V.21, N.6, 1993, pp.485–491.

[9] A. C. Antoulas, Approximation of Large-Scale Dynamical Systems, Advances in Design and Control, SIAM, 2005, 481p.

[10] B.D.O. Anderson, J.B. Moore, Linear Optimal Control, Prentice-Hall, Englewood Cliffs, NJ, 1971, 450p.

[11] W.F. Arnold III, A.J. Laub, Generalized eigenproblem algorithms and software for algebraic Riccati equations, Pro., V.72, 1984, pp.1746–1754.

[12] Z. Bai, D. Day, Q. Ye, ABLE: an Adaptive Block Lanczos Method for Non-Hermitian Eigenvalue Problems, SIAM Journal on Matrix Analysis and Applications, V.20, 1999, pp.1060–1082.

[13] Z. Bai, J. Demmel, Using the matrix sign function to compute invariant subspaces, SIAM Journal on Matrix Analysis and Applications, V.19, 1998, pp.205–225.

[14] Z. Bai and R. Freund, A Partial Padé-via-Lanczos Method for Reduced-Order Modeling, Linear Algebra and Applications, V.332-334, 2001, pp.139-164.

[15] G.A. Baker Jr., P.R. Graves-Morris, Pade Approximants, 2nd Edition, Cambridge University Press, Cambridge, 1996, 746p.

[16] H. Barkouki, A. H. Bentbib and K. Jbilou, An adaptive rational block Lanczos-type algorithm for model reduction of large scale dynamical systems, Journal of Scientific Computing, V.67, 2016, pp.221–236.

[17] H. Barkouki, A. H. Bentbib and K. Jbilou, A matrix rational Lanczos method for model reduction, Numerical Linear Algebra with Applications, 2016, to appear.

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

[19] P. Benner, R.C. Li, N. Truhar, On the ADI method for Sylvester equations, J. Comput. Appl. Math., V.233, 2009, pp.1035–1045.
[20] P. Benner, R. Byers, Ralph Byers, E. S. Quintana-Orti and G. Quintana-Orti, Solving Algebraic Riccati Equations on Parallel Computers Using Newton's Method with Exact Line Search, Parallel Comput., V.26, N.10, 2000, pp.1345–1368.
[21] P. Benner, J. Li, and T. Penzl Numerical solution of large Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems, V.15, N.9, 2008, pp.755–777.
[22] P. Benner, H. Faβbender, An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem, Linear Algebra and its Applications, V.263, 1997, pp.75–111.
[23] P. Benner and R. Byers, An exact line search method for solving generalized continuous-time algebraic Riccati equations, IEEE Transaction on Automatic Control, V.43, 1998, pp.101–107.
[24] S. Bittanti, A. Laub, J.C. Willems, The Riccati equation, Springer-Verlag, Berlin, 1991, 338p.
[25] A. Bouhamidi M. Hached and K. Jbilou , A preconditioned block Arnoldi method for large scale Lyapunov and algebraic Riccati equations, Journal of Global Optimtimisation, V.65, N.1, 2016, pp.19–32.
[26] A. Bouhamidi M. Hached, M. Heyouni and K. Jbilou , A preconditioned block Arnoldi for large Sylvester matrix equations, Numerical Linear Algebra with Applications V.2, 2013, pp.208–219.
[27] A. Bouhamidi and K. Jbilou, On the convergence of inexact Newton methods for discrete-time algebraic Riccati equations , Linear Algebra and Applications, V.439, N.7, 2013, pp.2057–2069.
