Adv Comput Math (2018) 44:1821–1844
https://doi.org/10.1007/s10444-018-9608-6

Balanced truncation model order reduction in limited time intervals for large systems

Patrick Kürschner[1]

Received: 7 July 2017 / Accepted: 16 April 2018 /
Published online: 5 June 2018

Abstract In this article we investigate model order reduction of large-scale systems using time-limited balanced truncation, which restricts the well known balanced truncation framework to prescribed finite time intervals. The main emphasis is on the efficient numerical realization of this model reduction approach in case of large system dimensions. We discuss numerical methods to deal with the resulting matrix exponential functions and Lyapunov equations which are solved for low-rank approximations. Our main tool for this purpose are rational Krylov subspace methods. We also discuss the eigenvalue decay and numerical rank of the solutions of the Lyapunov equations. These results, and also numerical experiments, will show that depending on the final time horizon, the numerical rank of the Lyapunov solutions in time-limited balanced truncation can be smaller compared to standard balanced truncation. In numerical experiments we test the approaches for computing low-rank factors of the involved Lyapunov solutions and illustrate that time-limited balanced truncation can generate reduced order models having a higher accuracy in the considered time region.

Keywords Lyapunov equation · Rational Krylov subspaces · Model order reduction · Balanced truncation · Matrix exponential

Mathematics Subject Classification 2010 15A16 · 15A18 · 15A24 · 65F60 · 93A15 · 93C

Communicated by: Peter Benner

✉ Patrick Kürschner
kuerschner@mpi-magdeburg.mpg.de

[1] Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, Magdeburg, 39106, Germany

1 Introduction

1.1 Model reduction of linear systems

Consider continuous-time, linear, time-invariant (LTI) systems

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = 0, \tag{1a}$$

$$y(t) = Cx(t) \tag{1b}$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{p\times n}$. Typically, the vector functions $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $y(t) \in \mathbb{R}^p$ are referred to as state, control, and, output vector, respectively. We assume that $m, p \ll n$ and A is Hurwitz, i.e., $\Lambda(A) \subset \mathbb{C}_-$, such that (1) is asymptotically stable. Given a large state space dimension n, model order reduction aims toward finding a reduced order model

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{B}u(t), \quad \tilde{x}(0) = 0, \tag{2a}$$

$$\tilde{y}(t) = \tilde{C}\tilde{x}(t) \tag{2b}$$

with $A \in \mathbb{R}^{r\times r}$, $B \in \mathbb{R}^{r\times m}$, $C \in \mathbb{R}^{p\times r}$, $\tilde{x}(t) \in \mathbb{R}^r$ and a drastically reduced state dimension $r \ll n$. The smaller reduced system (2) should approximate the input-output behavior or the original system (1) well. Moreover, simulating the system, i.e., solving the differential equations in (2) for many different control functions $u(t)$ should be numerically much less expensive compared to the original system (1). For the approximation of (1) regarding the actual time domain behavior, it is desired that for all feasible input functions $u(t)$,

$$\tilde{y}(t) \approx y(t) \quad \text{for} \quad t \geq 0, \tag{3a}$$

in other words, $\|y(t)-\tilde{y}(t)\|$ should be small $\forall t \geq 0$ in some norm $\|\cdot\|$. With the help of the Laplace transformation, one can also formulate the approximation problem in the frequency domain, e.g., via

$$\begin{aligned} \tilde{H}(\imath\omega) &\approx H(\imath\omega) \quad \text{for} \quad \omega \in \mathbb{R}, \quad \imath^2 = -1, \\ \text{where} \quad H(s) &= C(sI-A)^{-1}B, \quad \tilde{H}(s) = \tilde{C}(sI_r - \tilde{A})^{-1}\tilde{B} \end{aligned} \tag{3b}$$

are the transfer function matrices of (1) and (2). There exist different model order reduction technologies and here we focus on balanced truncation (BT) [36] model order reduction. The backbone of BT are the infinite controllability and observability Gramians of (1) which are the the symmetric, positive semidefinite solutions P_∞, Q_∞ of the continuous-time, algebraic Lyapunov equations

$$AP_\infty + P_\infty A^T = -BB^T, \quad A^T Q_\infty + Q_\infty A = -C^T C. \tag{4}$$

The Hankel singular values (HSV) of (1) are the square roots of the eigenvalues of the product $P_\infty Q_\infty$ and are system invariants under state spac transformations. The magnitude of the HSV enables to identify components that are weakly controllable and observable. In BT this is achieved by first transforming (1) into a balanced realization such that $P_\infty = Q_\infty = \Sigma_\infty = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$. Dropping all states corresponding to small σ_j gives the reduced order model. Solving (4) for the

```
Algorithm 1: Square-root balanced truncation with low-rank factors
Input  : System matrices $A$, $B$, $C$ defining an asymptotically stable dynamical
         system (1).
Output: Matrices $\tilde{A}$, $\tilde{B}$, $\tilde{C}$ of the reduced system.
Compute $Z_P,\ Z_Q$ (e.g., with the methods described in [9, 46]), such that
    $Z_P Z_P^T \approx P_\infty$, $Z_Q Z_Q^T \approx Q_\infty$ in (4).
Compute thin singular value decomposition
    $Z_Q^T Z_P = X \Sigma Y^T = \left[\, X_1 \;\; X_2 \,\right] \operatorname{diag}(\Sigma_1,\ \Sigma_2) \left[\, Y_1 \;\; Y_2 \,\right]^T$
    with $\Sigma_1 = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$ containing the largest $r$ (approximate) HSV.
Construct $T := Z_P Y_1 \Sigma_1^{-\frac{1}{2}}$ and $S := Z_Q X_1 \Sigma_1^{-\frac{1}{2}}$.
Generate reduced order model
    $\tilde{A} := S^T A T, \quad \tilde{B} := S^T B, \quad \tilde{C} := C T.$    (5)
```

Gramians $P_\infty,\ Q_\infty$ is the computationally most challenging part of balanced truncation. For large-scale systems one therefore uses low-rank approximations of the Gramians instead, e.g., $Z_P Z_P^T \approx P_\infty$ with low-rank solution factors $Z_P \in \mathbb{R}^{n \times k_P}$, $\operatorname{rank}(Z_P) = k_P \ll n$, and likewise for Q_∞. This strategy is backed up by the often numerically observed and theoretically explained rapid eigenvalue decay of solutions of Lyapunov equations [2, 3, 27, 37, 48]. The computation of the low-rank factors Z_P, Z_Q of P_∞, Q_∞ can be done efficiently by state-of-the-art numerical algorithms for solving large Lyapunov equations [9, 46]. BT using low-rank factors Z_P, Z_Q of the Gramians (4) is illustrated in Algorithm 1.

The $\mathcal{H}_\infty$-norm

$$\|H\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} (\|H(\imath\omega)\|_2).$$

of a stable system (1) is the L_2 induced norm of the convolution operator. It connects to the time-domain behavior via $\|y\|_{L_2} \le \|H\|_{\mathcal{H}_\infty} \|u\|_{L_2}$. With exact Gramian factors, i.e., $Z_P Z_P^T = P_\infty$, $Z_Q Z_Q^T = Q_\infty$, BT is known to always generate a asymptotically stable reduced system for which the error bound

$$\|H - \tilde{H}\|_{\mathcal{H}_\infty} \le 2 \sum_{j=r+1}^{n} \sigma_j \tag{6}$$

holds.

1.2 Model reduction in limited time- and frequency intervals

The approximation paradigms (3) enforce that the reduced system (2) is accurate for all times $t \in \mathbb{R}_+$ and frequencies $\omega \in \mathbb{R}$. From a practical point of view, achieving (3) might overshoot a realistic objective. For instance, if (1) models a mechanical or electrical system, practitioners working with this model (and its approximation (2)) might only be interested in certain finite frequency intervals $0 \leq \omega_1 < \omega_2 < \infty$. Likewise, when the final goal is to carry out simulations of (1), i.e., acquire time-domain solutions for various controls $u(t)$, one is usually only interested in $y(t)$ for t smaller than some final time $t_e < \infty$. Hence, we consider time- and frequency restricted versions of (3) of the form

$$\tilde{y}(t) \approx y(t) \quad \text{for} \quad t \in \mathcal{T} \subset \mathbb{R}_+, \tag{7a}$$

$$\tilde{H}(i\omega) \approx H(i\omega) \quad \text{for} \quad \omega \in \Omega \subset \mathbb{R},\ \Omega = -\Omega, \tag{7b}$$

where the time- and frequency regions $\mathcal{T}$, Ω of interest should be provided by the underlying application. The expressions (7) demand that the reduced order model (2) is only accurate in $\mathcal{T}$, Ω but allow larger errors outside these regions. Compared to MOR approaches for the unrestricted setting (3), it is desired that MOR approaches for (7) achieve smaller approximation errors in $\mathcal{T}$, Ω with the same reduced order r. Alternatively, one demands that time- and frequency restricted MOR leads to comparable approximation errors in $\mathcal{T}$, Ω with reduced systems of smaller order r. A secondary question is how the added time- and frequency restrictions influence the computational effort compared to an unrestricted MOR method of the same type. This issue will be in our particular focus. Typically, the time region in (7a) has the form

$$\mathcal{T} = [0, t_e], \quad t_e < \infty \tag{8a}$$

which will also be the main situation in this work, but the more general restriction

$$\mathcal{T} = [t_s, t_e],\ 0 < t_s < t_e < \infty \tag{8b}$$

will also be briefly discussed. Regarding the frequency restricted setting (7b), the typical regions are

$$\Omega := \hat{\Omega} \cup -\hat{\Omega},\ \hat{\Omega} := \bigcup_{i=1}^{h} [\omega_i, \omega_{i+1}] \quad \text{with} \quad 0 \leq \omega_1 < \ldots < \omega_h < \omega_{h+1} < \infty.$$

Introducing time- or frequency restrictions into balanced truncation MOR has been originally proposed in [25] and further studied in, e.g., [8, 18, 21, 29]. In certain applications, e.g., circuit design, only single frequencies $\omega \in \mathbb{R}$ might be of interest and an associated version of balanced truncation is addressed in [19]. $\mathcal{H}_2$-MOR with limitations or weights on the frequency and time domain is investigated in, e.g., [11, 26, 32, 38, 39, 47, 49]. As continuation of our work in [8] regarding frequency-limited BT, we consider in this paper the numerically efficient realization of time-limited balanced truncation (TLBT) for large-scale systems.

