ITERATIVE MODEL REDUCTION
OF LARGE STATE-SPACE SYSTEMS

Marko HUHTANEN*

There exist criteria for reducing the order of a large state-space model based on the accuracy of the approximate solutions to the Lyapunov matrix equations and the Hankel operator. Iterative solution techniques for the Lyapunov equations with the Arnoldi method have been proposed in a number of papers. In this paper we derive error bounds for approximations to the solutions to the Lyapunov equations as well as for the Hankel operator that indicate how to precondition while solving these equations iteratively. These bounds show that the error depends on three terms: First, on the amount of invariance of the constructed subspace for A, second, on the eigenvalues of A at least in proportion to $1/|\text{Re}\,\lambda|$, and third, under a certain condition on projectors $P_l = W_l W_l^*$, on the factor $\min_{X \in \mathbb{C}^{l \times p}} \|B - (\lambda I - A)W_l X\|$ for λ on a path Γ surrounding the spectrum of A. Consequently, in order to compensate for those parts of the spectrum where $1/|\text{Re}\,\lambda|$ is not small, preconditioning or an inverse iteration is needed to keep the sizes of the matrices used in construction of a reduced-order model moderate.

Keywords: model reduction, iterative methods, Lyapunov matrix equations, Hankel operator, preconditioning, Hankel singular values.

1. Introduction and Notation

We analyze the problem of generating iteratively a reduced-order model for a stable linear time-invariant state-space model, denoted by $\Sigma(A, B, C, D)$, of the form

$$\dot{z}(t) = Az(t) + Bu(t)$$
$$y(t) = Cz(t) + Du(t)$$

$$(1)$$

where $A \in \mathbb{C}^{n \times n}$ is a large, possibly sparse, matrix, $B \in \mathbb{C}^{n \times p}$ is the control matrix, $C \in \mathbb{C}^{r \times n}$ is the observation matrix and $D \in \mathbb{C}^{r \times p}$ (typically, $p \ll n$ and/or $r \ll n$). Large-scale systems arise, for instance, from problems in continuum mechanics, where a discretization of a partial-differential operator using finite elements or finite differences leads to a matrix A which is typically large and sparse (Balas, 1982). Well-established model reduction methods, such as the optimal Hankel norm

by Glover (1984) and balanced truncation by Moore (1981), further extended in (Enns, 1984; Pernebo and Silverman, 1982), begin then by solving the Lyapunov equations

$$AL_B + L_B A^* = -BB^*, \quad A^* L_C + L_C A = -C^* C \tag{2}$$

and computing the square roots of the eigenvalues of $L_C L_B$. These are called the Hankel singular values of $G(s) = C(sI - A)^{-1}B + D$, the transfer function of the system (1). The point of all this is that the Hankel singular values yield an easily computable error bound if the original state-space model (1) is approximated with a reduced-order model $\Sigma(A_k, B_k, C_k, D)$, that is, with

$$\begin{aligned}
\dot{z}(t) &= A_k z(t) + B_k u(t) \\
y(t) &= C_k z(t) + D u(t)
\end{aligned} \tag{3}$$

where $A_k \in \mathbb{C}^{k \times k}$, $B_k \in \mathbb{C}^{k \times p}$ and $C_k \in \mathbb{C}^{r \times k}$ are obtained by using projectors originating from the solutions to (2). However, it may not be realistic to solve (2) by dense matrix techniques and some approximation process is needed at that point. In particular, since for very large systems atypical situations can arise as compared with small dimensional systems. For instance, the matrix A need not be available. More precisely, A (or A^*) can exist as a computer code or as an implicit approximate representation of A, so that only output information Ab (or $A^* b$) is available with vectors $b \in \mathbb{C}^n$. This means that we may not have A^* (or A) at our disposal. Obviously, in this case solving the Lyapunov equations is not feasible and the criteria for producing a reduced-order model may need to be rephrased. Another example is that the rank of either B or C is very large, e.g. if $C = I$, the identity matrix. Then approximative techniques may fail as typically a 'small-rank' assumption for both B and C is needed for solving both of the Lyapunov equations. In consequence, effectively only A or A^* is at one's disposal in these cases. (See (Phillips, 1998; Phillips et al., 1996) and references therein for examples of similar situations.)

Bearing these aspects in mind, in this paper we study how to construct an acceptable reduced-order model with iterative methods for large state-space systems by using the Hankel singular values to form a criterion. In particular, we derive error bounds that explicitly show that inverse iteration or preconditioning is necessary in iterative model reduction. For that purpose, suppose that we have an iterative process, typically preconditioned, yielding at the k-th step an orthogonal projector $P_k = W_k W_k^*$. Here W_k is of size $n \times k$ with orthonormal columns. Further, we assume that the range of P_k contains the range of B and/or the range of C^*. By setting $A_k = W_k^* A W_k$, $B_k = W_k^* B$ and $C_k = C W_k$, the corresponding Lyapunov equations are as follows:

$$A_k L_B^k + L_B^k A_k^* = -B_k B_k^*, \quad A_k^* L_C^k + L_C^k A_k = -C_k^* C_k \tag{4}$$

Here L_B^k and L_C^k denote the projected solutions and thus they are small, of size $k \times k$. Approximations to the Lyapunov equations by using non-preconditioned Krylov subspaces were proposed in (Jaimoukha and Kasenally, 1994) where the analysis was based on the residual error, see also (Hu and Reichel, 1992; Saad, 1990). Hochbruck and Starke (1995) proposed preconditioning while solving iteratively the Lyapunov

equations by rewriting them as a linear system by using the Kronecker product and then using preconditioning with the obtained large linear system. Then the approximative solution is also of size $n \times n$. We, instead, work with *small* matrices L_B^k and L_C^k, and derive error bounds directly for the differences

$$L_B - W_k L_B^k W_k^*, \quad L_C - W_k L_C^k W_k^* \tag{5}$$

to control whether the constructed reduced-order model will be acceptable or not. An advantage of this is that the error bounds in this form are more informative as to the behaviour and the origin of the error. Under a certain technical condition on projectors, see (10), we obtain error bounds depending, roughly, on three terms. First, on the invariance of the subspace $P_k \mathbb{C}^n$ for A measured via $\|(I - P_k)AP_k\|$. Second, on the factor $-1/\operatorname{Re}\lambda$ for $\lambda \in \Gamma$ surrounding the spectrum of A. And third, on $\min_{X \in \mathbb{C}^{l \times p}} \|B - (\lambda I - A)W_l X\|$ for $\lambda \in \Gamma$. To make this more concrete, for L_B we show that

$$\left\| L_B - W_k L_B^k W_k^* \right\|$$

$$\leq \epsilon_k \frac{\|B\|\hat{M}}{\pi} \int_\Gamma \frac{-1}{\operatorname{Re}(\lambda - \hat{\omega})} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|} \min_{X \in \mathbb{C}^{l \times p}} \left\| B - (\lambda I - A)W_l X \right\| \mathrm{d}|\lambda|$$

where $\hat{\omega}$ is a positive constant. Here ϵ_k denotes $\|(I - P_k)AP_k\|$, i.e., the largest singular value of an $n \times k$ matrix $(I - W_k W_k^*)AW_k$. The factor $-1/\operatorname{Re}(\lambda - \hat{\omega})$ means that eigenvalues with a large negative real-part contribute less to the error. To compensate this close to the imaginary axis, $\min_{X \in \mathbb{C}^{l \times p}} \|B - (\lambda I - A)W_l X\|$ should be small. This, in turn, can be achieved by preconditioning or via inverse iteration. Namely, if the range of P_k contains $(\lambda I - A)^{-1}B$, then this factor is zero at λ. There are other reasons for preconditioning: First, it may be needed for keeping the approximation scheme stable (Grimme *et al.*, 1995; Huhtanen, 1997). In particular for strongly non-symmetric systems preconditioning seems almost indispensable. Second, the controllable subspace of the original system modelled by the partial-differential operator, which is approximated by (1), can, in general, be found only via inverse iteration.

