The balanced truncation error bound in Schatten norms

Mark R. Opmeer and Timo Reis

Abstract—The first main result in this article provides an error bound for balanced truncation where the matrix norm used is a general Schatten norm rather than the usual operator norm. The second main result in this article is that for the Schatten 1-norm (the trace class norm) this bound, for systems with a semi-definite Hankel operator, is in fact an equality. This class of systems for which we obtain equality includes state space symmetric systems.

Index Terms—balanced realization, balanced truncation, Hankel operator, error bound, model reduction, linear time-invariant systems.

I. INTRODUCTION

The well-known error bound for balanced truncation

\[ \sup_{\zeta \in \mathbb{C} : \Re\zeta > 0} \| G(\zeta) - G_r(\zeta) \| \leq 2 \sum_{j=r+1}^{\ell} \mu_j, \]  

(1)

where \( \{\mu_1, \ldots, \mu_{\ell}\} \) are the distinct Hankel singular values of \( G \) and \( G_r \) is the balanced truncation of \( G \), is known to be an equality for single-input single-output (SISO) state space symmetric systems, i.e. if \( G(s) = C(sI - A)^{-1}B \) with \( A = A^* \in \mathbb{C}^{n \times n} \) negative definite, \( C^* = B \in \mathbb{C}^n \), then

\[ \sup_{\zeta \in \mathbb{C} : \Re\zeta > 0} \| G(\zeta) - G_r(\zeta) \| = 2 \sum_{j=r+1}^{\ell} \mu_j, \]

(2)

(see e.g. [10, Theorem 4.1] and [18, Theorem 4.4]). It is also known that in this case (this follows e.g. from [15, Corollary 2.2]) the Hankel singular values of \( G \) all have multiplicity one. Moreover, it is known that for multi-input multi-output (MIMO) state space symmetric systems, i.e. if \( G(s) = C(sI - A)^{-1}B \) with \( A = A^* \in \mathbb{C}^{n \times n} \) negative definite, \( C^* = B \in \mathbb{C}^{n \times m} \) with \( m > 1 \), strict inequality may hold (see e.g. [10, Remark 4.1] and [18, Section 4]). In this article we further investigate this MIMO case.

In the MIMO case, different matrix norms can be chosen for the matrix \( G(\zeta) - G_r(\zeta) \). In (1) the operator norm is chosen (i.e. the largest singular value of \( G(\zeta) - G_r(\zeta) \)). The first result of this article is that if we instead choose the Schatten \( p \)-norm, i.e.

\[ \| T \|_p := \left( \sum_{k=1}^{n} |\sigma_k(T)|^p \right)^{1/p}, \quad p \in [1, \infty), \]

where \( \sigma_1(T) \geq \ldots \geq \sigma_m(T) \) are the singular values of \( T \in \mathbb{C}^{m \times m} \), then the following error-bound holds for balanced truncation

\[ \sup_{\zeta \in \mathbb{C} : \Re\zeta > 0} \| G(\zeta) - G_r(\zeta) \|_p \leq 2 \sum_{j=r+1}^{\ell} m_j^{1/p} \mu_j, \]  

(2)

where \( m_j \) is the multiplicity of \( \mu_j \) as a singular value of the Hankel operator of \( G \). For the case \( p = \infty \) (with the convention that \( 1/\infty = 0 \)) this reduces to (1). The most interesting other case is \( p = 1 \):

\[ \sup_{\zeta \in \mathbb{C} : \Re\zeta > 0} \| G(\zeta) - G_r(\zeta) \|_1 \leq 2 \sum_{j=r+1}^{\ell} m_j \mu_j, \]  

(3)

(see above that the right-hand side is equal to \( 2 \sum_{k=q+1}^{n} \lambda_k(H_G) \), where \( q := \sum_{j=1}^{r} m_j \) is the dimension of the balanced truncation and \( \lambda_k(H_G) \) denotes the \( k \)-th singular value of the Hankel operator \( H_G \) of \( G \).