1.3 Overview of this article

We start in Section 2 by reviewing the general concept of time-limited BT from [25], mainly for restrictions of the form (8a). This includes the associated time-limited Gramians as well as the respective Lyapunov equations. Similar to standard BT, executing TLBT for large systems heavily relies on how well the time-limited Gramians can be approximated by low-rank factorizations. This issue is investigated in Section 3.1 where we have a particular interest in the question when t_e induces significant differences between the infinite and time-limited Gramians. The actual issue of numerically dealing with the arising matrix functions and computing the low-rank factors of the time-limited Gramians is topic of Sections 3.2 and 3.3, respectively. Motivated by the promising results for frequency-limited BT studied in [8], we again employ rational Krylov subspace methods for this task. Section 4 collects different generalizations of TLBT including general state-space and certain differential algebraic systems, the time restriction (8b), and stability preservation. In Section 5, the proposed numerical approach is tested numerically with respect to the approximation of the Gramians as well as the reduction of the dynamical system.

1.4 Notation

The real and complex numbers are denoted by, respectively, $\mathbb{R}$ and $\mathbb{C}$, $\mathbb{R}_-$, $(\mathbb{R}_+)$, $\mathbb{C}_-$ $(\mathbb{C}_+)$ are the sets of strictly negative (positive) real numbers and the open left (right) half plane. The space of real (complex) Matrices of dimension $n \times m$ is $\mathbb{R}^{n\times m}$ $(\mathbb{C}^{n\times m})$. For any complex quantity $X = \mathrm{Re}\,(X) + \imath\,\mathrm{Im}\,(X)$, we denote by $\mathrm{Re}\,(X)$ and $\mathrm{Im}\,(X)$ its real and, respectively, imaginary parts with $\imath$ being the imaginary unit, and its the complex conjugate is $\overline{X} = \mathrm{Re}\,(X) - \imath\,\mathrm{Im}\,(X)$. The absolute value of any real or complex scalar is denoted by $|z|$. Unless stated otherwise, $\|\cdot\|$ is the Euclidean vector- or subordinate matrix norm (spectral norm). By A^T and $A^H = \overline{A}^T$ we indicate the transpose and complex conjugate transpose of a real and complex matrix, respectively. If $A \in \mathbb{C}^{n\times n}$ is a nonsingular, its inverse is A^{-1}, and $A^{-H} = (A^H)^{-1}$. Expressions of the form $x = A^{-1}b$ are always to be understood as solving a linear system $Ax = b$ for x. The identity matrix of dimension n is indicated by I_n, and the vector of ones is denoted by $\mathbf{1}_m := (1,\ldots,1)^T \in \mathbb{R}^m$. The notation $A \succ 0$ $(\prec 0)$ indicates symmetric positive (negative) definiteness of a symmetric or Hermitian matrix A, $\succeq$ $(\preceq)$ refers to semi-definiteness, and $A \succeq (\preceq) B$ means $A - B \succeq (\preceq) 0$. The spectrum of a matrix A is denoted by $\Lambda(A)$.

2 Gramians and balanced truncation for finite time horizons

Since A is assumed to be Hurwitz, the infinite Gramians P_∞, Q_∞ in (4) can be represented in integral form as

$$P_\infty = \int_0^\infty \mathrm{e}^{At}\, BB^T\, \mathrm{e}^{A^T t}\,\mathrm{d}t, \quad Q_\infty = \int_0^\infty \mathrm{e}^{A^T t}\, C^T C\, \mathrm{e}^{At}\,\mathrm{d}t. \tag{9}$$

Restricting the integration limits in the integrals (9) to a time interval $[t_s, t_e]$ immediately yields the *time-limited Gramians* $P_{\mathcal{T}}$, $Q_{\mathcal{T}}$.

Definition 1 (**Time-limited Gramians** [25]) The time-limited reachability and observability Gramians of (1) with respect to the time-interval $\mathcal{T} = [t_s, t_e],\ 0 \le t_e < t_e < \infty$ are defined by

$$P_{\mathcal{T}} = \int_{t_s}^{t_e} \mathrm{e}^{At}\, BB^T\, \mathrm{e}^{A^T t}\, \mathrm{d}t, \quad Q_{\mathcal{T}} = \int_{t_s}^{t_e} \mathrm{e}^{\mathcal{T}}\, C^T C\, \mathrm{e}^{At}\, \mathrm{d}t. \tag{10}$$

Theorem 1 (**Lyapunov equations for the time-limited Gramians** [25]) *The time-limited Gramians $P_{\mathcal{T}}$ and $Q_{\mathcal{T}}$ according to Definition 1 are equivalently given in the following ways.*

a) *The finite time Gramians $P_{\mathcal{T}},\ Q_{\mathcal{T}}$ from* (10) *are given by*

$$P_{\mathcal{T}} = \mathrm{e}^{At_s}\, P_\infty\, \mathrm{e}^{A^T t_s} - \mathrm{e}^{At_e}\, P_\infty\, \mathrm{e}^{A^T t_e}, \tag{11a}$$

$$Q_{\mathcal{T}} = \mathrm{e}^{A^T t_s}\, Q_\infty\, \mathrm{e}^{At_s} - \mathrm{e}^{A^T t_e}\, Q_\infty\, \mathrm{e}^{At_e}, \tag{11b}$$

where P_∞ and Q_∞ are the infinite reachability and observability Gramians (4).

b) *The time-limited Gramians $P_{\mathcal{T}},\ Q_{\mathcal{T}}$ satisfy the time-limited reachability and observability Lyapunov equations*

$$AP_{\mathcal{T}} + P_{\mathcal{T}}A^T = -B_{t_s}B_{t_s}^T + B_{t_e}B_{t_e}^T, \quad B_{t_s} := \mathrm{e}^{At_s}\, B,\ B_{t_e} := \mathrm{e}^{At_e}\, B \tag{12a}$$

$$A^T Q_{\mathcal{T}} + Q_{\mathcal{T}}A = -C_{t_s}^T C_{t_s} + C_{t_e}^T C_{t_e}, \quad C_{t_s} := C\, \mathrm{e}^{At_s}, C_{t_e} := C\, \mathrm{e}^{At_e}. \tag{12b}$$

Note that the time-limited Gramians (10) also exist for unstable A and if $\Lambda(A) \cap \Lambda(-A) = \emptyset$ they still solve the Lyapunov (12), see [40]. Hence, TLBT might be one possible approach to reduce unstable systems. Except for one brief numerical experiment, we will not pursue this topic any further. In analogy to the infinite time horizon case, the square roots of the eigenvalues of the product $P_{\mathcal{T}}Q_{\mathcal{T}}$ are called *time-limited Hankel singular values*. A basic calculation shows that the time-limited Hankel singular values are, as the standard Hankel singular values, invariant with respect to state-space transformations. By comparing the Lyapunov equations for the infinite Gramians (4) with (12), one immediately sees that the only differences are the inhomogeneities, while the left hand sides are the same unchanged (adjoint) Lyapunov operators. This raises the question how much different the time-limited Gramians are from the infinite ones, and how this depends on the time interval of interest. We will pursue this in the next section, especially regarding the numerically important issue of how well the time-limited Gramians can be approximated by low-rank factorizations $P_{\mathcal{T}} \approx Z_{P_{\mathcal{T}}}Z_{P_{\mathcal{T}}}^T$, $Q_{\mathcal{T}} \approx Z_{Q_{\mathcal{T}}}Z_{Q_{\mathcal{T}}}^T$. Of course, before approximately solving the time-limited Lyapunov equations, the actions of the matrix exponentials e^{At_i} to B (as well as $\mathrm{e}^{A^T t_i}$ to C^T) have to be dealt with numerically. Numerical approaches for handling the matrix exponential and computing low-rank solution factors $Z_{P_{\mathcal{T}}},\ Z_{Q_{\mathcal{T}}}$ are topic of Section 3.

A square-root version of TLBT is simply carried out by substituting Step 1 of Algorithm 1 by the code snippet shown in Algorithm 2 below and using the low-rank solution factors $Z_{P_\mathcal{T}}$, $Z_{Q_\mathcal{T}}$ in the remaining steps. Depending on the time region of interest, TLBT might in general not be a stability preserving method and, thus, there is also no $\mathcal{H}_\infty$-error bound similar to (6). This can be regained by modifying TLBT further [29] and we discuss this issue in Section 4. Without this modification it is possible to establish an error bound in the $\mathcal{H}_2$-norm [40].

```
Algorithm 2: Required changes in Algorithm 1 for TLBT
1a Compute (approximation of) $B_{t_i} := \mathrm{e}^{At_i}\, B,\ C_{t_i} = C\,\mathrm{e}^{A^T t_i},\ i \in \{s, e\}$.
1b Compute $Z_{P_\mathcal{T}},\ Z_{Q_\mathcal{T}}$, such that $P_\mathcal{T} \approx Z_{P_\mathcal{T}} Z_{P_\mathcal{T}}^T,\ Q_\mathcal{T} \approx Z_{Q_\mathcal{T}} Z_{Q_\mathcal{T}}^T$ in (12).
```

3 Numerical computation of low-rank factors of time-limited Gramians

This section is concerned with the actual numerical computation of low-rank factors of the time-limited Gramians. Before algorithms for dealing with the matrix exponentials and the Lyapunov equations are discussed, we briefly investigate the numerical ranks of the time-limited Gramians. For the sake of brevity, all considerations will be mostly restricted to the reachability Gramian because the observability Gramian can be dealt with similarly by replacing A, B with A^T, C^T, respectively. Moreover, only the situation (8a) will be discussed, i.e., $t_s = 0$ and $0 < t_e < \infty$.

3.1 The difference of the infinite and time-limited Gramians

In order to approximate $P_\mathcal{T}$, $Q_\mathcal{T}$ by low-rank factorizations, it is desirable that their eigenvalues decay rapidly. For investigating this decay we assume from now that eigenvalues of symmetric (positive definite) matrices are given in a non-increasing order $\lambda_1 \geq \ldots \geq \lambda_n$. The inhomogeneities of the time-limited Lyapunov (12) have up to twice the rank of their unlimited counterparts (4). Hence, by the available theory on the eigenvalue decay of solutions of matrix equations [2, 27, 37, 45] one expects that the eigenvalues of the time-limited Gramians decay somewhat slower than those of the infinite ones. We will see in the numerical experiments that, similar to the frequency-limited Gramians, in most situations it is the opposite case: the eigenvalues of the time-limited Gramians exhibit a faster decay and, consequently, have smaller numerical ranks. Assuming that (A, B) is controllable ($\mathrm{rank}[A - \lambda I, B] = n,\ \forall \lambda \in \mathbb{C}$) it holds $P_\infty \succ 0$ and, using (10), the relation (11a) yields $0 \preceq P_\mathcal{T} = P_\infty - \mathrm{e}^{At_e}\, P_\infty\, \mathrm{e}^{A^T t_e}$. Since $E(t_e) := \mathrm{e}^{At_e}\, P_\infty\, \mathrm{e}^{A^T t_e} \succ 0$ it holds

$$P_\infty \succeq P_\mathcal{T} \quad \text{and also} \quad \lambda_1(P_\infty) = \|P_\infty\|_2 \succeq \|P_\mathcal{T}\|_2 = \lambda_1(P_\mathcal{T}).$$

Due to the stability of A, the difference $E(t_e) = P_\infty - P_\mathcal{T}$ will decay for increasing values of t_e. Tight bounds for the eigenvalue behavior of $P_\mathcal{T}$ and $E(t_e)$ with respect to the parameter t_e are difficult to derive and is a research topic of its own. Here, for simplicity we restrict the discussion to the case $m = 1$ and present a basic investigation of how the decay of $E(t_e)$ depends on t_e.