There is an alternative way to compute an approximation to the Hankel singular values and to control whether a reduced-order model is acceptable. This is based on the Hankel operator. The Hankel operator of (1) is given by

$$(\Gamma u)(t) = \int_0^\infty C e^{A(t+s)} B u(s) \, \mathrm{d}s \tag{6}$$

from $L^2([0, \infty); \mathbb{C}^p)$ to $L^2([0, \infty); \mathbb{C}^r)$ and the Hankel singular values equal the singular values of Γ. The corresponding approximative Hankel operator is then $(\Gamma_k u)(t) = \int_0^\infty C W_k e^{A_k(t+s)} W_k^* B u(s) \, \mathrm{d}s$. Again, we obtain estimates consisting of the above three terms: $\|(I - P_k)AP_k\|$ and $-1/\operatorname{Re}\lambda$ as well as $\min_{X \in \mathbb{C}^{l \times p}} \|B - (\lambda I - A)W_l X\|$ for λ on the path of integration Γ. Consequently, preconditioning or inverse iteration is needed for making the last term small for λ near the imaginary axis.

The paper is organized as follows. In Section 2, we derive estimates to holomorfic functional calculus that are needed while approximating the Lyapunov equations and the Hankel operator. In Section 3, we derive bounds for the Lyapunov equations and the Hankel operator. In Section 4, we briefly explain why inverse iteration is the right choice when distributed-parameter systems are being approximated. Then we end the paper with two realistic numerical examples, one symmetric and another non-symmetric. Without preconditioning the sizes of matrices used in the construction of a reduced-order model would become unbearably large in both cases. Instead, if P_k is generated via inverse iteration, the orders remain moderate.

2. An Approximation to the Block-Holomorfic Functional Calculus

We study the problem of projecting the state-space model (1) to a smaller dimensional system (3). Typically, this problem arises when $\Sigma(A, B, C, D)$ needs to be of a high-order so as to be an accurate approximation of the physical system being modelled, let us say, via partial-differential equations. However, because the order of the discretized model is high, it is often unrealistic to compute its dynamics numerically and therefore a smaller dimensional system is needed for practical control. Then another problem arises: How to effectively and efficiently construct a low-order model that is close to the original one? Recently projection-based formulations using Krylov subspace techniques have become generally accepted.

Let us suppose that we have an orthogonal projector P_k and the dynamics of (1) is projected onto the subspace $P_k \mathbb{C}^n$ via (3). If $P_k \mathbb{C}^n$ is an invariant subspace of A such that $P_k B = B$, then we are satisfied for the following reason. The dynamics of (1) is completely characterized by the projected system as $P_k \mathbb{C}^n$ contains the controllable subspace of $\Sigma(A, B, C, D)$ *and* $\Sigma(A_k, B_k, C_k, D)$ yields a realization of the transfer function of (1). If $P_k \mathbb{C}^n$ is *not* invariant for A, then the remaining alternative is: How invariant is $P_k \mathbb{C}^n$ for A? One possibility is to derive estimates using the invariance of a subspace based on *quasitriangularity* (Halmos, 1968). In the context of numerical linear algebra this provides, roughly speaking, a global residual for a set of Ritz pairs. The applicability of quasitriangularity to numerics was noticed in (Nevanlinna, 1995; Nevanlinna and Vainikko, 1996) where it was used in spectral approximation, see also (Huhtanen, 1997; 1998) and the recent paper (Knyazev, 1998). Based on this approach, we apply the concept to obtain an approximation to the holomorfic functional calculus for matrices. More precisely, we obtain estimates for the accuracy of the approximation in the subspace $P_k \mathbb{C}^n$, not just for a single vector. This yields a way to control the error when equations involving matrices, like the Lyapunov equations or the Hankel operator, are being solved approximately. Thereby we can also assess whether the constructed low-order model $\Sigma(A_k, B_k, C_k, D)$ is close to the original one.

Let us assume that we have constructed a projector $P_k = W_k W_k^*$ by using an iterative method. That is, we suppose that W_k is an $n \times k$ matrix having orthonormal columns that span the range of P_k, and that the columns of W_k are obtained via an iterative scheme. Our basic assumption in this paper is that preconditioning of

some sort is used while generating W_k, but, to give an example, a simple way to produce W_k is to use the Arnoldi method. For the definition of the Arnoldi method, see e.g. (Saad, 1992). Then the columns of W_k form an orthonormal basis of the Krylov subspace

$$\mathcal{K}_k(A;b) = \text{span}\{b, Ab, \ldots, A^{k-1}b\} \tag{7}$$

generated with A from a vector $b \in \mathbb{C}^n$. Let us designate by A_k the restriction of $P_k A P_k$ to $P_k \mathbb{C}^n$ with the matrix representation $W_k^* A W_k$. In our notation $\|\cdot\|$ denotes the spectral norm and $\text{rank}(A)$ stands for the rank of A. Furthermore, let us introduce the quantity

$$\epsilon_k := \left\| (I - P_k) A P_k \right\| = \left\| (I - W_k W_k^*) A W_k \right\| \tag{8}$$

to measure the invariance of $P_k \mathbb{C}^n$ for A. This gives a compromise in case $P_k \mathbb{C}^n$ is not an invariant subspace for A, but there is, instead, a 'small overflow.' To put it in another way, if ϵ_k is small, then the subspace $P_k \mathbb{C}^n$ is roughly invariant for A. Clearly, one obtains this quantity inexpensively as it is given by the largest singular value of the $n \times k$ matrix $(I - W_k W_k^*) A W_k$. Note that this quantity is also used in the Arnoldi method: For a single vector b of unit length this means that when the columns of W_k form an orthonormal basis of (7), then A_k is a Hessenberg matrix fulfilling

$$AW_k = W_k A_k + \epsilon_k b_{k+1} e_k^*$$

where e_k is the k-th unit vector in $\mathbb{C}^n$.

While working with the matrix representation A_k, we need the following relation to the original matrix A. Here $\rho(A)$ denotes the resolvent set of A.

Lemma 1. (Nevanlinna and Vainikko, 1996) *Let* $A \in \mathbb{C}^{n \times n}$ *and* $\lambda \in \rho(A)$. *If* $\epsilon_k \|(\lambda I - A)^{-1}\| < 1$, *then* $\lambda \in \rho(A_k)$ *and*

$$\left\| (\lambda I_k - A_k)^{-1} \right\| \leq \frac{\|(\lambda I - A)^{-1}\|}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|}$$

First of all, once ϵ_k is small, this gives a sufficient condition for the existence of the inverse of $\lambda I_k - A_k$. Second, this yields a sufficient condition for the spectrum of A_k to belong to a neighborhood of $\sigma(A)$, the spectrum of A. This neighborhood is determined by the level sets of the norm of the resolvent matrix, i.e. the ϵ-*pseudospectra* of A (Trefethen, 1992). Third, by using this, it can be shown that $W_k(\lambda I_k - A_k)^{-1} W_k^*$ is, for a small ϵ_k, a good approximation to $(\lambda I - A)^{-1}$ restricted to $P_k \mathbb{C}^n$.