For systems with a semi-definite Hankel operator (this includes state space symmetric systems), we show that for this trace class norm (Schatten 1-norm) equality hold:

\[ \sup_{\zeta \in \mathbb{C} : \Re\zeta > 0} \| G(\zeta) - G_r(\zeta) \|_1 = 2 \sum_{j=r+1}^{\ell} m_j \mu_j, \]  

(3)

were we again note that the right-hand side can alternatively be written as \( 2 \sum_{k=q+1}^{n} \sigma_k(H_G) \). The proof of this is based on the inequality

\[ 2 \left( \sum_{k=q+1}^{n} \lambda_k(H_G) \right) \leq \sup_{\zeta \in \mathbb{C} : \Re\zeta > 0} \| G(\zeta) - G_r(\zeta) \|_1, \]  

(4)

where \( \lambda_k(H_G) \) are the eigenvalues of the Hankel operator \( H_G \) of \( G \) and as above \( q := \sum_{j=1}^{r} m_j \) equals the dimension of the balanced truncation. This inequality holds for systems with a self-adjoint Hankel operator. The SISO case of (4) is the main result of [12] (note that in the SISO case all Schatten norms are the same and that in the real-valued SISO case every Hankel operator is self-adjoint).

In Section II we first discuss the notation and terminology used. In Section III the upper-bound (2) is proven, in Section IV the lower-bound (4) is proven and in Section V the equality (3) is proven. Finally, Section VI contains comments on balanced singular perturbation approximation and on the case of non-rational transfer functions.

M.R. Opmeer is with the Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom, e-mail: m.opmeer@bath.ac.uk.

T. Reis is with the Department of Mathematics, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany, e-mail: timo.reis@uni-hamburg.de.
II. NOTATION AND TERMINOLOGY

The set of \( p \times m \)-matrices with entries in the field of complex rational functions is denoted by \( \mathbb{C}(s)^{p \times m} \). We call \( G \in \mathbb{C}(s)^{p \times m} \) stable if \( G \) is proper and all its poles have negative real part.

The impulse response \( h \) is the inverse Laplace transform of \( G \in \mathbb{C}(s)^{p \times m} \). The Hankel operator of a stable \( G \in \mathbb{C}(s)^{p \times m} \) is given by

\[
H: L^2(0, \infty; \mathbb{C}^m) \to L^2(0, \infty; \mathbb{C}^p), \quad u \mapsto (Hu)(t) = \int_0^\infty h_0(t + s)u(s) \, ds,
\]

where \( h_0 \) is the function part of the impulse response, i.e., the inverse Laplace transform of the strictly proper part of \( G \).

The nonzero singular values of the Hankel operator of \( G \) are called the Hankel singular values of \( G \). We denote the sequence of Hankel singular values by \( (\sigma_k)_{k=1}^\infty \), the sequence of distinct Hankel singular values by \( (\mu_j)_{j=1}^\ell \) and the sequence of multiplicities of the Hankel singular values by \( (m_j)_{j=1}^\ell \) (i.e. \( m_j \) is the number of times that \( \mu_j \) appears in the sequence \( (\sigma_k)_{k=1}^\infty \)). We choose the ordering of these sequences to be compatible in the following sense

\[
\sigma_{1+\sum_{j=1}^{i-1} m_j} = \ldots = \sigma_{\sum_{j=1}^{i} m_j} = \mu_i, \quad i = 1, \ldots, \ell.
\]

We denote the sequence of nonzero eigenvalues of the Hankel operator by \( (\lambda_k)_{k=1}^n \) and call these the Hankel eigenvalues of \( G \). We note that if the Hankel operator is self-adjoint, then the absolute values of the Hankel eigenvalues equal the Hankel singular values (including multiplicities). In this case we choose the ordering of these sequences to be compatible in the sense that

\[
|\lambda_k| = \sigma_k, \quad k = 1, \ldots, n.
\]

Giving the ordering of the Hankel singular values this may not uniquely determine the ordering of the Hankel eigenvalues, but for our purposes this particular non-uniqueness is irrelevant.