Lemma 1 *Let $B \in \mathbb{R}^n$, (A, B) controllable, A be diagonalizable, i.e., $A = X\Lambda X^{-1}$, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, and define $w := X^{-1}B$, $X_B := X\operatorname{diag}(w)$, $N(t_e) := X\,\mathrm{e}^{\Lambda t_e}\operatorname{diag}(w)$. Then $E(t_e) = \mathrm{e}^{At_e}\, P_\infty\, \mathrm{e}^{A^T t_e} = N(t_e)\mathcal{C}N(t_e)^H$,*

$$P_{\mathcal{T}} = P_\infty - E(t_e) = P_\infty - N(t_e)\mathcal{C}N(t_e)^H = X_B\left(\mathcal{C} - \mathrm{e}^{\Lambda t_e}\,\mathcal{C}\,\mathrm{e}^{\Lambda^H t_e}\right)X_B^H, \qquad (13)$$

and $\mathcal{C} := \left(\frac{-1}{\lambda_i + \overline{\lambda_j}}\right)_{i,j=1}^{n}$ is a Hermitian positive definite Cauchy matrix.

Proof Apply the eigendecomposition of A and $P_\infty = X_B\mathcal{C}X_B^H$ from [2, Lemma 3.2] to (11a). □

Consider the impulse response of (1),

$$y_\delta(t) = C\eta(t), \quad \eta(t) := \mathrm{e}^{At}\,B = X\,\mathrm{e}^{\Lambda t}\,w = N(t)\mathbf{1}_n,$$

indicating that the *impulse-to-state-map* $\eta(t)$ and $N(t)$ decay at a similar rate. With the *spectral abscissa* $R := \max\limits_{\lambda \in \Lambda(A)} \operatorname{Re}(\lambda)$, the basic point wise bounds

$$\|\eta(t)\| \le \mathrm{e}^{Rt}\,\|X\|\|w\|, \quad \|N(t)\| \le \mathrm{e}^{Rt}\,\|X\|\|w\|_\infty$$

make this more visible. Using this and (13), we can conclude that for increasing t, $E(t)$ is getting smaller at a similar speed as $\eta(t)$. Consequently, a significant difference between $P_{\mathcal{T}}$ and P_∞ might be only observed when t_e is chosen small enough in the sense that $\eta(t_e)$ has not reached an almost stationary state. The handling of the case $m > 1$ can be carried out similarly. Moreover, with a similar argumentation the decay rate of the time-limited Hankel singular values can be roughly connected to the decay of $y_\delta(t)$. To conclude, TLBT might be only practicable for small time horizons or for weakly damped systems. A similar investigation regarding TLBT for unstable systems is given in [47].

3.2 Approximating the products with the matrix exponential

There are several numerical approaches available to approximate the action of the matrix exponential to (a couple of) vectors, see, e.g., [1, 6, 12–15, 23, 24, 30, 31, 34, 35, 44]. Here, we are mostly interested in projection methods using (block) rational Krylov subspaces of the form

$$\mathcal{RK}_k = \operatorname{span}\{q_1, \ldots, q_k\}, \quad q_k = \left[\prod_{j=1}^{k}(A - s_j I)^{-1}\right]B \qquad (14)$$

where $s_k \in \mathbb{C}_+ \cup \imath\mathbb{R} \cup \{\infty\}$ are called shifts. They represent the poles of a rational approximation $r_k = \psi_{k-1}/\phi_{k-1}$ of e^z with polynomials ψ_{k-1}, ϕ_{k-1} of degree at most $k-1$.

Let $Q_k \in \mathbb{R}^{n\times km}$ have orthonormal columns and range $(Q_k) = \mathcal{RK}_k$. Then a Galerkin approximation [44] of $\mathrm{e}^{At_e}\, B$ takes the form

$$\begin{aligned} B_{t_e} \approx B_{t_e,k} &:= Q_k \hat{B}_{t_e,k}, \quad \hat{B}_{t_e,k} := \mathrm{e}^{H_k t_e}\, B_k,\\ H_k &:= Q_k^T A Q_k, \quad B_k := Q_k^T B. \end{aligned} \tag{15}$$

Note that for $m = 1$, the rational function r_k underlying the rational Krylov approximation $Q_k \left(\mathrm{e}^{H_k t_e}\right) Q_k^T B$ interpolates $f(z) = \mathrm{e}^{zt}$ at $\Lambda(H_k)$ (rational Ritz values) [31, Theorem 3.3]. Further information on the approximation properties can be found in, e.g., [6, 10, 14–16, 30]. In the remainder we assume $s_1 = \infty$ s.t. $q_1 = B$, range $(B) \subseteq$ range (Q_k) and $B_k = [\beta^T, 0]^T \in \mathbb{R}^{km\times m}$, where $\beta \in \mathbb{R}^{m\times m}$. The orthonormal basis matrix Q_k and the restriction H_k can be efficiently computed by a (block) rational Arnoldi process [41]. Since H_k is low-dimensional, $\mathrm{e}^{H_k t_e}$ can be computed by standard dense methods for the matrix exponential [34]. The choice of shifts s_k $(k > 1)$ is crucial for a rapid convergence and several strategies exists for this purpose [31]. In this work we exclusively use adaptive shift generation techniques [14–16] because this appeared to be the safest strategy in the majority of experiments. For the situation $m = 1$ and after step k of the rational Arnoldi process, the next shift s_{k+1} is selected via

$$s_{k+1} = \underset{\partial S_k}{\operatorname{argmax}}\, |r_k(s)|, \quad r_k(s) = \prod_{j=1}^{k} \tfrac{s-z_j}{s-s_j}, \tag{16}$$

where z_j are the eigenvalues of H_k (Ritz values of A); s_j, $j = 1, \ldots, k$ are the previously used shifts; and ∂S_k marks a set of discrete points from the boundary of S_k approximating $\Lambda(A)$. We follow [16] and use S_k as the convex hull of $\Lambda(H_k)$. In the symmetric case, one can also use the spectral interval given by the largest and smallest eigenvalue of A [14, 15, 31]. For $m > 1$ we simply use each s_j in the denominator of r_k in (16) m times as in [16]. A different technique to deal with $m > 1$ includes, e.g., tangential rational Krylov methods [17]. The rational Krylov subspace method was chosen not only because of the good approximation properties, but also because the generated subspace is also a good candidate to acquire low-rank solution factors of (12).

3.3 Computing the low-rank Gramian factors

Using (14), (15), a Galerkin approximation of the time-limited Gramian is $P_{\mathcal{T},k} = Q_k Y_k Q_k^T \approx P_{\mathcal{T}}$, where Y_k solves the projected time-limited Lyapunov equation

$$H_k Y_k + Y_k H_k^T = -B_k B_k^T + \hat{B}_{t_e,k} \hat{B}_{t_e,k}^T, \tag{17}$$

(cf. [43]). It holds $B_k = \hat{B}_{0,k} = e^{H_k t_s} B_k$ for the special case $t_s = 0$ considered here, indicating already how $t_s \neq 0$ can be included as well. Hence, after $B_{t_e,k}$ of sufficient accuracy is found, we follow the idea in [8] by recycling the rational Krylov basis and continuing the rational Arnoldi process for (12a) until $P_{\mathcal{T},k}$ leads to a sufficiently

Algorithm 3: Rational Krylov subspace method for time-limited Lyapunov (12a)

```
Input  : $A,\ B,\ t_e$ as in (12a), tolerances $0 < \tau_f,\ \tau_P \ll 1$.
Output: $Z_k \in \mathbb{R}^{n\times\ell}$ such that $Z_kZ_k^T \approx P_{\mathcal{T}}$ with $\ell \leq mk \ll n$.
1  $B = q_1\beta$ s.t. $q_1^Tq_1 = I_m$, $Q_1 = q_1$.
2  for $k = 1, 2, \ldots$ do
3      Get new shift $s_{k+1}$ via, e.g., (16).
4      Solve $(A - s_{k+1}I)g = q_k$ for $g$.
5      Real, orthogonal expansion of $Q_k$: $g_+ = \mathrm{Re}\ (g) - Q_k(Q_k^T\ \mathrm{Re}\ (g))$.
6      $q_{k+1} = g_+\beta_k$ s.t. $q_{k+1}^Tq_{k+1} = I_m$, $Q_{k+1} = [Q_k,\ q_{k+1}]$.
7      if $\mathrm{Im}\ (s_{k+1}) \neq 0$ then
8          $k = k + 1$, $g_+ = \mathrm{Im}\ (g) - Q_k(Q_k^T\ \mathrm{Im}\ (g))$
9          $q_{j+1} = g_+\beta_k$ s.t. $q_{k+1}^Tq_{k+1} = I_m$, $Q_{k+1} = [Q_k,\ q_{k+1}]$.
10     $H_k = Q_k^TAQ_k$, $B_k = Q_k^TB$.
11     $\hat{B}_{t_e,k} = \mathrm{e}^{H_kt_e}\ B_k$, $B_{t_e,k} = Q_k\hat{B}_{t_e,k}$
12     if $\|B_{t_e,k} - B_{t_e,k-1}\|/\|B_{t_e,k}\| < \tau_f$ then
13         Solve $H_kY_k + Y_kH_k^T + B_kB_k^T - \hat{B}_{t_e,k}\hat{B}_{t_e,k}^T = 0$ for $Y_k$.
14         Set $\mu_k := \frac{\|A(Q_kY_kQ_k^T)+(Q_kY_kQ_k^T)A^T+BB^T-B_{t_e,k}B_{t_e,k}^T\|}{\|B_kB_k^T-\hat{B}_{t_e,k}\hat{B}_{t_e,k}^T\|}$.
15         if $\mu_k < \tau_P$ then
16             $Y_k = S\Gamma S^T$, $S^TS = I_{km}$, $\Gamma = \mathrm{diag}\ (\gamma_1, \ldots, \gamma_{km})$.
17             Truncate if necessary: $\Gamma = \mathrm{diag}\ (\gamma_1, \ldots, \gamma_\ell)$, $S = S(:, 1 : \ell)$,
               $\ell \leq mk$.
18             return low-rank solution factor $Z_k = Q_kS\Gamma^{\frac{1}{2}}$.
```

small norm of the Lyapunov residual. For the Lyapunov stage, the same adaptive shift generation (16) can be used [16]. The rational Krylov subspace method for generating a low-rank approximation $Z_kZ_k^T \approx P_{\mathcal{T}}$ is illustrated in Algorithm 3.