Proposition 1. *Let* $A \in \mathbb{C}^{n \times n}$ *and* $\lambda \in \rho(A)$. *If* $\epsilon_k \|(\lambda I - A)^{-1}\| < 1$, *then for* $x \in P_k \mathbb{C}^n$ *of unit length we have*

$$\left\| W_k(\lambda I_k - A_k)^{-1} W_k^* x - (\lambda I - A)^{-1} x \right\| \leq \epsilon_k \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|}$$

Proof. By Lemma 1 we have $\lambda \in \rho(A_k)$. Thus we obtain

$$(\lambda I - A) \frac{W_k(\lambda I_k - A_k)^{-1}W_k^* x}{\|W_k(\lambda I_k - A_k)^{-1}W_k^* x\|} = W_k(\lambda I_k - A_k)W_k^* \frac{W_k(\lambda I_k - A_k)^{-1}W_k^* x}{\|W_k(\lambda I_k - A_k)^{-1}W_k^* x\|} + y$$

$$= \frac{x}{\|W_k(\lambda I_k - A_k)^{-1}W_k^* x\|} + y$$

for some $y \in \mathbb{C}^n$ of norm at most ϵ_k. So as $\lambda \in \rho(A)$, we get

$$\frac{1}{\|(\lambda I - A)^{-1}\|} \left\| \left(W_k(\lambda I_k - A_k)^{-1}W_k^* - (\lambda I - A)^{-1} \right) x \right\|$$

$$\leq \left\| (\lambda I - A) \left(W_k(\lambda I_k - A_k)^{-1}W_k^* - (\lambda I - A)^{-1} \right) x \right\|$$

$$= \|y\| \left\| W_k(\lambda I_k - A_k)^{-1}W_k^* x \right\| \leq \epsilon_k \left\| (\lambda I_k - A_k)^{-1} \right\|$$

from which the claim follows after multiplying both sides by $\|(\lambda I - A)^{-1}\|$ and using Lemma 1 again. ∎

Plugging the above approximation into Dunford's integral formula, we obtain an approximation to the holomorfic functional calculus for matrices. Here we assume that f is analytic in the neighborhood G of the spectrum of A and $\epsilon_k \|(\lambda I - A)^{-1}\| < 1$ on the path of integration Γ.

Corollary 1. *If $x \in P_k \mathbb{C}^n$ is of unit length, then*

$$\left\| W_k f(A_k) W_k^* x - f(A)x \right\| \leq \frac{\epsilon_k}{2\pi} \int_\Gamma |f(\lambda)| \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k\|(\lambda I - A)^{-1}\|} \, \mathrm{d}|\lambda|$$

where Γ is a path surrounding $\sigma(A)$ in G.

Proof. As $\epsilon_k\|(\lambda I - A)^{-1}\| < 1$, we are in the resolvent set of A_k as well, hence

$$W_k f(A_k) W_k^* x = \frac{1}{2\pi i} \int_\Gamma f(\lambda) W_k(\lambda I_k - A_k)^{-1}W_k^* x \, \mathrm{d}\lambda$$

so that, after subtracting $f(A)x = \frac{1}{2\pi i} \int_\Gamma f(\lambda)(\lambda I - A)^{-1}x \, \mathrm{d}\lambda$ from it and taking the norms, we get the claim. ∎

For the case $f(\lambda) = e^\lambda$ and for a single vector this problem has been studied in a number of papers, see e.g. (Saad, 1992). In (Hochbruck and Lubich, 1997) estimates are derived based on the use of the Faber polynomials such that the path of integration Γ surrounds the field of values of A. Recall that the field of values of A is the set $\{x^*Ax \in \mathbb{C} : x^*x = 1\}$ (Horn and Johnson, 1991). These papers cover the Arnoldi and the non-symmetric Lanczos methods for A. Since we are concerned with possibly very non-symmetric systems $\Sigma(A, B, C, D)$, possibly with $\mathrm{rank}(B) > 1$ and $\mathrm{rank}(C) > 1$, preconditioning may be needed and therefore we need more general bounds to cover block-analogies. In particular, we will show later that it is important, for obtaining good bounds, to be able to use paths Γ cutting through the field of values of A. The condition $\epsilon_k\|(\lambda I - A)^{-1}\| < 1$ allows us to do this.

Corollary 1 states that as soon as ϵ_k is small, $f(A)$ is well approximated in the whole subspace $P_k \mathbb{C}^n$ and not just for a single vector. In particular, the variation of constants formula provided by the control matrix B is well approximated if the range of P_k contains the range of B. However, for better bounds, notice the following property: For any $X \in \mathbb{C}^{k \times p}$ we have

$$\left((\lambda I - A)^{-1} - W_k (\lambda I_k - A_k)^{-1} W_k^* \right) B$$

$$= \left((\lambda I - A)^{-1} - W_k (\lambda I_k - A_k)^{-1} W_k^* \right) \left(B - (\lambda I - A) W_k X \right) \quad (9)$$

since $W_k^* (\lambda I - A) W_k = (\lambda I_k - A_k)$. For our bounds we need, additionally, that the range of $(\lambda I - A) W_k X$ be contained in the range of P_k, so we make the following assumption: We say that the scheme for producing W_k's has the property $\mathcal{P}_{l,k}$ for $l < k$, if:

The first l columns of W_k equal W_l and $\text{span}\{AW_l\} \subset \text{span}\{W_k\}$. (10)

This restriction is obviously not the most general for the needed condition, since we need only to see that X is such that $(\lambda I - A) W_k X$ is contained in the range of P_k. However, this condition is very easy to check and several important schemes fulfil it. For instance, in the block version of the Arnoldi method columns for W_k are formed from subspaces $\mathcal{K}_k(A; B) = \text{span}\{B, AB, \dots, A^{k-1}B\}$ so that it clearly has the property $\mathcal{P}_{l,k}$. Another scheme that fulfils this is the inverse iteration defined as follows: Using the Arnoldi method with A^{-1} applied to B, construct from

$$\mathcal{K}_l(A^{-1}; B) = \text{span}\{B, A^{-1}B, \dots, A^{-l+1}B\}$$

W_l and then generate W_k from this by using the subspace $\text{span}\{W_l, AB\}$. This scheme obviously has the property $\mathcal{P}_{l,k}$.

With this notation we obtain the following bounds: Again we assume that f is analytic in the neighborhood G of the spectrum of A and $\epsilon_k \|(\lambda I - A)^{-1}\| < 1$ on the path of integration.

Corollary 2. *Suppose that $P_k B = B$ and the scheme has the property $\mathcal{P}_{l,k}$. Then*

$$\left\| W_k f(A_k) W_k^* B - f(A) B \right\|$$

$$\leq \frac{\epsilon_k}{2\pi} \int_\Gamma |f(\lambda)| \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|} \min_{X \in \mathbb{C}^{l \times p}} \left\| B - (\lambda I - A) W_l X \right\| \mathrm{d}|\lambda|$$

where Γ is a path surrounding $\sigma(A)$ in G.

Proof. Thus, using again Dunford's integral formula, we obtain from Proposition 1 in a similar manner

$$\left\| W_k f(A_k) W_k^* B - f(A) B \right\| \leq \frac{1}{2\pi} \int_\Gamma |f(\lambda)| \left\| \left(W_k (\lambda I_k - A_k)^{-1} W_k^* - (\lambda I - A)^{-1} \right) B \right\| \mathrm{d}\lambda$$

$$= \frac{1}{2\pi} \int_\Gamma |f(\lambda)| \left\| \left(W_k (\lambda I_k - A_k)^{-1} W_k^* - (\lambda I - A)^{-1} \right) (B - (\lambda I - A) W_l X) \right\| \mathrm{d}\lambda$$

since $P_k(\lambda I - A)W_l X = (\lambda I - A)W_l X$ for all $X \in \mathbb{C}^{l \times p}$. Consequently, using Proposition 1 we obtain

$$\left\| \left(W_k(\lambda I_k - A_k)^{-1}W_k^* - (\lambda I - A)^{-1}\right)(B - (\lambda I - A)W_l X)\right\|$$

$$\leq \epsilon_k \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k\|(\lambda I - A)^{-1}\|}\left\|(B - (\lambda I - A)W_l X)\right\|$$

and the claim follows. ∎

After the projector is fixed, an interesting question clearly is: How to derive simple pointwise bounds to $|f(\lambda)| \min_{X \in \mathbb{C}^{l \times p}} \|B - (\lambda I - A)W_l X\|$ for $\lambda \in \Gamma$? Of course, if we take $X = 0$ for every $\lambda \in \Gamma$, we trivially get $|f(\lambda)|\|B\|$. A further aspect becomes apparent too: How well B is approximated by $(\lambda I - A)W_k X$ depends strongly on the way in which W_k is generated, that is, on the chosen iterative method and preconditioning.