A realization of \( G \in \mathbb{C}(s)^{m \times n} \) is a quadruple \( \{A, B, C, D\} \) consisting of \( A \in \mathbb{C}^{n \times n} \), \( B \in \mathbb{C}^{n \times m} \), \( C \in \mathbb{C}^{p \times n} \), \( D \in \mathbb{C}^{p \times m} \) with

\[
G(s) = C(sI - A)^{-1}B + D.
\]

Conversely, \( G \) is called the transfer function of \( \{A, B, C, D\} \).

A realization \( \{A, B, C, D\} \) is called stable if all eigenvalues of \( A \) have negative real part. The reachability map \( \Phi : L^2(0, \infty; \mathbb{C}^m) \to \mathbb{C}^n \) and the observability map \( \Psi : \mathbb{C}^n \to L^2(0, \infty; \mathbb{C}^p) \) of a stable realization \( \{A, B, C, D\} \) are defined by

\[
\Phi u = \int_0^\infty e^{At}Bu(t) \, dt, \quad \Psi z = z \mapsto Ce^{At}z.
\]

We note that the Hankel operator equals the product of the observability and reachability maps: \( H = \Psi \Phi \).

The stable realization \( \{A, B, C, D\} \) is called balanced, if \( \Phi \Phi^* \in \mathbb{C}^{n \times n} \) and \( \Psi^* \Psi \in \mathbb{C}^{p \times n} \) satisfy

\[
\Phi \Phi^* = \Psi^* \Psi = \text{diag}(\sigma_1, \ldots, \sigma_n),
\]

where we recall that \( (\sigma_k)_{k=1}^n \) is the sequence of Hankel singular values.

It is well-known that a stable \( G \in \mathbb{C}(s)^{p \times m} \) has a balanced realization (see e.g. [1, Section 7.1]). Note that by our definition a balanced realization is necessarily a minimal realization.

Let \( \{A, B, C, D\} \) be a realization of \( G \in \mathbb{C}(s)^{p \times m} \). Let \( r \in \{1, \ldots, \ell\} \) and \( q := \sum_{j=1}^r m_j \). Then the balanced truncation of \( G \) of dimension \( q \) is defined by the transfer function \( G_r \) of \( \{A_r, B_r, C_r, D_r\} \), where, for \( Z_r = [I_\ell | C] \in \mathbb{C}^{n \times q} \), the matrices \( A_r \in \mathbb{C}^{q \times q} \), \( B_r \in \mathbb{C}^{q \times m} \) and \( C_r \in \mathbb{C}^{p \times q} \) are defined by \( A_r := Z_r^*AZ_r \), \( B_r := Z_r^*B \), \( C_r := CZ_r \). The balanced truncation \( G_r \) depends only on \( G \), the ordering of the distinct Hankel singular values and \( r \) (and not on the particular balanced realization chosen).

Note that the balanced truncation depends on the ordering of the sequence of distinct Hankel singular values. We assume that such an ordering is given (the customary one is the one with \( \mu_1 > \mu_2 > \ldots > \mu_\ell > 0 \); in which case \( G_r \) depends only on \( G \) and \( r \), but other orderings are permitted).

We refer the reader to [1, Chapter 7], [7, Chapter 9], [21, Chapter 7] or [22, Chapter 7] and the main original contributions [13], [17], [6], [2] for background material on balanced realizations and balanced truncations.

III. THE UPPER BOUND IN SCHATTEN NORMS

In this section we consider the upper bound in Schatten norms (2).