In the presence of a complex shift s_{k+1}, it is implicitly assumed that $\overline{s_{k+1}}$ is the subsequent shift. For this situation, the orthogonal expansion in Steps 5–9 of the already computed basis matrix Q_k by real basis vectors goes back to [42]. The projected matrix H_k in Step 10, as well as the Lyapunov residual norm in Step 14 can be computed without accessing the large matrix A, see, e.g., [10, 16, 31, 41]. The small Lyapunov equation in Step 13 can be solved by standard methods for dense matrix equations, e.g., the Bartels-Stewart [4] method which we employ here. Once the scaled Lyapunov residual norm falls below the desired threshold, the rational Arnoldi process is terminated and the Steps 16–18 bring the computed low-rank Gramian approximation in the desired form $Z_kZ_k^T$ and allow a rank truncation. Note that typically, once the approximation of B_{t_e} is found, the generated subspace is already good enough to acquire a low-rank Gramian approximation without many additional iterations of the rational Arnoldi process. Often, the criteria in Steps 12–15 are satisfied in the same iteration step.

4 Extensions and further problems

4.1 Multiple time values

Computing low-rank factors of the time-limited Gramians (12) for the more general approximation setting (8b) with a nonzero start time t_s, i.e., $\mathcal{T} = [t_s, t_e]$, can also be done using Algorithm 3 with minor adjustments. Having computed a rational Krylov basis for approximating $\mathrm{e}^{At_e} B$ for some time value t_e, the same basis typically also provides good approximations for any other time values [31]. Consequently, Algorithm 3 has to be simply changed by adding $\hat{B}_{t_s,k} = \mathrm{e}^{H_k t_s} B_k$, $B_{t_s,k} = Q_k \hat{B}_{t_s,k}$ and appropriately adjusting the steps regarding the Gramian approximation. Also the methods relying on Taylor approximations [1] can be easily modified to handle multiple time values. However, even if a nonzero t_s does not yield additional computational complications, this does not imply that TLBT will produce a accurate approximation of the transient behavior of (1) in $[t_s, t_e]$. Already the original TLBT paper [25] states that TLBT in $[t_s, t_e]$ is expected to only give good approximations of the impulse response ($u(t) = v\delta(t)$, $v \in \mathbb{R}^m$) $y_\delta(t) = C\,\mathrm{e}^{At} Bv$ and numerical experiments confirm this. For an intuitive explanation assume for simplicity that no truncation is done in TLBT, i.e. in Step 2 of Algorithm 1. Then it holds $\mathrm{range}\,(Q_k) = \mathrm{range}\,(T)$, $\mathrm{e}^{At_s} B \approx B_{k,t_s} \in \mathrm{range}\,(Q_k)$, $\mathrm{e}^{At_e} B \approx B_{k,t_e} \in \mathrm{range}\,(Q_k)$, and likewise for $C\,\mathrm{e}^{At_s}$, $C\,\mathrm{e}^{At_s}$ and $\mathrm{range}\,(S)$. Hence, the impulse response is accurately approximated at the relevant times.

For the response of an arbitrary input $u(t)$, $y_u(t) = C \int\limits_0^t \mathrm{e}^{A(t-\tau)} Bu(\tau)\mathrm{d}\tau$, such argumentation clearly does not automatically hold. The value of $y_u(t)$ with respect to a general $u(t)$ depends, in general, on the values at the times before t. In the present form TLBT restricts the approximation to the time frame $\mathcal{T}$ and, thus, $y_u(t)$ will be poorly approximated for $t \leq t_s$ which, in turn, makes it difficult to acquire good approximations in $\mathcal{T}$.

One approach is to apply a time translation to the underlying system (1) such that the requested time-interval is transformed to $[0, t_e - t_s]$. However, this time translation will also introduce an inhomogeneous initial value x_0, which is an additional difficulty for model order reduction. Some strategies to cope with nonzero initial values are given in [5, 33] and we plan to investigate the incorporation to time-limited BT in the future.

4.2 Generalized state-space and index-one descriptor systems

Consider generalized state-space systems

$$M\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = 0, \tag{18a}$$
$$y(t) = Cx(t) \tag{18b}$$

with $M \in \mathbb{R}^{n\times n}$ nonsingular. Simple manipulations reveal that, similar to unrestricted [7] and frequency-limited BT [8], the time-limited Gramians of (18) are $P_{\mathcal{T}}$,

$M^T Q_{\mathcal{T}} M$, where $P_{\mathcal{T}}$, $Q_{\mathcal{T}}$ solve the time-limited generalized Lyapunov equations

$$AP_{\mathcal{T}}M^T + MP_{\mathcal{T}}A^T = -B_{t_s}B_{t_s}^T + B_{t_e}B_{t_e}^T, \; B_{t_i} := M\,\mathrm{e}^{M^{-1}At_i}\,M^{-1}B, \tag{19a}$$

$$A^T Q_{\mathcal{T}}M + M^T Q_{\mathcal{T}}A = -C_{t_s}^T C_{t_s} + C_{t_e}^T C_{t_e}, \; C_{t_i} := C\,\mathrm{e}^{M^{-1}At_i} \tag{19b}$$

for $t_i = t_s, t_e$. Note that $B_{t_i} := \mathrm{e}^{AM^{-1}t_i}\,B$. Low-rank factors of $P_{\mathcal{T}}$, $Q_{\mathcal{T}}$ can be computed by methods for generalized Lyapunov equations. In particular, the rational Krylov subspace methods which we employ for approximating B_{t_i} will implicitly work on $M^{-1}A$, $M^{-1}B$ or alternatively, if $M = LL^T \succ 0$, on $L^{-1}AL^{-T}$, $L^{-1}B$. With low-rank approximations $P_{\mathcal{T}} \approx Z_{P_{\mathcal{T}}}Z_{P_{\mathcal{T}}}^T$, $Q_{\mathcal{T}} \approx Z_{Q_{\mathcal{T}}}Z_{Q_{\mathcal{T}}}^T$, the SVD of $Z_{Q_{\mathcal{T}}}^T M Z_{P_{\mathcal{T}}}$ has to be used in Step 2 of Algorithm 1.

In some of the numerical experiments we will encounter the situation $M = \left[\begin{smallmatrix} M_1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, $A = \left[\begin{smallmatrix} A_1 & A_2 \\ A_3 & A_4 \end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix} B_1 \\ B_2 \end{smallmatrix}\right]$, $C = [\,C_1\; C_2\,]$ with $M_1 \in \mathbb{R}^{n_f \times n_f}$, $A_4 \in \mathbb{R}^{n-n_f \times n-n_f}$ nonsingular, s.t. (18) becomes a semi-explicit index-one descriptor system. Eliminating the algebraic constraints leads to a general state-space system defined by M_1, $\hat{A} := A_1 - A_2A_4^{-1}A_3$ and $\hat{B} := B_1 - A_2A_4^{-1}B_2$, $\hat{C} := C_1 - C_2A_4^{-1}A_3$, and an additional feed-through term $-C_2A_4^{-1}B_2$, see [22]. Time-limited and unrestricted BT can be applied right away to this system via (19) defined by M_1, $\hat{A}$, $\hat{B}$, $\hat{C}$. However, the matrix $\hat{A}$ will in general be dense and, thus, solving the linear systems in Algorithm 3 can be very expensive. In the context of unrestricted BT, the authors of [22] exploit in a LR-ADI iteration for the Gramians that the arising dense linear systems $(\hat{A} - sM_1)\hat{V} = \hat{W}$ are equivalent to the sparse linear systems $(A - sM)V = W$ with $V = [\hat{V}^T, \Psi]^T$, $W = [\hat{W}^T, 0]^T$ which are easier to solve numerically. We use the same trick within Step 4 of Algorithm 3.

4.3 Stability preservation and modified TLBT

Because of the altered and sometimes indefinite right hand sides of (12), (19), TLBT is in general not stability preserving. Only when the used time interval is long enough such that the right hand sides are negative semi-definite, TLBT will produce an asymptotically stable reduced order model [25, Condition 1]. For the general situation, a stability preserving modification of TLBT is proposed in [29] using the Lyapunov equations

$$\begin{aligned} AP_{\mathcal{T}}^{\mathrm{mod}}M^T + MP_{\mathcal{T}}^{\mathrm{mod}}A^T &= -B_{\mathrm{mod}}B_{\mathrm{mod}}^T, \\ A^T Q_{\mathcal{T}}^{\mathrm{mod}}M + M^T Q_{\mathcal{T}}^{\mathrm{mod}}A &= -C_{\mathrm{mod}}^T C_{\mathrm{mod}}, \end{aligned} \tag{20}$$

$$B_{\mathrm{mod}} := Q_B\,\mathrm{diag}\left(|\lambda_1^B|, \ldots, |\lambda_{r_B}^B|\right)^{\frac{1}{2}}, \quad C_{\mathrm{mod}} := \mathrm{diag}\left(|\lambda_1^C|, \ldots, |\lambda_{r_C}^C|\right)^{\frac{1}{2}} Q_C^T.$$

where $Q_B \in \mathbb{R}^{n \times r_B}$, $Q_C \in \mathbb{R}^{n \times r_C}$ contain the eigenvectors corresponding to the $r_B \le 2m$, $r_C \le 2p$ nonzero eigenvalues λ_i^B, λ_i^C of the right hand sides of (12a), (19a) and, respectively, (12b), (19b). The rational Krylov subspace method in Algorithm 3 for $M = I$, $t_s = 0$ can be easily modified for (20) by replacing Step 13 with the steps shown in Algorithm 4.

```
Algorithm 4: Changes in Algorithm 3 for modified time-limited Gramians
13a Compute partial eigendecomposition
    $\left(\begin{bmatrix} \beta\beta^T & 0 \\ 0 & 0 \end{bmatrix} - \hat{B}_{t_e,k}\hat{B}_{t_e,k}^T\right) Q_B = Q_B \operatorname{diag}\left(\lambda_1^B, \ldots, \lambda_{r_B}^B\right),\ Q_B^T Q_B = I_{r_B},\ \lambda_i^B \neq 0.$
13b Factor of projected modified rhs $B_{\mathrm{mod},k} := Q_B \operatorname{diag}\left(|\lambda_1^B|, \ldots, |\lambda_{r_B}^B|\right)^{\frac{1}{2}}.$
13c Solve $H_k Y_k + Y_k H_k^T + B_{\mathrm{mod},k} B_{\mathrm{mod},k}^T$ for $Y_k$.
```

Generalization for $M \neq I$ and $0 < t_s < t_e$ are straightforward. Note that because the modified time-limited Gramians do not fulfill a relation of the form (10) or (11), we cannot expect a fast singular value decay similar to $P_{\mathcal{T}},\ Q_{\mathcal{T}}$. As observed in [8, 29] for modified frequency-limited BT, modified TLBT might also lead to less accurate approximations in the considered time region compared to unmodified TLBT.