[28] R. Bouyouli, K. jbilou, R. Sadaka and H. Sadok, Convergence properties of some block Krylov methods, Journal of Computational and Applied Mathematics, V.196, 2006, pp.498–511.
[29] C. Brezinski, Computational Aspects of Linear Control, Kluwer, Dordrecht, 2002, 295p.
[30] C. Brezinski, Pade-Type Approximation and General Orthogonal Polynomials, ISNM Vol. 50, Birkhauser-Verlag, Basel, 1980, 249p.
[31] Bultheel and B. De Moore, Rational approximation in linear systems and control, Journal of Computational and Applied Mathematics, V.121, 2000, pp.355–378.
[32] R. Byers, Solving the algebraic Riccati equation with the matrix sign function, Linear Algebra and its Applications, V.85, 1987, pp.267–279.
[33] D. Calvetti, L. Reichel, Application of ADI iterative methods to the restoration of noisy images, SIAM Journal on Matrix Analysis and Applications, V.17, 1996, pp.165–186.
[34] 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.
[35] B.N. Datta, Numerical Methods for Linear Control Systems Design and Analysis, Elsevier Academic Press, 2003, 696p.
[36] V. Druskin, C. Lieberman, M. Zaslavsky, On adaptive choice of shifts in rational Krylov subspace reduction of evolutionary problems, SIAM Journal od Scientific Computing, V.32, N.5, 2010, pp.2485–2496.
[37] V. Druskin and V. Simoncini,Adaptive rational Krylov subspaces for large-scale dynamical systems, Systems & Control Letters, V.60, 2011, pp.546–560.
[38] V. Druskin, L. Knizhnerman, and V. Simoncini, Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation, SIAM Journal on Numerical Analysis, V.49, 2011, pp.1875–1898.
[39] L. Fortuna, G. Nunnari and A. Gallo , Model Order Reduction Techniques with Applications in Electrical Engineering, Springer-Verlag London, 1992, 232p.
[40] Freund, Roland W, Krylov-subspace methods for reduced-order modeling in circuit simulation, Journal of Computational and Applied Mathematics, V.123, 2000, pp.395–421.
[41] A. El Guennouni, K. Jbilou and A.J. Riquet, Block Krylov subspace methods for solving large Sylvester equations, Numerical Algorithms , V.29, 2002, pp.75–96.
[42] M. Frangos, I. M. Jaimoukha, Adaptive rational interpolation: Arnoldi and Lanczos-like equations, European Journal of Control, V.14, N.4, 2008, pp.342–354.
[43] F. Feitzinger, T. Hylla, E. W. Sachs, Inexact Kleinman Newton method for Riccati equations, SIAM Journal on Matrix Analysis and Applications, V.3, N.2, 2009, pp.272–288.
[44] K. Gallivan, E. Grimme and P. Van Dooren, A rational Lanczos algorithm for model reduction, Numerical Algorithms, V.12, 1996, pp.33-63.
[45] K. Gallivan, E. Grimme and P. Van Dooren, Padé Approximation of Large-Scale Dynamic Systems with Lanczos Methods, Proceedings of the 33rd IEEE Conference on Decision and Control, IEEE Publications, Piscataway, NJ, 1994, pp.443-448.
[46] W. Gawronski and J.-N. Juang, Model reduction in limited time and frequency intervals, International Journal of Systems Sciences, V.21, N.2, 1990, pp.349–376.
[47] K. Glover, All optimal Hankel-norm approximations of linear multivariable systems and their L-infinity error bounds, International Journal of Control, V.39, 1984, pp.1115–1193.