3. Approximations to the Lyapunov Equations and the Hankel Operator

In this section, we use the estimates of Section 2 to derive error bounds for approximations to the Lyapunov equations and the Hankel operator. Our purpose is to derive bounds that show what kind of preconditioning strategy should be considered while reducing the order of a large state-space system in an effective manner.

3.1. Bounds for the Lyapunov Equations

Let us first study how the solutions to the Lyapunov equations

$$AL_B + L_B A^* = -BB^*, \quad A^* L_C + L_C A = -C^* C$$

can be approximated if the analysis is based on Corollary 2. For that purpose, we suppose that $A \in \mathbb{C}^{n \times n}$ is exponentially stable with $\|e^{At}\| \leq Me^{-\omega t}$, where $M \geq 1$ and $\omega > 0$, and that the approximation scheme yielded by projectors P_k is uniformly stable, i.e.

$$\left\|e^{A_k t}\right\| \leq M_0 e^{-\omega_0 t} \quad \text{for} \quad M_0 \geq 1 \quad \text{and} \quad \omega_0 > 0 \tag{11}$$

Theorem 1. *Suppose that (11) holds, $P_k B = B$ and the scheme has the property $\mathcal{P}_{l,k}$. Then with $\epsilon_k = \|(I - P_k)AP_k\|$ and $\epsilon_k\|(\lambda I - A)^{-1}\| < 1$ on Γ we have*

$$\left\|L_B - W_k L_B^k W_k^*\right\|$$

$$\leq \epsilon_k \frac{\|B\|\hat{M}}{\pi} \int_\Gamma \frac{-1}{\text{Re}(\lambda - \hat{\omega})} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k\|(\lambda I - A)^{-1}\|} \min_{X \in \mathbb{C}^{l \times p}} \left\|B - (\lambda I - A)W_l X\right\| d|\lambda|$$

where $\hat{\omega} = \min\{\omega, \omega_0\}$, $\hat{M} = \max\{M, M_0\}$ and Γ surrounds $\sigma(A)$ in $\mathbb{C}^-$.

Proof. From the representation of the corresponding solutions in the form

$$L_B = \int_0^\infty e^{As} BB^* e^{A^*s}\, ds, \quad W_k L_B^k W_k^* = \int_0^\infty W_k e^{A_k s} W_k^* BB^* W_k e^{A_k^* s} W_k^*\, ds$$

we obtain after subtracting

$$
\begin{aligned}
L_B - W_k L_B^k W_k^* &= \int_0^\infty e^{As} BB^* e^{A^*s} - W_k e^{A_k s} W_k^* BB^* W_k e^{A_k^* s} W_k^*\, ds \\
&= \int_0^\infty e^{As} BB^* (e^{A^*s} - W_k e^{A_k^* s} W_k^*)\, ds \\
&\quad + \int_0^\infty (e^{As} - W_k e^{A_k s} W_k^*) BB^* W_k e^{A_k^* s} W_k^*\, ds
\end{aligned}
$$

where we added and subtracted $e^{As} BB^* W_k e^{A_k^* s} W_k^*$ from the integrand. Consequently, this yields, after taking the norms and using $P_k B = B$,

$$\left\| L_B - W_k L_B^k W_k^* \right\| \leq \|B\| \int_0^\infty \left(\|e^{As}\| + \|e^{A_k s}\| \right) \left\| (W_k e^{A_k s} W_k^* - e^{As}) B \right\|\, ds$$

$$\leq \|B\| \hat{M} \frac{1}{\pi} \int_\Gamma \frac{-1}{\operatorname{Re}(\lambda - \hat\omega)} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|} \min_{X \in \mathbb{C}^{l \times p}} \left\| B - (\lambda I - A) W_l X \right\|\, d|\lambda|$$

by Corollary 2 so that the claim follows then after changing the order of integration by Fubini's theorem and performing the integration. ∎

A simple, but crude, upper bound is obtained if we set $X = 0$ for every λ, so that then $\|B - (\lambda I - A) W_l X\| = \|B\|$ on Γ. Moreover, note that if $P_k \mathbb{C}^n$ is invariant for A, with the constraint $P_k B = B$, then $\epsilon_k = 0$ and the Lyapunov equation is solved exactly. This corresponds to the case when there exists a smaller dimensional realization of the plant $\Sigma(A, B, C, D)$.

If the subspace construction is made with C, then in a similar vein we obtain the following for the solution L_C to the second Lyapunov equation. Then A_k designates the adjoint of the restriction of $P_k A^* P_k$ to $P_k \mathbb{C}^n$. More precisely, $W_k^* A^* W_k$ is the matrix representation of the adjoint of A_k.

Corollary 3. *Suppose that (11) holds, $P_k C^* = C^*$ and the scheme has the property $\mathcal{P}_{l,k}$ for A^*. For $\hat\epsilon_k = \|(I - P_k) A^* P_k\|$ and $\hat\epsilon_k \|(\lambda I - A^*)^{-1}\| < 1$ on Γ we get*

$$\left\| L_C - W_k L_C^k W_k^* \right\|$$

$$\leq \hat\epsilon_k \frac{\|C\| \hat{M}}{\pi} \int_\Gamma \frac{-1}{\operatorname{Re}(\lambda - \hat\omega)} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \hat\epsilon_k \|(\lambda I - A)^{-1}\|} \min_{X \in \mathbb{C}^{l \times r}} \left\| C^* - (\lambda I - A^*) W_l X \right\|\, d|\lambda|$$

where $\hat\omega = \min\{\omega, \omega_0\}$, $\hat{M} = \max\{M, M_0\}$ and Γ surrounds $\sigma(A^*)$ in $\mathbb{C}^-$.

Again, if $P_k \mathbb{C}^n$ is invariant for A^*, with the constraint $P_k C^* = C^*$, then the Lyapunov equation is solved exactly.