5 Numerical experiments

Here, we illustrate numerically the results of Section 3 regarding the numerical rank of the frequency-limited Gramians as well as the numerical method for computing low-rank factors of $P_{\mathcal{T}},\ Q_{\mathcal{T}}$. Afterwards, the quality of the approximations obtained by TLBT with the low-rank factors is evaluated and compared against unlimited BT. Further topics like nonzero starting times and the modified TLBT scheme are also examined along the way. All experiments are done in MATLAB ® 8.0.0.783 (R2012b) on a Intel®Xeon®CPU X5650 @ 2.67GHz with 48 GB RAM. Table 1 list the used examples and their properties. For time-domain simulation of (1), (18) and the reduced order models for a given input function $u(t)$, an implicit midpoint rule with a fixed small time step Δt is used. Because the impulse response of (1) for $u(t) = \delta(t)v,\ v \in \mathbb{R}^m,\ x_0 = 0$ can be expressed as $y_\delta(t) = C\,\mathrm{e}^{At}\,Bv$, it is computed by the same integrator applied to the uncontrolled system ($u(t) = 0$) with initial condition $x_0 = Bv$. For the impulse, or step responses, the vector distributing the control to the columns of B is set to $v = \mathbf{1}_m$.

Table 1 Dimensions, final integration time t_f, step size Δt, matrix properties, and source of the test systems

Example	n	m	p	t_f	Δt	properties	source
`bips_606`	7135	4	4	20	0.04	index-1, $n_f = 606$, $M_1 = I_{n_f}$	morwiki[a], [22]
`bips_3078`	21128	4	4	20	0.04	index-1, $n_f = 3078$, $M_1 = I_{n_f}$	morwiki[a], [22]
`vertstand`	16626	6	6	600	0.6	$A \prec 0$, $M \succ 0$, C random	morwiki[a], [28]
`rail`	79841	7	6	400	0.4	$A \prec 0$, $M \succ 0$	Oberwolfach Collection[b], ID=38881

[a] https://morwiki.mpi-magdeburg.mpg.de/morwiki

[b] http://portal.uni-freiburg.de/imteksimulation/downloads/benchmark

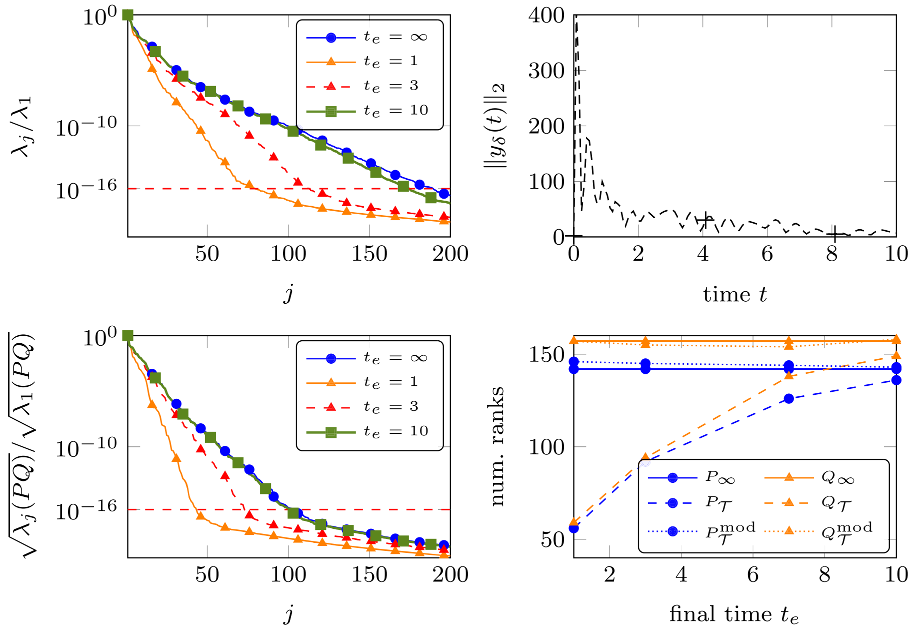

Fig. 1 Scaled eigenvalues of $P_\mathcal{T}$ (top left), Hankel and time-limited Hankel singular values $\sqrt{\lambda(P_\mathcal{T} Q_\mathcal{T})}$ (bottom left), norm of impulse response $y_\delta(t)$ (top right), and numerical ranks of infinite, (modified) time-limited Gramians against varying final times t_e (bottom right) for the bips_606 system

Because the bips systems have eigenvalues very close to the imaginary axis which caused difficulties for all considered numerical methods, the shifted matrix $A - 0.08M$ is used instead as in [22]. The generalized state-space systems vertstand and rail are dealt with Cholesky factorizations $M = LL^T$ as explained in Section 4.2. There, the sparse Cholesky factors L are computed by the MATLAB command chol(M,'vector'). Matrix exponentials and Lyapunov equations defined by smaller dense matrices (including the projected ones in Algorithm 3) are solved directly by the expm and lyap routines.

5.1 Decay of the eigenvalues of the Gramians and the time-limited singular values

At first we investigate how the end time t_e influences the eigenvalue decay of the time-limited Gramians. The index one descriptor system bips_606 is used for this experiment and reformulated into an equivalent state space system of dimension $n_f = 606$ as explained in Section 4.2. This comparatively small size allows a direct computation of the matrix exponentials and the Gramians. The top left plot in Fig. 1 shows the scaled and ordered eigenvalues λ_j/λ_1 of the infinite reachability Gramian P_∞ and the time-limited one $P_\mathcal{T}$ for $t_e = 1, 3, 10$. Obviously, a distinctly faster eigenvalue decay of $P_\mathcal{T}$ is only observed for small time values $t_e = 1, 3$. As the final time increases, the eigenvalues move closer to the ones of P_∞. The eigenvalues of the observability Gramians exhibit a similar behavior. This observation is even more drastic for the decay of the time-limited Hankel singular values shown in the bottom left plot. For the largest value $t_e = 10$, hardly any difference to the Hankel singular values is visible. In the top right plot the point wise norm of the impulse response $y_\delta(t)$ shows that after $t_e = 10$, $y_\delta(t)$ has already almost reached its stationary phase. This confirms the expectation that significant differences between infinite and time-limited Gramians occur only for times t_e which are small with respect to the behavior of the impulse response. The bottom right plots shows the numerical rank of infinite, time-limited and modified time-limited Gramians against t_e. The numerical ranks of $P_{\mathcal{T}},\ Q_{\mathcal{T}}$ clearly move toward the numerical ranks of $P_\infty,\ Q_\infty$ as t_e increases. In contrast, the numerical ranks of $P_{\mathcal{T}}^{\mathrm{mod}},\ Q_{\mathcal{T}}^{\mathrm{mod}}$ (Section 4.3) are always close to the ones of $P_\infty,\ Q_\infty$ and appear to be largely unaffected by different values of t_e. This is a very similar behavior as observed for the modified frequency-limited Gramians in [8].

5.2 Computing low-rank factors of the time-limited Gramians

We proceed by testing the computation of low-rank factors of the infinite and (modified) time-limited reachability Gramians by the rational Krylov subspace method in Algorithm 3. The stopping criteria for matrix function and Gramian approximations use the thresholds $\tau_f = \tau_P = 10^{-8}$. To save some computational cost, the projected matrix exponentials and Lyapunov equations (Steps 11 and 13 in Algorithm 3) are only dealt with every 5th iteration step.

Table 2 summarizes the used time values t_e, the produced subspace dimensions d, the ranks r of the low-rank approximations, and the total computing times t_{rk} for the Gramians P_∞, $P_{\mathcal{T}}$, and $P_{\mathcal{T}}^{\mathrm{mod}}$. Apparently, larger rational Krylov subspaces and approximately twice as long computation times are needed to obtain low-rank factors of the time-limited Gramians, but the final ranks of the low-rank approximations of the time-limited Gramians $P_{\mathcal{T}}$ are in all cases smaller compared to P_∞. The modified time-limited Gramians $P_{\mathcal{T}}^{\mathrm{mod}}$ do not show this behavior as they have ranks similar to the infinite Gramians. For the `rail` example, increasing the time horizon t_e does not change the required subspace dimension d for the approximation of the time-limited

Table 2 Results of the numerical numerical computation of low-rank factors of the different Gramians: time horizon t_e, generated subspace dimension d, rank r of the low-rank approximations, and computing time t_{rk} in seconds

		P_∞			$P_{\mathcal{T}}$			$P_{\mathcal{T}}^{\mathrm{mod}}$		
Example	t_e	d	r	t_{rk}	d	r	t_{rk}	d	r	t_{rk}
`bips_3078`	3	308	245	18.33	664	131	105.89	664	241	106.17
`vertstand`	300	180	164	6.18	300	140	13.42	300	183	13.55
`rail`	10	245	224	34.71	420	168	74.06	420	228	76.06
	100	245	224	34.71	420	198	76.53	420	227	74.68

Gramian $P_\mathcal{T}$, but its rank r is clearly larger. In all cases, no additional steps of the rational Krylov method were necessary once the approximation of $\mathrm{e}^{At_e}B$ was found. The results for the observability Gramians were largely similar. For the `bips_3078` example, the observability Gramians appeared to be more demanding for Algorithm 3 than the reachability Gramians.