[48] 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 Journal on Control Optimisation, V.29, 1991, pp.283–324.
[49] G. H. Golub, S. Nash and C. Van Loan, A Hessenberg-Schur method for the problem $A\,X + X\,B = C$, IEEE Transaction on Automatic Control, AC-24, 1979, pp.909–913.
[50] G. H. Golub, R. Underwood, The block Lanczos method for computing eigenvalues, Mathematical Software III, J. R. Rice, ed., Academic Press, New York, pp.361–377, 1977.
[51] E. Grimme, Krylov Projection Methods for Model Reduction, Ph.D. Thesis, ECE Dept., University of Illinois, Urbana-Champaign, 1997, 223p.
[52] E. Grimme, K. Gallivan, A rational lanczos algorithm for model reduction ii: Interpolation point selection, Numerical Algorithms., V.12, 1998, pp.33–63.
[53] E. J. Grimme, D. C. Sorensen and P. Van Dooren, Model reduction of state space systems via an implicitly restarted Lanczos method, Numerical Algorithms, V.12, 1996, pp.1–32.
[54] S. Gugercin, D.C. Sorensen, and A.C. Antoulas, A modified low-rank Smith method for large-scale Lyapunov equations, Numerical Analysis, V.32, N.1, 2003, pp.27–55.
[55] S. Gugercin and A.C. Antoulas, A survey of model reduction by balaned trun- cation and some new results, International Journal of Control, V.77, N.8, 2004, pp.748–766.
[56] S. Gugercin and A.C. Antoulas and C. Beattie, H_2 model reduction for large-scale dynamical systems, SIAM Journal on Matrix Analysis and Applications, V.30, N.2, 2008, pp.609–638.
[57] C.H. Guo, P. Lancaster, Analysis and modification of Newton's method for algebraic Riccati equations, Mathematics on Computation, V.67, 1998, pp.1089–1105.
[58] C.H. Guo and A.J. Laub, On a Newton-Like Method for Solving Algebraic Riccati Equations, SIAM Journal on Matrix Analysis and Applications, V.21, N.2, 2000, pp.694–698.
[59] S. J. Hammarling, Numerical solution of the stable, nonnegative definite Lyapunov equation, IMA Journal on Numerical Analysis, V.2, 1982, pp.303–323.
[60] M. Heyouni and K. Jbilou, An extended block Arnoldi algorithm for large-scale solutions of the continuous-time algebraic Riccati equation, Electronic Transaction on Numerical Analysis, V.33, 2009, pp.53–62.
[61] M. Heyouni, Extended Arnoldi methods for large low-rank Sylvester matrix equations, Applied Numerical Mathematics, V.60, N.11, 2010, pp.1171–1182.
[62] M. Heyouni and K. Jbilou, Matrix Krylov subspace methods for large scale model reduction problems, Applied Mathematics and Computation, V.181, 2006, pp.1215-1228.
[63] M. Heyouni and K. Jbilou, A. Messaoudi and K. Tabaa, Model reduction in large scale MIMO dynamical systems via the block Lanczos method, Computational Applied Mathematics, V.27, 2008, pp.211–236.
[64] I.M. Jaimoukha and E.M. Kasenally, Krylov subspace methods for solving large Lyapunov equations, SIAM Journal on Matrix Analysis with Applications, V.31, N.1, 1994, pp.227–251.
[65] K. Jbilou, A. Messaoudi and H. Sadok, Global FOM and GMRES algorithms for matrix equations, Applied Numerical Mathematics, V.31, 1999, pp.49–63.
[66] K. Jbilou, A. J. Riquet, Projection methods for large Lyapunov matrix equations, Linear Algebra and its Applications, V.415, N.2, 2006, pp.344–358.
[67] K. Jbilou, ADI preconditioned Krylov methods for large Lyapunov matrix equations, Linear Algebra and its Applications, V.432, N.10, 2010, pp.2473–2485.
[68] K. Jbilou, Low rank approximate solutions to large Sylvester matrix equations Applied Mathematics and Computation, V.177, 2006, pp.365–376.
[69] K. Jbilou, Block Krylov subspace methods for large continuous-time algebraic Riccati equations, Numerical Algorithms, V.34, 2003, pp.339–353.
[70] K. Jbilou, An Arnoldi based method for large algebraic Riccati equation, Applied Mathematic Letters, V.19, 2006, pp.437–444.
[71] K. Jbilou and A. Messaoudi, A computational method for symmetric Stein matrix equations, In Numerical Linear Algebra in Signals, Systems and Control, V.80, 2011, pp.295–311.
[72] R E . Kalman, Contributions to the theory of optimal control, Bol. Soc. Mat. Mex., V.5, 1960, pp.102–119.
[73] R.E. Kalman, On the general theory of control systems, in Proc.I st International Congress on Automatic Control, Moscow, 1960; Butterworths, London, V.1, 1961, pp.481–492.
[74] D.L. Kleinman, On an iterative technique for Riccati equation computations, IEEC Transaction on Automatic Control, V.13, 1968, pp.114–115.
[75] C. Kelly, Iterative methods for linear and nonlinear equations, SIAM, Philadelphia, 1995, 263p.
[76] P. Lancaster and M. Tismenetsky, The Theory of Matrices, Academic press, London, 1985, 563p.
[77] P. Lancaster and L. Rodman, Algebraic Riccati equations, Oxford Univ. Press, 1995, 480p.
[78] C.-A Lin and T.-Y Chiu, Model reduction via frequency weighted balanced realization, Control Theory and Advanced Technol., V.8, 1992, pp.341–351.