To get an illustration of these bounds, let us consider a simple example where we approximate just the integral term

$$\int_\Gamma \frac{-1}{\mathrm{Re}\,\lambda} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|} \, \mathrm{d}|\lambda| \tag{12}$$

Example 1. In this example, we study the contribution of a single eigenvalue to the error. Let us consider the stabilized heat equation

$$\begin{cases} \dot{z}(t) = \dfrac{\partial^2}{\partial x^2} z - 4\pi^2 z + Bu(t) \\[2mm] y(t) = Cz(t) + Du(t) \end{cases}$$

with the boundary conditions $z(0,t) = z(1,t) = 0$. Suppose that A is a discretization of the stabilized partial differential operator $\partial^2(\cdot)/\partial x^2 - 4\pi^2$ such that A has the eigenvalues $-(4 + n^2)\pi^2$ for $n = 0, 1, \ldots, 100$. If we surround the n-th eigenvalue of A with the circle Γ_n of radius one, then, as soon as $\epsilon_k \leq 1/2$,

$$\int_{\Gamma_n} \frac{-1}{\mathrm{Re}\,\lambda} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|} \, \mathrm{d}|\lambda| \leq \frac{4\pi}{(4 + n^2)\pi^2 - 1} \approx \frac{4}{\pi n^2}$$

However, we can do much better. If, instead, we take Γ_n to be a circle surrounding the n-th eigenvalue of A with radius $(n^2 - (n-1)^2)\pi^2/2 = (n - 1/2)\pi^2$, i.e. Γ_n is chosen such that the minimum distance to the n-th and $(n-1)$-th eigenvalue is the same, then

$$\int_{\Gamma_n} \frac{-1}{\mathrm{Re}\,\lambda} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|} \, \mathrm{d}|\lambda| \leq \frac{4}{\pi^3} \frac{1}{(n^2 - \frac{3n}{2} + \frac{5}{4})(n - \frac{1}{2})} \approx \frac{4}{\pi^3 n^3}$$

$$\blacklozenge$$

From the above example it is apparent that there are two factors that can make the contribution of (12) to the error small. First, the real parts of the eigenvalues are large negative. Second, if the eigenvalues are isolated and clustered such that they can be surrounded by a large circle. Above we encircled just one eigenvalue by cutting the field of values of A and then used the fast decay of the term $\|(\lambda I - A)^{-1}\|^2$ far from the eigenvalues. In (Hochbruck and Lubich, 1997) there are estimates for the analytic functional calculus for the Arnoldi method with A applied to B, i.e. without preconditioning, but only with paths surrounding the field of values of A. To be able to integrate inside the field of values of A is important since it is relatively easy to construct examples with nilpotent weighted shifts that have a huge field of values and an arbitrarity small ϵ_k. As soon as ϵ_k is of a reasonable size, we can integrate with paths inside the field of values of A.

3.2. Bounds for the Hankel Operator

By the previous subsection it seems that two sequences of projectors are needed for a good approximation to the Hankel singular values: One with A using B and another with A^* using C^*. Then, with the computed approximations to L_B and L_C, a reduced-order model is constructed with, for instance, balanced truncation. However, the construction of two sequences of projectors complicates the model reduction problem and, needless to say, makes it more expensive. Another factor that can make this approach unrealistic is the assumption that all dense matrix manipulations are feasible. For instance, the availability of the adjoint of A need not hold in large-scale computations. A very good description of situations of this type can be found in (Phillips, 1998; Phillips *et al.*, 1996). Therefore, to generate a reduced-order model in these cases, we next show that a single sequence of projectors suffices and that one can compute approximations to Hankel singular values by using A (or A^*) only. Then a justified way to create a reasonable reduced-order model is to set $A_k = W_k^* A W_k$ (or $A_k^* = W_k^* A^* W_k$, respectively), $B_k = W_k^* B$ and $C_k = C W_k$.

Thus, let us consider the Hankel operator of $\Sigma(A, B, C)$ defined by

$$(\Gamma u)(t) = \int_0^\infty C e^{A(t+s)} B u(s) \, \mathrm{d}s \tag{13}$$

from $L^2([0,\infty); \mathbb{C}^p)$ to $L^2([0,\infty); \mathbb{C}^r)$. The corresponding approximative Hankel operator equals $(\Gamma_k u)(t) = \int_0^\infty C W_k e^{A_k(t+s)} W_k^* B u(s) \, \mathrm{d}s$. To start with, we need the following auxiliary result for the pointwise error. To simplify the notations we put

$$G(\lambda) := \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|} \min_{X \in \mathbb{C}^{l \times p}} \|B - (\lambda I - A) W_l X\|$$

Lemma 2. *Suppose that $P_k B = B$ and the scheme has the property $\mathcal{P}_{l,k}$. Then with $\epsilon_k = \|(I - P_k) A P_k\|$ and $\epsilon_k \|(\lambda I - A)^{-1}\| < 1$ on Γ we get*

$$\|(\Gamma u - \Gamma_k u)(t)\| \leq \frac{\epsilon_k}{2\pi} \|C\| \, \|u\|_2 \left(\int_\Gamma \frac{-|e^{\lambda t}|}{\operatorname{Re} \lambda} G(\lambda) \, \mathrm{d}|\lambda| \right)^{1/2} \left(\int_\Gamma |e^{\lambda t}| G(\lambda) \, \mathrm{d}|\lambda| \right)^{1/2}$$

where Γ surrounds $\sigma(A)$ in $\mathbb{C}^-$.

Proof. After subtracting the solutions and then applying Hölder's inequality we obtain

$$\|(\Gamma u - \Gamma_k u)(t)\| \leq \int_0^\infty \left\| C(e^{A(t+s)} - W_k e^{A_k(t+s)} W_k^*) B u(s) \right\| \, \mathrm{d}s$$

$$\leq \int_0^\infty \|C\| \left\| (e^{A(t+s)} - W_k e^{A_k(t+s)} W_k^*) B \right\| \, \|u(s)\| \, \mathrm{d}s$$

$$\leq \|C\| \, \|u\|_2 \left(\int_0^\infty \left\| (e^{A(t+s)} - W_k e^{A_k(t+s)} W_k^*) B \right\|^2 \, \mathrm{d}s \right)^{1/2}$$

$$\leq \frac{\epsilon_k}{2\pi} \|C\| \, \|u\|_2 \left(\int_0^\infty \int_\Gamma |e^{\lambda_1(t+s)}| G(\lambda_1) \, \mathrm{d}|\lambda_1| \right.$$

$$\left. \times \int_\Gamma |e^{\lambda_2(t+s)}| G(\lambda_2) \, \mathrm{d}|\lambda_2| \, \mathrm{d}s \right)^{1/2}$$

by using Corollary 2, since $P_k B = W_k W_k^* B = B$. Then from Fubini's theorem, while integrating with respect to s, and by using the approximation (again, Γ is supposed to surround the spectrum of A in $\mathbb{C}^-$)

$$\frac{-|e^{(\lambda_1 + \lambda_2)t}|}{\operatorname{Re}(\lambda_1 + \lambda_2)} \leq \frac{-|e^{(\lambda_1 + \lambda_2)t}|}{\operatorname{Re}\lambda_1}$$

the claim follows.   ∎

Using this we obtain the following error bound for the Hankel operator approximation. The norm is the operator norm from $L^2([0, \infty); \mathbb{C}^p)$ to $L^2([0, \infty); \mathbb{C}^r)$.

Theorem 2. *Suppose that $P_k B = B$ and the scheme has the property $\mathcal{P}_{l,k}$. Then with $\epsilon_k = \|(I - P_k)A P_k\|$ and $\epsilon_k \|(\lambda I - A)^{-1}\| < 1$ on Γ we have*

$$\|\Gamma - \Gamma_k\| \leq \frac{\epsilon_k}{2\pi} \|C\| \int_\Gamma \frac{-1}{\operatorname{Re}\lambda} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \epsilon_k \|(\lambda I - A)^{-1}\|} \min_{X \in \mathbb{C}^{s \times p}} \|B - (\lambda I - A)W_l X\| \, d|\lambda|$$

where Γ surrounds $\sigma(A)$ in $\mathbb{C}^-$.