Theorem 1. Let \( G \in \mathbb{C}(s)^{m \times n} \) be stable. Let \( (\mu_j)_{j=1}^\ell \) denote the sequence of distinct Hankel singular values of \( G \) with multiplicities \( (m_j)_{j=1}^\ell \). For \( r \in \{1, \ldots, \ell\} \) let \( G_r \) be the balanced truncation of \( G \) of dimension \( q := \sum_{j=1}^r m_j \). Then

\[
\sup_{\zeta \in \mathbb{C} \setminus \mathbb{R} ; \Re \zeta > 0} \|G(\zeta) - G_r(\zeta)\|_p \leq \sum_{j=r+1}^\ell m_j^{1/p} \mu_j,
\]

with the convention that \( 1/\infty = 0 \).

Proof. This follows from a trivial modification of the proof of the standard error-bound (1) as given in [7, Sections 9.4.2 and 9.4.3] (which in turn is a slight sharpening of the proof given by Enns [2]). We indicate where the proof from [7] has to be adapted. On the bottom of page 328, instead of taking the operator norm of \( E(\iota \omega) \), the Schatten \( p \)-norm should be taken. From the arguments in [7, Lemma 9.4.2] it follows that \( \hat{U}(i\omega) = E(\iota \omega) \) and \( \hat{U}(\iota \omega) \) both have \( m_j \) nonzero singular values which are the same and equal \( \mu_j \) (note that [7, Lemma 9.4.2] denotes our \( \mu_j \) by \( \sigma \)). Therefore it follows that the Schatten \( p \)-norm of both \( \hat{U}(i\omega) + E(\iota \omega) \) and \( \hat{U}(\iota \omega) \) is \( (\sum_{j=1}^m \mu_j^p)^{1/p} \), which equals \( (m_j \mu_j^p)^{1/p} \), which equals \( m_j^{1/p} \mu_j \). Therefore, the Schatten \( p \)-norm of \( E(\iota \omega) \) is bounded from above by \( 2m_j^{1/p} \mu_j \). Using this replacement of [7, Lemma 9.4.2] in [7, Theorem 9.4.3] (where again Schatten norms instead of operator norms must be taken) gives the result.

Remark 2. As mentioned, Theorem 1 follows from slightly modifying the proof of the standard balanced truncation error bound originally given by Enns [2]. There are other proofs.
of the standard balanced truncation error bound. However, for example the one based on the bounded real lemma (see [1, Section 7.2.1]) does not seem to be suited for the generalization given in Theorem 1.

IV. SELF-ADJOINT SYSTEMS

In this section we consider self-adjoint systems and prove the lower-bound (4).

Definition 3. A rational function $G \in \mathbb{C}(s)^{m \times m}$ is called self-adjoint if $G = G^\dagger$, where $G^\dagger$ is defined by

$$G^\dagger(s) := (G(s))^*.$$

Remark 4. Note that any SISO system with real coefficients has a self-adjoint transfer function. Also note that the transfer function of a state space symmetric system (i.e. with $A = A^*$, $C = B^*$ and $D = D^*$) is self-adjoint.

Lemma 5. The following are equivalent for any stable and strictly proper $G \in \mathbb{C}(s)^{m \times m}$.

1) $G$ is self-adjoint.
2) The impulse response $h$ is self-adjoint (that is, $h(t) = h(t)^*$ for all $t \in [0, \infty)$).
3) The Hankel operator is self-adjoint.