5.3 Model reduction results

Now we execute BT [36] as well as (modified) TLBT [25, 29] using the square-root method (Algorithm 1) with the low-rank factors of the reachability and observability Gramians computed in the previous section. For this purpose, we restrict ourselves to the reduction to fixed specified orders r. The approximation quality of the constructed reduced order models is assessed via the point wise and maximal relative error norms

$$\mathcal{E}(t) := \tfrac{\|y(t)-y_r(t)\|_2}{\|y(t)\|_2},\ t \leq t_f, \quad \mathcal{E}_\mathcal{T} := \max_{t\in[0,t_e]} \mathcal{E}(t)$$

of the output responses $y(t)$, $\tilde{y}(t)$ of original and reduced order models. We consider the response $y_\delta(t)$ for the impulse input $u(t) = \delta(t)\mathbf{1}_m$ for all examples. Moreover, for each used test system, the transient response with respect to an additional input signal $u(t)$ is also considered. We use step like input signals $u(t) = \mathbf{1}_m$ and $u(t) = 50\mathbf{1}_m$ for the `bips_3078` and, respectively, `rail` example, and $u_* := [5 \cdot 10^4 \cdot 0.198(\sin(t\pi/100)^2), 4, 2, 1, 3, 1]^T$ for `vertstand`.

The results are given in Table 3 listing the largest relative error $\mathcal{E}_\mathcal{T}$ in $[0, t_e]$ and the overall computation time t_{mor}, i.e., the computation time for computing low-rank factors of both reachability and observability Gramians by Algorithm 3 plus the time for Algorithm 1 to execute the BT variants. We also indicate whether the produced reduced order models are asymptotically stable ($s = 1$) or unstable ($s = 0$). For some selected settings of $u(t)$ and t_e, the system responses and point wise relative

Table 3 Model reduction results by BT, TLBT, and modified TLBT using different t_e and $u(t)$: reduced order r, largest relative error $\mathcal{E}_\mathcal{T}$ in $[0, t_e]$, overall computation time t_{mor}, and $s \in \{0, 1\}$ indicates if reduced system is asymptotically stable or unstable

			BT			TLBT			mod. TLBT		
Example	u	t_e	$\mathcal{E}_\mathcal{T}$	t_{mor}	s	$\mathcal{E}_\mathcal{T}$	t_{mor}	s	$\mathcal{E}_\mathcal{T}$	t_{mor}	s
`bips_3078`,	δ	3	5.10e-04	291.3	1	1.08e-06	331.9	0	5.15e-04	399.6	1
$r = 100$	1	3	6.90e-06	291.3	1	6.33e-09	331.9	0	1.20e-05	399.6	1
`vertstand`,	δ	300	8.90e-02	13.2	1	2.44e-02	30.4	1	6.48e-02	31.6	1
$r = 20$	u_*	300	8.26e-03	13.2	1	9.18e-04	30.4	1	9.95e-03	31.6	1
`rail`,	δ	10	6.60e-01	62.2	1	5.66e-04	130.8	0	6.59e-01	137.4	1
$r = 50$	δ	100	6.60e-01	62.2	1	7.90e-03	135.7	1	6.60e-01	136.4	1
	50	10	1.78e-03	62.2	1	6.08e-07	130.8	0	2.61e-03	137.4	1
	50	100	1.78e-03	62.2	1	2.57e-05	135.7	1	2.15e-03	136.4	1

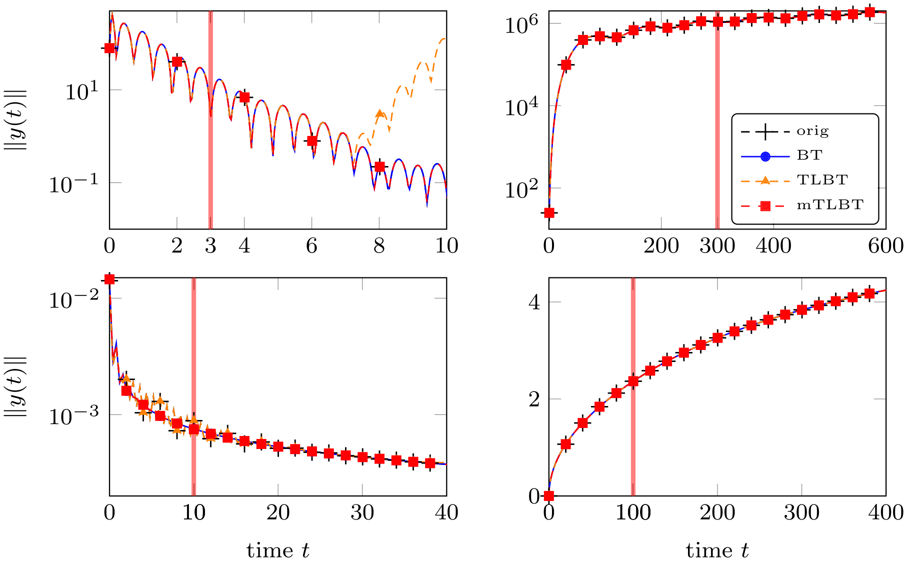

Fig. 2 Responses of original and reduced systems obtained by different BT versions: (top left) `bips_3078`, $t_e = 3$, $u(t)$ =impulse, $r = 100$, (top right) `vertstand`, $t_e = 300$, $u(t) = u_*$, $r = 20$, (bottom left) `rail`, $t_e = 10$, $u(t)$ =impulse, $r = 50$, (bottom left) `rail`, $t_e = 100$, $u(t) = 50$, $r = 50$

errors $\mathcal{E}(t)$ are plotted against the time t in Figs. 2 and 3, respectively. Figure 4 shows the behavior of $\mathcal{E}_\mathcal{T}$ as the reduced order r increases.

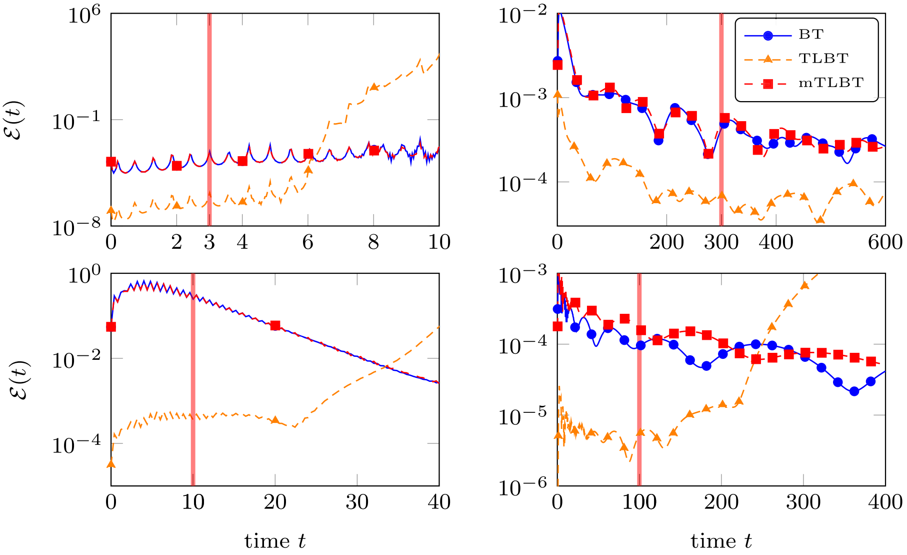

Fig. 3 Relative errors $\mathcal{E}(t)$ obtained by different BT versions: (top left) `bips_3078`, $t_e = 3$, $u(t)$ =impulse, $r = 100$, (top right) `vertstand`, $t_e = 300$, $u(t) = u_*$, $r = 20$, (bottom left) `rail`, $t_e = 10$, $u(t)$ =impulse, $r = 50$, (bottom left) `rail`, $t_e = 100$, $u(t) = 50$, $r = 50$

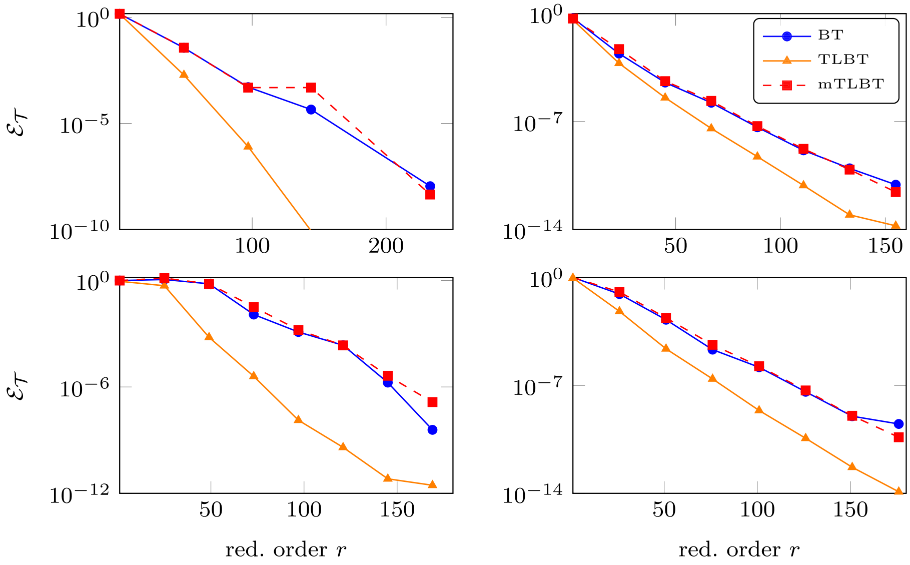

Fig. 4 Maximum relative errors $\mathcal{E}_{\mathcal{T}}$ in $[0, t_e]$ against increasing reduced orders r for different $u(t)$ and BT variants: (top left) `bips_3078`, $t_e = 3$, $u(t)$ =impulse, (top right) `vertstand`, $t_e = 300$, $u(t) = u_*$, (bottom left) `rail79k`, $t_e = 10$, $u(t)$ =impulse, (bottom left) `rail79k`, $t_e = 100$, $u(t) = 50$

Apparently, for the chosen orders r and in the time regions of interest, the largest relative errors $\mathcal{E}_{\mathcal{T}}$ of the reduced order models returned by TLBT are in most experiments more than one order of magnitude smaller compared to standard and modified time-limited BT. The plots in Fig. 4 also indicate that much larger reduced order models are needed for BT and modified TLBT to achieved the same accuracy as unmodified TLBT. Figure 3 shows that after leaving the time region $\mathcal{T}$, TLBT delivers larger errors. However, Table 3 also reveals that executing TLBT and its modified version is more time consuming than standard BT. This is a direct consequence of the higher computation times for getting the low-rank factors of the (modified) time-limited Gramians which was pointed out in the previous subsection (Table 2).