[79] Y. Lin and V. Simoncini, Minimal residual methods for large scale Lyapunov equations, Applied Numerical Mathematics, V.72, 2013, pp.52–71.
[80] Y. Lin and V. Simoncini, A new subspace iteration method for the algebraic Riccati equation, Numerical Linear Algebra with Applications, V.22, N.1, 2015, pp.26–47.
[81] 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 and its Applications, V.415, 2006, pp.235-261.
[82] J.R. Li and J. White, Low rank solution of Lyapunov equations, SIAM Journal on Matrix Analysis and Applications, V.24, 2002, pp.26--280.
[83] A. Lu and E. Wachspress, Solution of Lyapunov equations by alternating direction implicit iteration, Computational and Applied Mathematics, V.21, 1991, pp.43–58.
[84] A.J. Laub, A Schur method for solving algebraic Riccati equations, IEEE Transaction on Automatic Control, Vol. AC-24, 1979, pp.913–921.
[85] V. Mehrmann, The autonomous Linear Quadratic Control Problem, theory and Numerical Solution, Number in Lecture Notes in Control and Information Sciences, Springer-Verlag, Heidelberg, 1991, 174p.
[86] B.C. Moore, Principal component analysis in linear systems: controllability, observability and model reduction, IEEE Transaction on Automatic Control, AC-26, 1981, pp.17–32.
[87] C. T. Mullis and R. A. Roberts, Synthesis of minimum roundoff noise fixed-point digital filters, IEEE Transaction on Circuits Systems, CAS-23, 1976, pp.551–562.
[88] R. Ober, Balanced parametrization of classes of linear systems, SIAM Journal Control Optimisation, V.29, 1991, pp.1251–1287.
[89] V. Papakos and I.M. Jaimoukha, A deflated implicitly restarted Lanczos algorithm for model reduction, In Proc. IEEE Conference on Decision and Control, V.3, 2003, pp.2902– 2907.
[90] D.W. Peaceman and H.H. Rachford, The numerical solution of parabolic and elliptic differential equations, Journal of the Society for Industrial and Applied Mathematics, V.3, 1955, pp.28–41.
[91] T. Penzl, A cyclic low-rank Smith method for large Lyapunov equations, SIAM Journal on Scientific Computing, V.21, N.4, 2000, pp.1401–1418.
[92] T. Penzl, Eigenvalue decay bounds for solutions of Lyapunov equations: The symmetric case, System & Control Letters, V.40, 2000, pp.139–144.
[93] T. Penzl, *LYAPACK: A MATLAB toolbox for large Lyapunov and Riccati Equations, Model Reduction Problems, and Linear-quadratic Optimal Control Problems*, http://www.tu-chemintz.de/sfb393/lyapack.
[94] V. Popov, Hyperstability and optimality of automatic systems with several control functions, Rev. Roum. Sci. Tech., V.9, 1964, pp.629–690.
[95] N. Radojicic, M. Maric and S. Bozovic, Two efficient hybrid metaheuristic methods for solving the load balance problem, Applied Computational Mathematics, V.13, N.3, 2014, pp.332–349.
[96] D.J. Roberts Linear model reduction and solution of the algebraic Riccati equation by use of the sign function, International Journal on Control, V.32, 1998, pp.677–687.
[97] A. Ruhe, Rational Krylov sequence methods for eigenvalue computation, Linear Algebra and its Applications, V.58, 1984, pp.391–405.
[98] Y. Saad, Iteratives methods for sparse lineair system, PWS press, New york, 1996, 557p.
[99] 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., Birk., Boston, pp. 503–511, 1990.
[100] Y. Shamash, Stable reduced-order models using Padé type approximations, IEEE Transaction on Automatic Control, AC-19, 1974, pp.615–616,.
[101] V. Simoncini, A new iterative method for solving large-scale Lyapunov matrix equations, SIAM Journal on Scientific Computing, V.29, N.3, 2007, pp.1268–1288.
[102] V. Simoncini, On the numerical solution of AX - XB = C, BIT, V.36, N.4, 1996, pp.814–830.
[103] V. Simoncini, Daniel B. Szyld and Marlliny Monsalve, On two numerical methods for the solution of large-scale algebraic Riccati equations, IMA Journal of Numerical Analysis, V.34, N.3, 2014, pp.904–920.
[104] V. Simoncini, Computational methods for linear matrix equations, to appear in SIAM Review.
[105] J.-G. Sun, Residual bounds of approximate solutions of the algebraic Riccati equation, Numerish Mathematics, V.76, 1997, pp.249–263.
[106] P. Van Dooren, The generalized eigenstructure problem in linear system theory , IEEE Transactions on Automatic Control, AC-26, 1981, pp.111–129.
[107] E.L. Wachspress, Extended application of alternating direction implicit iteration model problem theory, J. Soc. Indust. Appl. Math., V.11, 1963, pp.994–1016.
[108] W.M. Wonham, On a matrix Riccati equation of stochastic control, SIAM Journal on Control, V.6, 1968, pp.681–697.

[109] K. Zhou, Frequency-weighted $\mathcal{L}_\infty$ norm and optimal Hankel norm model reduction, IEEE Transaction on Automatic Control, V.40, N.10, 1995, pp.1687-1699.
[110] K. Zhou, G, Salomon and E. Wu, Balanced realization and model reduction for unstable systems, International Journal on Robust Nonlinear Control, V.9, 1999, pp.183-198.