Proof. Taking the inner product gives, after using Lemma 2,

$$\int_0^\infty \left\| (\Gamma u - \Gamma_k u)(t) \right\|^2 dt$$

$$\leq \frac{\epsilon_k^2}{4\pi^2} \|C\|^2 \|u\|_2^2 \int_0^\infty \int_\Gamma \frac{-|e^{\lambda_1 t}|}{\operatorname{Re}\lambda_1} G(\lambda_1) \, d|\lambda_1| \int_\Gamma |e^{\lambda_2 t}| G(\lambda_2) \, d|\lambda_2| \, dt$$

Then from Fubini's theorem, while integrating with respect to t and by using now the approximation

$$\frac{-1}{\operatorname{Re}(\lambda_1 + \lambda_2)} \leq \frac{-1}{\operatorname{Re}\lambda_2}$$

we obtain the claim.   ∎

It is clear how the Hankel singular values are approximated as they equal the singular values of Γ, since after adding and subtracting $\Gamma \Gamma_k^*$ we have

$$\|\Gamma \Gamma^* - \Gamma_k \Gamma_k^*\| \leq \left(\|\Gamma\| + \|\Gamma_k\| \right) \|\Gamma - \Gamma_k\| \leq \left(2\|\Gamma\| + \|\Gamma - \Gamma_k\| \right) \|\Gamma - \Gamma_k\|$$

In some cases it may be worthwhile to construct P_k with C^*. This could be the case if the rank of C is smaller than B or if there exists a low dimensional realization of $\Sigma(A, B, C, D)$ based on a small rank P_k for which $(I - P_k)A^* P_k = 0$ with $P_k C^* = C^*$. For that purpose, we state the adjoint claim which is proven in a similar vein.

Corollary 4. *Suppose that $P_k C^* = C^*$ and the scheme has the property $\mathcal{P}_{l,k}$ for A^*. Then with $\hat{\epsilon}_k = \|(I - P_k)A^* P_k\|$ and $\hat{\epsilon}_k \|(\lambda I - A^*)^{-1}\| < 1$ on Γ we have*

$$\|\Gamma - \Gamma_k\| \leq \frac{\hat{\epsilon}_k}{2\pi} \|B\| \int_\Gamma \frac{-1}{\operatorname{Re}\lambda} \frac{\|(\lambda I - A)^{-1}\|^2}{1 - \hat{\epsilon}_k \|(\lambda I - A)^{-1}\|} \min_{X \in \mathbb{C}^{s \times r}} \|C^* - (\lambda I - A^*)W_l X\| \, d|\lambda|$$

where Γ surrounds $\sigma(A^*)$ in $\mathbb{C}^-$.

As a summary for an iterative Hankel operator approximation, we have the following: There are two factors, in addition to the amount of invariance of a constructed subspace, that make the contribution of (12) to the error small. First, (12) is small for the eigenvalues with large negative real parts and, second, it is small for clustered eigenvalues which can be surrounded within a large circle. Therefore, in order to reduce the effect of the eigenvalues close to the imaginary axis, preconditioning that takes into account B or/and C^* is needed for keeping

$$g(\lambda) := \min_{X \in \mathbb{C}^{l \times p}} \left\| B - (\lambda I - A) W_l X \right\| \tag{14}$$

small on Γ near the imaginary axis.

4. Inverse Iteration and Preconditioning in Reduction of the Order of a Model

In this final section, we first motivate the use of the inverse iteration, or some appropriate preconditioning, in iterative model reduction purely from a mathematical point of view. Finally, we give numerical examples.

4.1. Construction of Projectors for Discretizations of Partial Differential Equations

Before giving numerical examples, we briefly describe why projectors obtained by simply applying A to B may not be a good choice for iterative model reduction of discretizations of systems modelled by partial-differential equations. For that purpose, we have to recall some aspects of infinite dimensional system theory.

It is well-known that $\mathcal{K}(A; B) = \mathrm{span}\{B, AB, \ldots, A^{n-1}B\}$ equals the controllable subspace of the system (1). For computation of the controllable subspace via iterative methods, see (Boley and Golub, 1991) and references therein. Obviously, the Arnoldi method applied with A to B, yielding projectors P_k with the range

$$\mathcal{K}_k(A; B) = \mathrm{span}\{B, AB, A^2B, \ldots, A^{k-1}B\} \tag{15}$$

is probably the most attractive and inexpensive candidate to be used in iterative model reduction. However, for large systems originating from discretizations of partial-differential systems the quality of the subspace $\mathcal{K}_k(A; B)$ can be unsatisfactory. To give mathematical reasons for this statement, let us suppose that (1) is a discretization of a state-space model, denoted by $\Sigma(\hat{A}, \hat{B}, \hat{C}, \hat{D})$,

$$\dot{z}(t) = \hat{A}z(t) + \hat{B}u(t)$$
$$y(t) = \hat{C}z(t) + \hat{D}u(t) \tag{16}$$

where $\hat{A}$ is a generator of a strongly continuous semigroup $S(t)$ on a separable Hilbert space H. Further, $\hat{B}$ is a bounded linear operator from $\mathbb{C}^p$ to H, $\hat{C}$ is

a bounded linear operator from H to $\mathbb{C}^r$ and $\hat{D} : \mathbb{C}^p \to \mathbb{C}^r$. For $\Sigma(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ the states that can be approximatively reached with the controllability map $\mathcal{B}^\tau u = \int_0^\tau S(\tau - s)\hat{B}u(s) \, ds$, defined from $L^2([0, \tau]$ to H, are obtained as follows:

Definition 1. The approximatively controllable subspace of $\Sigma(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ is the closure of $\bigcup_{\tau > 0} \{\mathcal{B}^\tau u \in H : u \in L^2([0, \tau]; \mathbb{C}^p)\}$.

The following should be a well-known fact:

Lemma 3. (Curtain and Zwart, 1995) *The approximatively controllable subspace of* $\Sigma(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ *equals the smallest closed* $S(t)$-*invariant subspace containing the range of* $\hat{B}$.

As to iterative methods, the important point is that the smallest closed $S(t)$-invariant subspace containing the range of $\hat{B}$ is precisely

$$\mathcal{K}\big((\lambda I - \hat{A})^{-1}; \hat{B}\big) = \overline{\text{span}}\{\hat{B}, (\lambda I - \hat{A})^{-1}\hat{B}, (\lambda I - \hat{A})^{-2}\hat{B}, \dots\}$$

for λ in the component of the resolvent set of $\hat{A}$ containing the interval $[r, \infty)$ for an $r \in \mathbb{R}$. This can be seen from the following:

Theorem 3. *The approximatively controllable subspace of* $\Sigma(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ *equals* $\mathcal{K}((\lambda I - \hat{A})^{-1}; \hat{B})$.

Proof. First, a closed subspace is invariant for $S(t)$ if and only if it is invariant for $(\lambda I - \hat{A})^{-1}$ (Curtain and Zwart, 1995, Lemma 2.5.6). By construction, $\mathcal{K}((\lambda I - \hat{A})^{-1}; \hat{B})$ is an invariant subspace of $(\lambda I - \hat{A})^{-1}$ with the additional property that it is the smallest invariant subspace of $(\lambda I - \hat{A})^{-1}$ containing the range of $\hat{B}$. ∎

Clearly, for an unbounded $\hat{A}$, $\mathcal{K}(\hat{A}; \hat{B})$ does not, in general, make any sense. And even if it does, it need not have anything to do with the approximatively controllable subspace of $\Sigma(\hat{A}, \hat{B}, \hat{C}, \hat{D})$. Let us illustrate this with an example.