Proof. The definition of the impulse response yields

$$G(s)^* = \int_0^\infty e^{-st}h(t)^* \, dt, \quad G^*(s) = \int_0^\infty e^{-st}h(t)^* \, dt.$$ From this (and uniqueness of the inverse Laplace transform) we see the equivalence of 1 and 2. Since the adjoint of the Hankel operator is given by

$$(H^* u)(t) = \int_0^\infty h(t + s)^* u(s) \, ds,$$ we see that 2 implies 3. That 3 implies 2 follows from the fact that $H - H^*$ is the Hankel operator corresponding to the impulse response $h(t) - h(t)^*$ and the fact that the zero Hankel operator must have zero impulse response. □

Remark 6. For simplicity in Lemma 5 we considered only the strictly proper case; if there is a nonzero-feedthrough, then the impulse response is no longer a function and this slightly complicates the formulation. In that case 2 has to be replaced by the function part of the impulse response being self-adjoint and additionally the feedthrough operator being self-adjoint. The condition that the feedthrough operator must be self-adjoint must also be added to condition 3. All of the above can be proven by applying Lemma 5 to $G - G(\infty)$.

The following lemma shows that a balanced realization of a self-adjoint transfer function has a certain state space symmetry property.

Lemma 7. Let $G \in \mathbb{C}(s)^{m \times m}$ be stable and self-adjoint and let $[A \ B]$ be a balanced realization of $G$. Then there exists a unique self-adjoint operator $J$ such that

$$\begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}^* \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix}. \quad (5)$$

This operator $J$ is diagonal with diagonal entries

$$J_{ii} = \frac{\lambda_i}{|\lambda_i|},$$

where $(\lambda_k)_{k=1}^n$ are the Hankel eigenvalues of $G$.

Proof. By [20, Theorem III], there exists a unique $J = J^*$ such that (5) holds true. The definition of the reachability and observability map then gives rise to $\Phi = J\Psi\Psi^*$, and thus $\Phi\Psi = J\Psi^*\Psi$. Thereby, we get $J = \Phi\Psi(\Psi^*\Psi)^{-1}$. Since $[A \ B]$ is balanced, $(\Psi^*\Psi)^{-1}$ is a diagonal matrix with on the diagonal the reciprocals of the Hankel singular values. The operator $\Phi\Psi$ is the cross-Gramian (see [3] and [19]) and in a balanced realization, it is a diagonal matrix with the Hankel eigenvalues on the diagonal. Therefore $J$ is diagonal with the indicated diagonal entries. □

The chosen subset of the Hankel singular values is retained in balanced truncation. The following lemma shows that, in the self-adjoint case, the same is true for the Hankel eigenvalues. Moreover, the lemma shows that balanced truncation preserves self-adjointness.

Lemma 8. Let $G \in \mathbb{C}(s)^{m \times m}$ be stable and self-adjoint. Denote the Hankel eigenvalues of $G$ by $(\lambda_k)_{k=1}^n$. For $r \in \{1, \ldots, \ell\}$ let $G_r$ be the balanced truncation of $G$, and let $q := \sum_{j=1}^r m_j$, where $(m_j)_{j=1}^q$ denote the multiplicities of the Hankel singular values of $G$. Then $G_r$ is self-adjoint and the Hankel eigenvalues of $G_r$ are $(\lambda_k^r)_{k=1}^q$.

Proof. Define $J^r$ as the diagonal matrix of size $q$ with $\lambda_i/|\lambda_i|$ on the diagonal. Since the $\lambda_i$ are real (because the Hankel operator is self-adjoint), this is a self-adjoint matrix. It follows from (5) that

$$\begin{bmatrix} J^r & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_r & B_r \\ C_r & D_r \end{bmatrix} = \begin{bmatrix} A_r & B_r \\ C_r & D_r \end{bmatrix}^* \begin{bmatrix} J^r & 0 \\ 0 & I \end{bmatrix},$$

where $[A_r \ B_r]$ is the balanced truncation of $[A \ B]$. This equality implies that $G_r$ is self-adjoint. By the uniqueness part of Lemma 5 it follows that

$$J^r_{ii} = \frac{\lambda_i^r}{|\lambda_i^r|},$$

where the $\lambda_i^r$ are the Hankel eigenvalues of $G_r$. We conclude that

$$\frac{\lambda_i}{|\lambda_i|} = \frac{\lambda_i^r}{|\lambda_i^r|}.$$ Since for a self-adjoint operator the absolute values of the eigenvalues are the singular values and the Hankel singular values are preserved under balanced truncation, it follows that, in the case considered, the Hankel eigenvalues are preserved under balanced truncation: $\lambda_i^r = \lambda_i$ for $i = 1, \ldots, q$. □

The following result is a specialization of the main result of [4] to the rational case.