Using the concept of angles between subspaces or the modal assurance criterion (MAC) ($MAC(x, y) = |y^T x|^2/(\|x\|^2\|y\|^2)$, [20]) indicated that the spaces spanned by the projection matrices T_{TLBT}, S_{TLBT} in TLBT and T_{BT}, S_{BT} in unrestricted BT, respectively, are different. For example, computing the MAC for the right projection matrices $T_{\mathrm{TLBT}}, T_{\mathrm{BT}} \in \mathbb{R}^{n_f \times 100}$ for the `bips_3078` system showed that only a few of the columns of both matrices are well correlated to each other (i.e., $MAC(T_{\mathrm{TLBT}}e_i, T_{\mathrm{BT}}e_j) \approx 1$ for very few $i, j \in \{1, \ldots, 100\}$).

Moreover, albeit the higher accuracy in $[0, t_e]$, in some cases TLBT produces unstable reduced order models. This is especially visible in the upper left plot of Fig. 2 showing the impulse responses of `bips_3078`. For times $t \geq t_e$ outside the interval $[0, t_e]$, the impulse response of reduced order model generated by TLBT exhibits an exponential growth and departs from the original response. As illustrated

Table 4 Results of BT, TLBT, and modified TLBT reduction to $r = 15$ of `vertstand` example with respect to time frame $\mathcal{T} = [t_s, t_e] = [50, 100]$

	BT			TLBT			mod. TLBT		
input u	$\mathcal{E}_{\mathcal{T}}$	t_{mor}	s	$\mathcal{E}_{\mathcal{T}}$	t_{mor}	s	$\mathcal{E}_{\mathcal{T}}$	t_{mor}	s
δ	7.95e-04	14.2	1	3.07e-04	37.6	0	9.21e-04	32.5	1
u_*	6.34e-03	14.2	1	1.78e-02	37.6	0	6.85e-03	32.5	1

with the `rail` examples and proven in [25], using a higher end time t_e can already cure this. Modified TLBT does not generate unstable reduced systems, but its approximation quality in $[0, t_e]$ is very close to standard BT without time restrictions. Taking also into account the higher computational costs of modified TLBT given in Table 3, the introduction of the time restriction is rendered essentially redundant because no smaller errors are achieved in the targeted time region. Hence, if stability preservation in the reduced order model is crucial, we recommend to stick to standard BT.

To conclude, TLBT fulfills the goal to acquire smaller errors in the desired time interval $[0, t_e]$, but at the price of somewhat larger execution times because the computation of required low-rank Gramian factors is currently more costly.

Now we carry out one experiment to evaluate the approximation qualities of TLBT for nonzero t_s. The `vertstand` example is used and the reduced order model should approximate the output $y(t)$ with respect to $u(t) = \delta(t)v$ and $u(t) = u_* = [5 \cdot 10^4 \cdot 0.198(\sin(t\pi/100)^2), 4, 2, 1, 3, 1]^T$ in the time window $\mathcal{T} = [t_s, t_e] = [50, 100]$. The low-rank factors of the Gramians are computed as before using Algorithm 3 with the required small extensions mentioned in Section 4.1.

The results are summarized in Table 4 and Fig. 5 illustrates the relative errors plots as well as the largest relative error in $\mathcal{T}$ against the reduced order r. Compared to standard BT, TLBT delivers, as expected, more accurate results of the impulse response, but fails for $u(t) = u_*$. Moreover, the bottom left plot Fig. 5 shows that the decay of $\mathcal{E}_{\mathcal{T}}$ with respect to the impulse response is less monotonic as in the case $t_s = 0$ s.t. for some reduced orders r, especially larger ones, standard BT outperforms TLBT. The failure of TLBT to approximate the transient response for arbitrary inputs $u(t)$ was also observed in other experiments with different time intervals and examples, and occasionally even for the impulse response. In experiments with smaller systems, using exact matrix exponentials and Gramian factors did not yield any improvement of the reduction results such that the problems with TLBT are most likely not a result of inaccurate Gramian approximations. In summary, although TLBT in the current form does indeed deliver significantly more accurate reduced order models in the time intervals $[0, t_e]$, the same cannot be said when time intervals $[t_s, t_e]$, $t_s > 0$ are considered. Improving the performance of TLBT in this scenario is therefore subject of further research.

As the final experiment we briefly test the application of TLBT for the reduction of unstable systems. For this purpose a modification of the `bips_606` example is used with $\hat{A} := A + 0.2M$ leading to $\max \operatorname{Re}(\lambda) = 0.323$, $\lambda \in \Lambda(\hat{A})$. The final

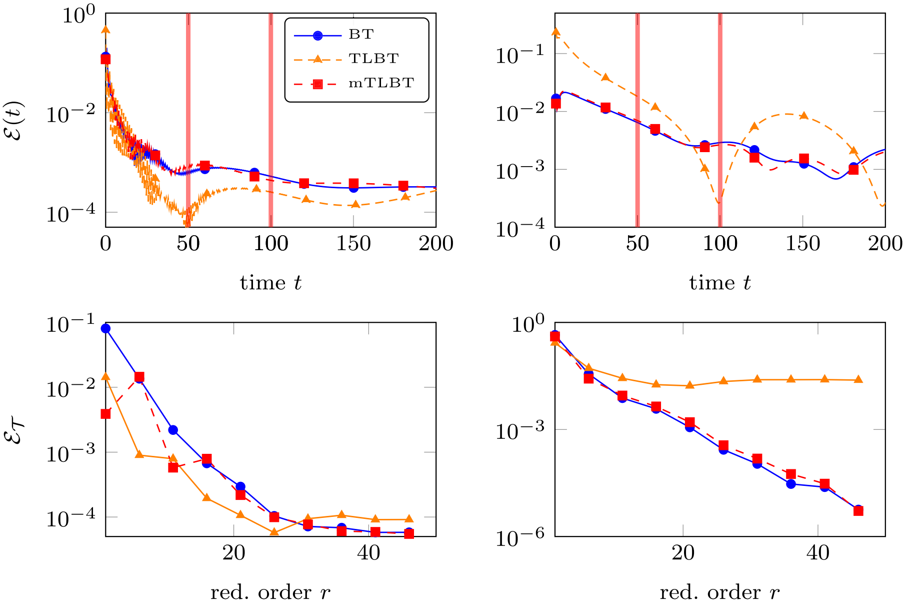

Fig. 5 Results for the reduction of the `vertstand` example in $[t_s, t_e] = [50, 100]$. The top plots show the relative error norms against t for $u(t) = \delta(t)v$ (left) and $u(t) = u_*$ (right) after a reduction to $r = 15$. The behavior of $\mathcal{E}_{\mathcal{T}}$ in $[t_s, t_e]$ for increasing reduced dimensions r is illustrated in the bottom plots

time is set to $t_e = 5$ and a reduced order model with $r = 100$ is generated. Due to the comparatively small system dimension ($n_f = 606$), the matrix exponentials and Gramians could be computed by direct, dense methods for these tasks. Moreover, the employed rational Krylov methods from the experiments before were not able to compute the required low-rank Gramians factors. The impulse response of exact and reduced system and the relative error are illustrated in Fig. 6.

Apparently, TLBT was able to reduce this mildly unstable system to a reduced order model with maximal relative error $\mathcal{E}_{\mathcal{T}} \approx 1.2 \cdot 10^{-2}$ in $[0, t_e]$. Hence, provided t_e is chosen in a reasonable way, e.g., with a sufficient distance to the exponential

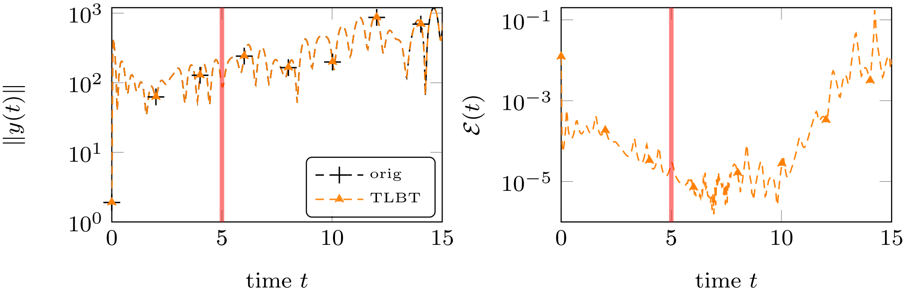

Fig. 6 Results for the reduction of the unstable variation of the `bips_606` example for $t_e = 5$, $r = 100$, $u = \delta(t)v$ (left: system response, right: relative error)

growth of $y(t)$, TLBT appears to be a potential candidate from model order reduction of unstable systems. However, the applications to large-scale unstable systems is currently still difficult because algorithms for computing low-rank solutions of Lyapunov equations usually require that A is (anti)stable. Advances in this direction are, therefore, necessary to pursue this type of reduction further.

6 Conclusions

BT model order reduction for large-scale systems restricted to finite time intervals [25] was investigated. The resulting Lyapunov equations that have to be solved numerically also include the matrix exponential in their inhomogeneities. We first showed that the difference of time-limited and infinite Gramians decays for increasing times in a similar way as the impulse response of the underlying system. Hence, for small time intervals, a reduced numerical rank of the time-limited Gramians can be observed. Future research should further investigate the influence of the chosen end time on the eigenvalue decay of the Gramians.

As in frequency-limited BT [8], we proposed to handle the matrix exponentials and the Lyapunov equations by an efficient rational Krylov subspace method incorporating a subspace recycling idea. While this numerical approach already led to satisfactory results, there is some room for improvement, especially regarding the approximation of the action of the matrix exponential. In this context, further work could include, for instance, enhanced strategies like tangential directions [17], different inner products [23], or using different methods [1, 12] altogether.

The numerical reduction experiments indicated that TLBT is able to acquire several orders of magnitude more accurate reduced order models in time intervals of the form $[0, t_e]$ at a somewhat higher, although still comparable, numerical effort. Similar techniques were applied for stability preserving modified TLBT [29] which, however, could not keep up with standard or time-limited BT in terms of efficiency and accuracy of the reduced order models. Hence, we recommend to use standard BT if the preservation of stability is an irrevocable goal. For keeping the high accuracy in the time-interval of interest, a different way to make TLBT a stability preserving method has to be found. TLBT for time regions $[t_s, t_e]$ with nonzero start times $t_s > 0$ provided much worse results than with $t_s = 0$, except for approximating the impulse response. In the form introduced in [25], TLBT appears to be incapable of producing good reduced order models with respect to $[t_s, t_e]$ and, thus, further research is necessary in this direction. The results in [5, 33] might be one possible ingredient for this. A brief experiment also indicated that TLBT can be employed to reduce unstable systems. The efficient computation of low-rank Gramian factors in this case is currently not as advanced as in the stable situation, making this a further interesting research topic.