Example 2. This is an adaptation of (Zwart, 1989, Example I.6). Let $H = L^2(0, 1)$ and $\hat{A} = \partial^2(\cdot)/\partial x^2$ with the domain consisting of all $f \in H$ with the second derivative in H such that $f(0) = f(1) = 0$. With $\chi_{[1/2,1]}$, the indicator function of the interval $[1/2, 1]$, we have

$$g(s) = \hat{A}^{-1}\big(\chi_{[1/2,1]}\pi^2 \sin(\pi s)\big) = \begin{cases} -s, & s \in [0, 1/2] \\ -s + 1 - \sin(\pi s), & s \in [1/2, 1] \end{cases}$$

Take a smooth function $\hat{b}(s)$ which is close in L^2-norm to $\chi_{[1/2,1]}\pi^2 \sin(\pi s)$ such that it is equal to zero for $s \in [0, 1/2 - \epsilon]$ and which has derivatives $\hat{b}^j(1) = 0$ for all $j = 1, 2, \dots$. Thus, $\mathcal{K}(\hat{A}; \hat{b})$ is well-defined. However, all the elements of $\mathcal{K}(\hat{A}; \hat{b})$ are equal to zero in the interval $[0, 1/2 - \epsilon]$. But by the continuity of $\hat{A}^{-1}$, the point $\hat{A}^{-1}\hat{b}$ is close to $g(s)$, which is not equal to zero on that interval and thus the approximatively controllable subspace is very different from $\mathcal{K}(\hat{A}; \hat{b})$. ◆

Thus, with the heat operator and with a very simple control operator the approximatively controllable subspace of $\Sigma(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ can be completely different from the forward iterated subspace. This is a partial reason why the quality of $\mathcal{K}_k(A; B)$ may be insufficient for an approximation of the controllable subspace of (1), unless k is very large, that is, unrealistically large for practical computations.

To sum up, the inverse iteration, or preconditioning, is a good alternative for a projector construction in model reduction for the following reasons: First, in order to make (14) small on Γ near the imaginary axis. Second, it is needed for keeping the approximations stable. And third, the forward iteration is not generally well-defined for infinite-dimensional unbounded systems, and, even if it were, it would not yield approximatively controllable states of the system. In finite dimensions and for small k this shows up in the possible poor quality of the Krylov subspace $\mathcal{K}_k(A; B)$ while approximating the subspaces corresponding to the largest Hankel singular values.

Instead of using the inverse of A in the subspace construction, there is, as already mentioned, the less expensive possibility of using some sort of approximation for the inverse of A as a substitute. In other words, this means using a preconditioner for A while constructing subspaces. Recall that a preconditioner for A is a matrix M that has the property that AM is close to the unit matrix I in some manner. As to finding a good preconditioner for A, at the outset there is virtually no limits to available options for a good preconditioner. For example, preconditioners can be derived from knowledge of an original physical problem from which the linear system arises. Therefore, finding a good preconditioner is often viewed as a combination of art and science. Therefore it is also a huge topic in itself. Good references in this context are recent books (Saad, 1996) and (Greenbaum, 1997). Some preconditioning techniques used for model reduction can be found in (Grimme *et al.*, 1995; Huhtanen. 1997).

4.2. Numerical Examples

Finally, we illustrate the previous analysis by computing two numerical examples. The first example is non-symmetric and the second one is symmetric. In both the examples projectors with the range equal (15) generated via the Arnoldi method with A yield too large matrices for constructing a reduced model for $\Sigma(A, B, C, D)$.

We generate projectors in two different ways. First, we use the Arnoldi method with A to produce W_k for $P_k = W_k W_k^*$. At each step the range of P_k equals (15). This projector construction was proposed in (Jaimoukha and Kasenally, 1994). Then we compare this with P_k in which W_k is obtained from the inverse Arnoldi iteration

$$\mathcal{K}_k\big((\lambda I - A)^{-1}; B\big) = \operatorname{span}\{B, (\lambda I - A)^{-1}B, \ldots, (\lambda I - A)^{-(l-1)}B\} \quad (17)$$

with a $\lambda \in \mathbb{C}$, i.e. the Block-Arnoldi method with $(\lambda I - A)^{-1}$. Unless there is information about the spectrum of A, we use $\lambda = 0$ which seems to be a good choice e.g. for symmetric systems. As to the amount of operations, we make an LU-factorization of $\lambda I - A$ and use it throughout the process. This amounts approximatively to $2n^3/3$ arithmethic operations in the beginning and $O(n^2)$ for each inverse iterate.
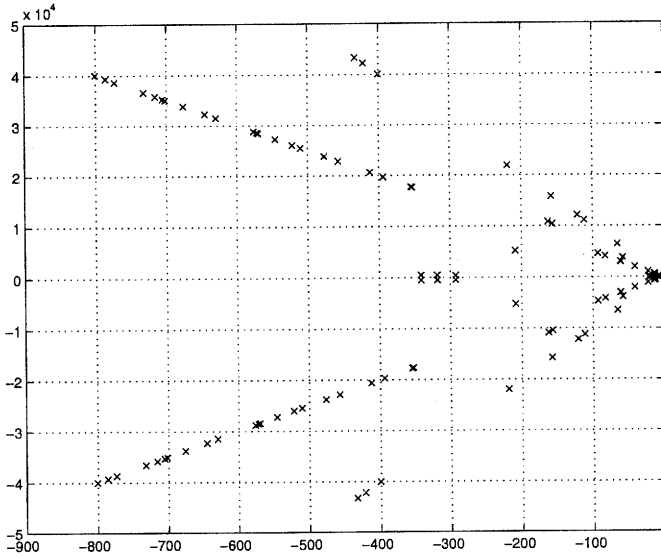
Fig. 1. The eigenvalues of Example 3, the matrix of the model for a CD-player regarding.

The computations were carried out with MATLAB and we used the ARE command to compute numerical solutions to the Lyapunov equations.

Example 3. In this example, A is obtained from a model for a CD-player (Wortel-boer, 1994) and it is of size 120×120. This matrix is also used in (Grimme *et al.*, 1995) as an example for which the Arnoldi and Lanczos methods with A yield un-stable reduced models practically for all steps, i.e. preconditioning for a reasonable reduced model is needed. In Fig. 1 we have plotted the eigenvalues of A and we emphasize that they spread wide open to the left half-plane (note the scales). In our computation we chose B to be a random vector of unit length. In Fig. 2 we have computed approximative solutions to the Lyapunov equation with two iterative methods. In the first case, W_k is constructed via the Arnoldi iteration with A (i.e. as in (Jaimoukha and Kasenally, 1994)) and in the second one we computed W_k via the inverse iteration at the origin, i.e. $\lambda = 0$.   ◆