Lemma 9. Let $G \in \mathbb{C}(s)^{m \times m}$ be stable and self-adjoint. Denote the Hankel eigenvalues of $G$ by $(\lambda_k)_{k=1}^n$. Then

$$\text{trace}(G(0) - G(\infty)) = 2 \sum_{k=1}^n \lambda_k.$$
Combining Lemma 9 with Lemma 8, we obtain the following.

**Proposition 10.** Let $G \in \mathbb{C}(s)^{m \times m}$ be stable and self-adjoint. Denote the Hankel eigenvalues of $G$ by $(\lambda_k^n)_{k=1}^n$. For $r \in \{1, \ldots, \ell\}$ let $G_r$ be the balanced truncation of $G$, and let $q := \sum_{j=1}^r m_j$, where $(m_j)_{j=1}^r$ denote the multiplicities of the Hankel singular values of $G$. Then

\[
2 \sum_{k=q+1}^n \lambda_k \leq \sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1.
\]

**Proof.** Using that $G(\infty) = G_r(\infty)$ and applying Lemma 9 to both $G$ and $G_r$ we have, with $\lambda_k^r$ the Hankel eigenvalues of $G_r$,

\[
2 \sum_{k=1}^n \lambda_k - 2 \sum_{k=1}^q \lambda_k^r = \text{trace}(G(0) - G_r(0)) - \text{trace}(G_r(0) - G_r(\infty)) = \text{trace}(G(0) - G_r(0)).
\]

By Lemma 8 we have $\lambda_k^r = \lambda_k$ for $k = 1, \ldots, q$, so that the left-hand side of (6) equals

\[
2 \sum_{k=q+1}^n \lambda_k.
\]

Using that the absolute value of the trace is smaller than or equal to the trace class norm, the absolute value of the right-hand side of (6) is smaller than or equal to $\|G(0) - G_r(0)\|_1$. In turn this is smaller than or equal to

\[
\sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1.
\]

We conclude that

\[
2 \sum_{k=q+1}^n \lambda_k \leq \sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1.
\]

\[
\begin{align*}
\sum_{k=q+1}^n \lambda_k &\leq \sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1. \\
&\leq \sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1.
\end{align*}
\]

**V. SYSTEMS WITH A SEMI-DEFINITE HANKEL OPERATOR**

The following is the main result of this article: it shows that for systems with a semi-definite Hankel operator the trace class balanced truncation error bound is in fact an equality.

**Theorem 11.** Let $G \in \mathbb{C}(s)^{n \times m}$ be stable and self-adjoint with a Hankel operator which is either positive semi-definite or negative semi-definite. Denote the Hankel singular values of $G$ by $(\sigma_k)_k=1$. For $r \in \{1, \ldots, \ell\}$ let $G_r$ be the balanced truncation of $G$, and let $q := \sum_{j=1}^r m_j$, where $(m_j)_{j=1}^r$ denote the multiplicities of the Hankel singular values of $G$. Then

\[
\sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1 = 2 \sum_{k=q+1}^n \sigma_k.
\]

Moreover the above also equals $\|G(0) - G_r(0)\|_1$.

**Proof.** That $\sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1 \leq 2 \sum_{k=q+1}^n \sigma_k$ was shown in Theorem 1.