Acknowledgments Open access funding provided by Max Planck Society. I thank the referees for their helpful comments. Moreover, I am grateful for the constructive discussions with Maria Cruz Varona, Serkan Gugercin, and Stefan Guettel.

References

1. Al-Mohy, A.H., Higham, N.J.: Computing the action of the matrix exponential, with an application to exponential integrators. SIAM J. Sci. Comput. **33**(2), 488–511 (2011)
2. Antoulas, A.C., Sorensen, D.C., Zhou, Y.: On the decay rate of Hankel singular values and related issues. Syst. Cont. Lett. **46**(5), 323–342 (2002)
3. Baker, J., Embree, M., Sabino, J.: Fast singular value decay for Lyapunov solutions with nonnormal coefficients. SIAM J. Matrix Anal. Appl. **36**(2), 656–668 (2015)
4. Bartels, R.H., Stewart, G.W.: Solution of the matrix equation $AX + XB = C$: Algorithm 432. Comm. ACM **15**, 820–826 (1972)
5. Beattie, C., Gugercin, S., Mehrmann, V.: Model reduction for systems with inhomogeneous initial conditions. Syst. Control Lett. **99**, 99–106 (2017). https://doi.org/10.1016/j.sysconle.2016.11.007
6. Beckermann, B., Reichel, L.: Error estimates and evaluation of matrix functions via the faber transform. SIAM J. Numer. Anal. **47**(5), 3849–3883 (2009). https://doi.org/10.1137/080741744
7. Benner, P.: Solving large-scale control problems. IEEE Control Syst. Mag. **14**(1), 44–59 (2004)
8. Benner, P., Kürschner, P., Saak, J.: Frequency-limited balanced truncation with low-rank approximations. SIAM J. Sci. Comput. **38**(1), A471–A499 (2016). https://doi.org/10.1137/15M1030911
9. Benner, P., Saak, J.: Numerical solution of large and sparse continuous time algebraic matrix Riccati and Lyapunov equations: a state of the art survey. GAMM Mitteilungen **36**(1), 32–52 (2013). https://doi.org/10.1002/gamm.201310003
10. Berljafa, M., Güttel, S.: Generalized rational krylov decompositions with an application to rational approximation. SIAM J. Matrix Anal. Appl. **36**(2), 894–916 (2015)
11. Breiten, T., Beattie, C., Gugercin, S.: Near-optimal frequency-weighted interpolatory model reduction. Sys. Control Lett. **78**, 8–18 (2015)
12. Caliari, M., Kandolf, P., Ostermann, A., Rainer, S.: Comparison of software for computing the action of the matrix exponential. BIT **54**(1), 113–128 (2014)
13. Davies, P.I., Higham, N.J.: Computing $f(A)b$ for matrix functions f. In: Boriçi, A., Frommer, A., Joó, B., Kennedy, A., Pendleton, B. (eds.) QCD and Numerical Analysis III, Lect. Notes Comput. Sci. Eng., vol. 47, pp. 15–24. Springer, Berlin (2005). https://doi.org/10.1007/3-540-28504-0_2
14. Druskin, V., Knizhnerman, L., Zaslavsky, M.: Solution of large scale evolutionary problems using rational krylov subspaces with optimized shifts. SIAM J. Sci. Comput. **31**(5), 3760–3780 (2009)
15. Druskin, V., Lieberman, C., Zaslavsky, M.: On adaptive choice of shifts in rational krylov subspace reduction of evolutionary problems. SIAM J. Sci. Comput. **32**(5), 2485–2496 (2010)
16. Druskin, V., Simoncini, V.: Adaptive rational Krylov subspaces for large-scale dynamical systems. Syst. Control Lett. **60**(8), 546–560 (2011)
17. Druskin, V., Simoncini, V., Zaslavsky, M.: Adaptive tangential interpolation in rational Krylov subspaces for MIMO dynamical systems. SIAM J. Matrix Anal. Appl. **35**(2), 476–498 (2014). https://doi.org/10.1137/120898784
18. Du, X., Benner, P.: Finite-Frequency Model Order Reduction of Linear Systems via Parameterized Frequency-dependent Balanced Truncation. Tech. Rep. 1602.04408. ArXiv e-prints (2016)
19. Du, X., Benner, P., Yang, G., Ye, D.: Balanced truncation of linear time-invariant systems at a single frequency. Preprint MPIMD/13-02, Max Planck Institute Magdeburg. Available from http://www.mpi-magdeburg.mpg.de/preprints/ (2013)
20. Ewins, D.: Modal testing: theory, practice, and application (2nd edn) Research Study Press LTD (2000)
21. Fehr, J., Fischer, M., Haasdonk, B., Eberhard, P.: Greedy-based approximation of frequency-weighted Gramian matrices for model reduction in multibody dynamics. Z. Angew. Math. Mech. **93**(8), 501–519 (2013)
22. Freitas, F., Rommes, J., Martins, N.: Gramian-based reduction method applied to large sparse power system descriptor models. IEEE Trans. Power Syst. **23**(3), 1258–1270 (2008)

23. Frommer, A., Lund, K., Szyld, D.B.: Block Krylov subspace methods for computing functions of matrices applied to multiple vectors. Tech. Rep. 17-03-21, Department of Mathematics Temple University (2017)
24. Frommer, A., Simoncini, V.: Matrix functions. In: Schilders, W., van der Vorst, H., Rommes, J. (eds.) Model Order Reduction: Theory, Research Aspects and Applications, Mathematics in Industry, vol. 13, pp. 275–303. Springer, Berlin (2008). https://doi.org/10.1007/978-3-540-78841-6_13
25. Gawronski, W., Juang, J.: Model reduction in limited time and frequency intervals. Int. J. Syst. Sci. **21**(2), 349–376 (1990). https://doi.org/10.1080/00207729008910366
26. Goyal, P., Redmann, M.: Towards time-limited $\mathcal{H}_2$-optimal model order reduction. Tech. Rep. WIAS Preprint No. 2441 (2017)
27. Grasedyck, L.: Existence of a low rank or H-matrix approximant to the solution of a Sylvester equation. Numer. Lin. Alg. Appl. **11**, 371–389 (2004)
28. Großmann, K., Städel, C., Galant, A., Mühl, A.: Berechnung von Temperaturfeldern an Werkzeugmaschinen. Zeitschrift fü,r Wirtschaftlichen Fabrikbetrieb **107**(6), 452–456 (2012)
29. Gugercin, S., Antoulas, A.C.: A survey of model reduction by balanced truncation and some new results. Internat. J. Control **77**(8), 748–766 (2004). https://doi.org/10.1080/00207170410001713448
30. Güttel, S.: Rational Krylov Methods for Operator Functions. Ph.D. thesis, Technische Universität Bergakademie Freiberg, Germany. Available from http://nbn-resolving.de/urn:nbn:de:bsz:105-qucosa-27645 (2010)
31. Güttel, S.: Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection. GAMM-Mitteilungen **36**(1), 8–31 (2013). https://doi.org/10.1002/gamm.201310002
32. Halevi, Y.: Frequency weighted model reduction via optimal projection. IEEE Trans. Automat. Control **37**(10), 1537–1542 (1992)
33. Heinkenschloss, M., Reis, T., Antoulas, A.C.: Balanced truncation model reduction for systems with inhomogeneous initial conditions. Automatica **47**(3), 559–564 (2011). https://doi.org/10.1016/j.automatica.2010.12.002
34. Higham, N.: Functions of matrices: Theory and Computation. Society for Industrial and Applied Mathematics, Philadelphia (2008)
35. Knizhnerman, L.A.: Calculation of functions of unsymmetric matrices using Arnoldi's method. Comput. Math. Math. Phys. **31**(1), 1–9 (1992)
36. Moore, B.C.: Principal component analysis in linear systems: controllability, observability, and model reduction. IEEE Trans. Autom. Control **AC–26**(1), 17–32 (1981). https://doi.org/10.1109/TAC.1981.1102568
37. Penzl, T.: Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case. Syst. Cont. Lett. **40**, 139–144 (2000). https://doi.org/10.1016/S0167-6911(00)00010-4
38. Petersson, D.: A Nonlinear Optimization Approach to H2-Optimal Modeling and Control. Ph.D. thesis, Linköping University. Available from http://www.diva-portal.org/smash/get/diva2:647068/FULLTEXT01.pdf (2013)
39. Petersson, D., Löfberg, J.: Model reduction using a frequency-limited $\mathcal{H}_2$-cost. Sys. Control Lett. **67**, 32–39 (2014)
40. Redmann, M., Kürschner, P.: An $\mathcal{H}_2$-Type Error Bound for Time-Limited Balanced Truncation. Tech. Rep. 1710.07572v1. ArXiv e-prints (2017)
41. Ruhe, A.: Rational Krylov sequence methods for eigenvalue computation. Linear Algebra Appl. **58**, 391–405 (1984)
42. Ruhe, A.: The rational Krylov algorithm for nonsymmetric Eigenvalue problems. III: complex shifts for real matrices. BIT **34**, 165–176 (1994)
43. Saad, Y.: Numerical Solution of Large Lyapunov Equation. In: Kaashoek, M.A., van Schuppen, J.H., Ran, A.C.M. (eds.) Signal Processing, Scattering, Operator Theory and Numerical Methods, pp. 503–511, Birkhäuser (1990)
44. Saad, Y.: Analysis of some Krylov subspace approximations to the matrix exponential operator. SIAM J. Numer. Anal. **29**(1), 209–228 (1992)
45. Sabino, J.: Solution of large-scale Lyapunov equations via the block modified smith method. Ph.D. thesis, Rice University, Houston, Texas. http://www.caam.rice.edu/tech_reports/2006/TR06-08.pdf (2007)
46. Simoncini, V.: Analysis of the rational Krylov subspace projection method for large-scale algebraic Riccati equations. SIAM J. Matrix Anal. Appl. **37**(4), 1655–1674 (2016). https://doi.org/10.1137/16M1059382

47. Sinani, K., Gugercin, S.: Iterative Rational Krylov Algorithms for Unstable Dynamical Systems and Optimality Conditions for a Finite-Time Horizon (2017). http://meetings.siam.org/sess/dsp_talk.cfm?p=81168 SIAM CSE (2017)
48. Truhar, N., Veselić, K.: Bounds on the trace of a solution to the Lyapunov equation with a general stable matrix. Syst. Cont. Lett. **56**(7–8), 493–503 (2007). https://doi.org/10.1016/j.sysconle.2007.02.003
49. Vuillemin, P.: Frequency-limited model approximation of large-scale dynamical models. Ph.D. thesis, Université de Toulouse. https://hal.archives-ouvertes.fr/tel-01092051 (2014)