Example 4. In this example, A is a diagonal matrix of size 100×100 with elements $-j^2\pi^2$ for $j = 1, \ldots, 100$ and B is a random vector. In Fig. 3 we have computed approximative solutions to the Lyapunov equation with two iterative methods. In the first case, W_k is constructed via the Arnoldi iteration with A (i.e. as in (Jaimoukha and Kasenally, 1994)) and in the second one we computed W_k via the inverse iteration at the origin, i.e. $\lambda = 0$. Note that A is self-adjoint.   ◆
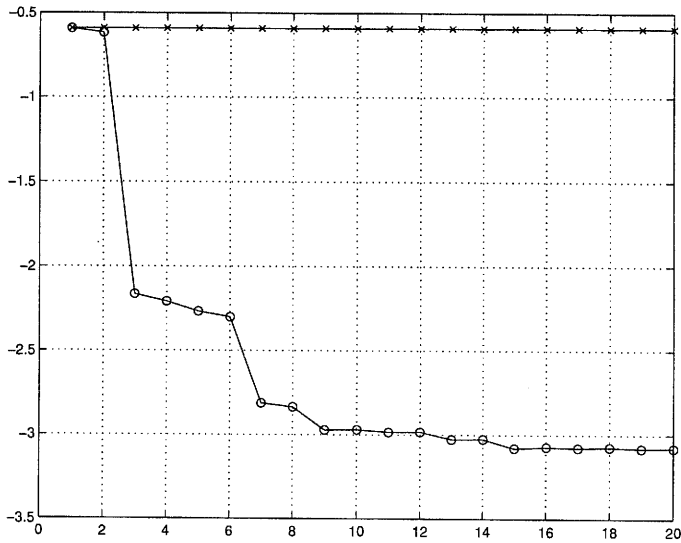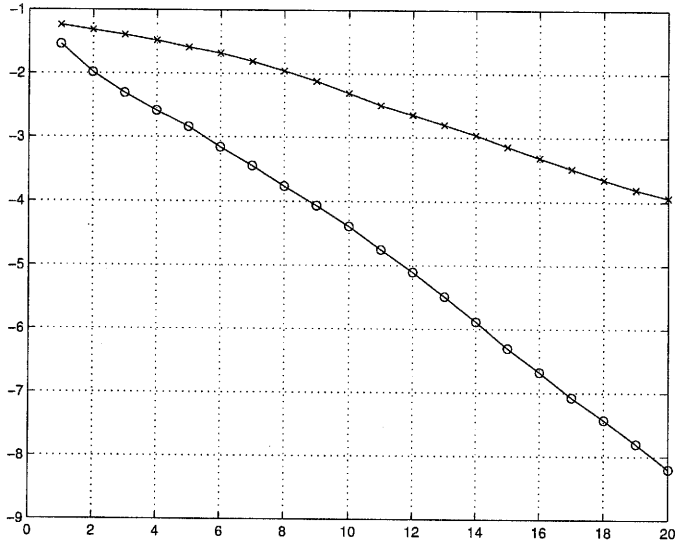
Fig. 2. The error in $\log_{10}$-scale of the Lyapunov equation of Example 3. Here the x-axis denotes the size of the approximation matrices (–×– denotes the Arnoldi method and – O – denotes the inverse iteration of (17)).

Fig. 3. The error in $\log_{10}$-scale of the Lyapunov equation of Example 4. Here the x-axis denotes the size of the approximation matrices (–×– denotes the Arnoldi method and – O – denotes the inverse iteration of (17)).

5. Conclusions

We have derived error bounds that can be used in reduction of the order of a large state-space system via iterative methods using the Hankel singular values as the criterion. The error terms show how the eigenvalues of A contribute to the error as a function of distance to the imaginary axis. Thereby the bounds explain why a simple Arnoldi iteration with A does not suffice while reducing the order of the system. The reason is that close to the imaginary axis the block-equation $(\lambda I - A)W_l X = B$ (or $(\lambda I - A^*)W_l X = C^*$) should be solved accurately. This, however, can be achieved if preconditioning or inverse iteration is used in the projector construction. Finally, we have given two realistic numerical examples showing that the Arnoldi iteration with A applied to B is not sufficient as the order of matrices used in model reduction can then become very large before being accurate enough.

Acknowledgment

References

Balas J. (1982): *Trends in large space structure control theory: Fondest hopes, wildest dreams.* — IEEE Trans. Automat. Contr., Vol.27, No.3, pp.522–535.

Boley D. and Golub G.H. (1991): *The nonsymmetric Lanczos-Arnoldi algorithm and controllability.* — Syst. Contr. Lett., Vol.16, No.2, pp.97–106.

Curtain R. and Zwart H. (1995): *An Introduction to Infinite-Dimensional Linear System Theory.* — New York: Springer-Verlag.

Enns D. (1984): *Model reduction with balanced realizations: An error bound and a frequency weighted geneeralization.* — Proc. 23-rd Conf. Contr., Las Vegas, NV, pp.127–132.

Glover K. (1984): *All optimal Hankel norm approximations of linear multivariable system and their L^∞-error bounds.* — Int. J. Contr., Vol.39, No.6, pp.1115–1193.

Greenbaum A. (1997): *Iterative Methods for Solving Linear Systems.* — Frontiers in Applied Mathematics, Philadelphia: SIAM.

Grimme E.J., Sorensen D.C. and Van Dooren P. (1995): *Model reduction of state space systems via an implicitly restarted Lanczos method.* — Num. Alg., Vol.12, No.1, pp.1–31.

Halmos P. (1968): *Quasitriangular operators.* — Acta Szeged., Vol.29, No.3–4, pp.283–293.

Hochbruck M. and Starke G. (1995): *Preconditioned Krylov subspace method for Lyapunov matrix equations.* — SIAM J. Matrix Anal. Appl., Vol.16, No.1, pp.156–171.

Hochbruck M. and Lubich C. (1997): *On Krylov approximations to the matrix exponential operator.* — SIAM J. Numer. Anal., Vol.34, No.5, pp.1911–1925.

Horn R. and Johnson C.R. (1991): *Topics in Matrix Analysis.* — Cambridge: Cambridge University Press.

Hu D.Y. and Reichel L. (1992): *Krylov subspace methods for the Sylvester equation.* — Lin. Alg. Appl., Vol.172, pp.283–313.

Huhtanen M. (1997): *Model reduction of state space systems via EN updating scheme.* — Proc. 4-th *European Control Conference*, Brussels, Belgium.

Huhtanen M. (1998): *On a resolvent condition in finite dimensional controller desing.* — Proc. MTNS, Padova, Italy.

Jaimoukha I. and Kasenally E. (1994): *Krylov subspace methods for solving large Lyapunov equations.* — SIAM J. Numer. Anal., Vol.31, No.1, pp.283–314.

Phillips J.R. (1998): *Model reduction of time-varying linear systems using approximate multipoint Krylov subspace projectors.* — Proc. Int. Conf. *Computer-Aided Design*, Santa Clara, CA, pp.96–102.

Phillips J.R., Chiprout E. and Ling D. (1996): *Efficient full-wave electromagnetic analysis via model-reduction of fast integral transforms.* — Proc. Design Automat. Conf., Las Vegas, NV, pp.377–382.

Knyazev A. (1998): *New estimates for Ritz vectors.* — Math. Comp., Vol.66, No.219, pp.985–995.

Moore B. (1981): *Principal component analysis in linear systems: controllability, observability and model reduction.* — IEEE Trans. Automat. Contr., Vol.AC-26, No.1, pp.17–31.

Nevanlinna O. (1995): *Hessenberg matrices in Krylov subspaces and the computation of the spectrum.* — Num. Funct. An. Opt., Vol.16, No.384, pp.443–473.

Nevanlinna O. and Vainikko G. (1996): *Limit spectrum of discretely converging operators.* — Num. Funct. An. Opt., Vol.17, No.788, pp.797–808.

Pernebo L. and Silverman L.M. (1982): *Model reduction via balanced state space representation.* — IEEE Trans. Automat. Contr., Vol.AC-27, No.2, pp.382–387.

Saad Y. (1990): *Numerical solution of large Lyapunov equations,* In: Signal Processing, Scattering, Operator Theory and Numerical Methods (M.A. Kaashoek, J.H. van Schuppen and A.C.M. Ran, Eds.). — Birkhäuser, Boston, pp.503–511.

Saad Y. (1992): *Analysis of some Krylov subspace approximations to the matrix exponential operator.* — SIAM J. Numer. Anal., Vol.29, No.1, pp.209–228.

Saad Y. (1996): *Iterative Methods for Sparse Linear Systems.* — Boston: PWS.

Trefethen L.N. (1992): *Pseudospectra of matrices,* In: Numerical Analysis 1991 (D.F. Griffiths and G.A. Watson, Eds.). — Harlow: Longman.

Wortelboer P. (1994): *Frequency-weighted Balanced Reduction of Closed-loop Mechanical Servo-systems: Theory and Tools.* — Ph.D. Th., Technical University of Delft.

Zwart H.J. (1989): *Geometric Theory for Infinite-Dimensional Systems.* — Lecture Notes in Control and Information Sciences, Vol.115, Berlin: Springer.