Consider the case where the Hankel operator is positive semi-definite. Then the eigenvalues of the Hankel operator are nonnegative and equal the singular values of the Hankel operator. Proposition 10 gives

\[
2 \sum_{k=q+1}^n \lambda_k \leq \sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1,
\]

where, since $\lambda_k = \sigma_k \geq 0$, the left-hand side equals $2 \sum_{k=q+1}^n \sigma_k$. We conclude that (7) holds. Moreover, the proof of Proposition 10 shows that

\[
2 \sum_{k=q+1}^n \lambda_k \leq \|G(0) - G_r(0)\|_1.
\]

Combining this with (7) gives that in fact equality holds.

If the Hankel operator is negative semi-definite then its eigenvalues are nonpositive and equal to the negatives of the Hankel singular values. The remainder of the argument is as above.

**Example 12.** It is easily seen that a state space symmetric system (that is, $A = A^* \in \mathbb{C}^{n \times n}$ negative definite, $C^* = B \in \mathbb{C}^{n \times m}$, $D = D^* \in \mathbb{C}^{m \times m}$) has a Hankel operator which is positive semi-definite. Therefore Theorem 11 applies to state space symmetric systems.

Systems with $A = A^*$ negative definite, $C^* = -B$ and $D = D^*$ have a Hankel operator which is negative semi-definite and therefore Theorem 11 applies to such systems as well.

The following corollary deals with the operator norm (the Schatten $\infty$-norm).

**Corollary 13.** Let $G \in \mathbb{C}(s)^{n \times m}$ be stable and self-adjoint with a Hankel operator which is either positive semi-definite or negative semi-definite. Denote the singular values of the Hankel operator of $G$ by $(\sigma_k)_k=1$, and the distinct Hankel singular values by $(\mu_j)_{j=1}^\ell$. For $r \in \{1, \ldots, \ell\}$ let $G_r$ be the balanced truncation of $G$ of dimension $q := \sum_{j=1}^r m_j$, where $(m_j)_{j=1}^\ell$ denote the multiplicities of the Hankel singular values. Then

\[
\frac{2}{m} \sum_{k=q+1}^n \sigma_k \leq \sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_\infty \leq 2 \sum_{j=r+1}^\ell \mu_j.
\]

**Proof.** The upper-bound is the standard balanced truncation error bound. For the lower-bound we use that for any $m$-by-$m$ matrix $T$ there holds

\[
\|T\|_1 \leq m \|T\|_\infty.
\]

This gives (using Theorem 11):

\[
\frac{2}{m} \sum_{k=q+1}^n \sigma_k = \frac{1}{m} \sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_1 \leq \sup_{\zeta \in \mathbb{C} \setminus \mathbb{R}_0^+} \|G(\zeta) - G_r(\zeta)\|_\infty.
\]
Remark 14. By choosing \( r = \ell - 1 \), we see in particular from the above corollary that \( \frac{1}{m} m_j \mu \leq 2 \mu \), which is equivalent to \( m \mu \leq m \). Since we can choose any ordering of the distinct Hankel singular values, we obtain that
\[
m_j \leq m, \quad j = 1, \ldots, \ell,
\]
i.e. the multiplicity of a nonzero singular value of a semi-definite Hankel operator is bounded above by the dimension of the input space.

VI. EXTENSIONS

In this section we briefly mention two extensions to the theory presented in this article. The first considers balanced singular perturbation approximation rather than balanced truncation and the second considers the case of non-rational functions.

A. Balanced singular perturbation approximation

Balanced realizations cannot only be used to define the balanced truncation, but also to define the balanced singular perturbation approximation [11]. The theorems presented in this article for the balanced truncation also hold for the balanced singular perturbation approximation. This follows easily using the reciprocal transformation [14]. Define \( G_{\text{recip}} \) by
\[
G_{\text{recip}}(s) := G(1/s).
\]
If \( \begin{bmatrix} A & B \\ C & D \end{bmatrix} \) is a realization of \( G \), then a realization of \( G_{\text{recip}} \) is
\[
\begin{bmatrix}
A^{-1} & -A^{-1}B \\
C & D - CA^{-1}B
\end{bmatrix}.
\]
It is shown in [11] that the reachability and observability maps of the system and its reciprocal are related by
\[
\Phi_{\text{recip}}(\Phi_{\text{recip}}) = \Phi \Phi^* \text{ and } (\Phi_{\text{recip}})^* \Phi_{\text{recip}} = \Psi \Psi^*.
\]
In particular, \( \begin{bmatrix} A & B \\ C & D - CA^{-1}B \end{bmatrix} \) is balanced, if and only if, \( \begin{bmatrix} 0 & -A^{-1}B \\ C & D - CA^{-1}B \end{bmatrix} \) is balanced. It is possible that \( G_{\text{recip}} \) has this same Hankel singular values (with the same multiplicities) as \( G \). It can be furthermore concluded from (5) that self-adjointness of \( G \) (which is clearly equivalent to the self-adjointness of \( G_{\text{recip}} \)), implies that a balanced realization of the reciprocal system fulfills
\[
\begin{bmatrix}
-J & 0 \\
0 & I
\end{bmatrix}
\begin{bmatrix}
A & -A^{-1}B \\
C & D - CA^{-1}B
\end{bmatrix} =
\begin{bmatrix}
A & -A^{-1}B \\
C & D - CA^{-1}B
\end{bmatrix}^*
\begin{bmatrix}
-J & 0 \\
0 & I
\end{bmatrix}.
\]
Lemma 7 then implies that the Hankel eigenvalues of \( G_{\text{recip}} \) are (with the same multiplicities) the negatives of the Hankel eigenvalues of \( G \). In particular, \( G \) has a positive (negative) semi-definite Hankel operator if, and only if, the Hankel operator of \( G_{\text{recip}} \) is negative (positive) semi-definite.

Let \( G_{\text{spa}} \) be the balanced singular perturbation approximation of \( G \) and let \( (G_{\text{recip}})^\ell \) be the balanced truncation of \( G_{\text{recip}} \). Then \( G_{\text{spa}} = (G_{\text{recip}})^\ell(1/s) \), see [14, Figure 1]. Therefore
\[
G(s) - G_{\text{spa}}(s) = G_{\text{recip}}(1/s) - (G_{\text{recip}})^\ell(1/s).
\]
Since \( s \mapsto 1/s \) is a bijection of the open right-half complex plane, we obtain that
\[
\sup_{\zeta \in \mathbb{C} \text{Re} \zeta > 0} \| G(\zeta) - G_{\text{spa}}(\zeta) \| = \sup_{\zeta \in \mathbb{C} \text{Re} \zeta > 0} \| G_{\text{recip}}(\zeta) - (G_{\text{recip}})^\ell(\zeta) \|,
\]
for any matrix norm.

The results in this article applied to the right-hand side then lead to the corresponding results for the left-hand side. The consequence is that we can simply replace \( G_r \) by \( G_{\text{spa}} \) in the statements of the theorems, except for the second part of Theorem 11, where \( \| G(0) - G_r(0) \|_1 \) must be replaced by \( \| G(\infty) - G_{\text{spa}}(\infty) \|_1 \).

B. The non-rational case

The theorems presented in this article continue to hold for non-rational matrix-valued functions as long as the Hankel operator is trace class, i.e. \( \sum_{k=1}^{\infty} \sigma_k < \infty \) (see e.g. [8], [9] for this class of systems). Theorem 1 can be proven as in [9], [8, Section 5.4] taking Schatten class norms instead of operator norms and using the finite-dimensional result given in Theorem 1 instead of the standard balanced truncation error-bound. Lemma 8 can be proven utilizing the discrete-time infinite-dimensional result [5, Theorem 5.1] translated to continuous-time using the usual linear fractional transformation (Cauchy transform) given in e.g. [16] as replacement for the reference to [20, Theorem II]. The remainder of the proofs can remain unchanged.

REFERENCES


