LUND UNIVERSITY

Balanced truncation of linear time-varying systems

Sandberg, Henrik; Rantzer, Anders

Link to publication

Total number of authors:
2

Balanced Truncation of Linear Time-Varying Systems

Henrik Sandberg, *Student Member, IEEE,* and Anders Rantzer, *Fellow, IEEE*

Abstract—In this paper, balanced truncation of linear time-varying systems is studied in discrete and continuous time. Based on relatively basic calculations with time-varying Lyapunov equations/inequalities we are able to derive both upper and lower error bounds for the truncated models. These results generalize well-known time-invariant formulas. The case of time-varying state dimension is considered. Input–output stability of all truncated balanced realizations is also proven. The method is finally successfully applied to a high-order model.

Index Terms—Balanced truncation, error bound, linear time-varying systems, model reduction.

I. INTRODUCTION

THIS paper treats model reduction of time-varying linear systems. Time-varying linear systems are of interest not only for modeling of time-varying physical processes, but also because of the fact that *time-invariant* nonlinear systems can be well approximated by *time-varying* linear systems around nominal trajectories. Linear time-varying systems have attained much attention lately, see for example the survey over periodic systems in [1] and the references therein.

A. Problem Statement

We will assume that a linear system G is given, either in continuous or discrete time. The system should have a finite-dimensional realization with n states. The objective is to find a system $\hat{G}$ with $\hat{n}$ states that approximates G well, where $\hat{n}$ ideally should be much smaller than n. One objective is to find simple candidates $\hat{G}$ for given G and $\hat{n}$. Another objective is to find simple functions $C_1(\cdot)$ and $C_2(\cdot)$, error bounds, such that

$$C_1(\hat{n}) \leq \|G - \hat{G}\| \leq C_2(\hat{n}) \qquad (1)$$

as this simplifies the selection of $\hat{G}$. The operator norm will be the induced 2-norm. Notice that we can always compute $\|G - \hat{G}\|$ to any wanted degree of accuracy once $\hat{G}$ is chosen. However, this is computationally expensive and involves bisection algorithms and solving time-varying Riccati equations, see for instance [2], which is hardly something we would like to do for each candidate $\hat{G}$. So bounds of the type (1) are helpful. Moreover, we would like essential properties of the original system G, such as stability, to be preserved for each candidate $\hat{G}$.

B. Previous Work

To reduce the order of linear time-invariant systems, balanced truncation is often used. Balanced realizations were introduced in [3], but were first used for the purpose of model reduction in [4]. A sufficient condition for asymptotic stability of truncated models was later given in [5]. Since then an error bound has been proven [6], [7], which gives a simple bound on the worst case error between the original and truncated model and justifies the approximation. The bound was first derived for continuous-time systems, but it also holds for discrete-time systems as proven in [8]. The bound is a sum of truncated Hankel singular values and the result is now considered to be standard and is included in most courses on robust control and identification.

Balanced realizations for time-varying linear systems have also received attention, see for example [9], [10], for some early references. For the related class of linear parameter-varying (LPV) systems, balanced truncation has been studied in for example [11]. However, until recently no error bound has been given for the time-variable case. To obtain bounds, methods for uncertain systems could be utilized, see for example [12]. However, these bounds would be conservative as the known time-variance is encapsulated in an uncertainty ball.

The first explicit error bound for balanced discrete time-varying models, to the authors' best knowledge, was given in [13] and later refined in [14]. There, an operator-theoretic framework was used to give bounds similar to those that apply to time-invariant models. For discrete time-periodic linear systems bounds have been proven in [15], [16]. There, a special form of lifting isomorphism was used.

C. Contributions of This Paper

In this paper, we will work directly with the time-varying observability and reachability Lyapunov inequalities [linear matrix inequalities (LMIs)] in both continuous and discrete time. It will be seen that it is natural to allow the state-space dimension to vary in size over time. In fact, a time-varying state dimension may be required for a minimal realization as is shown and used in for example [16] and [17]. The approach will give fairly simple calculations and more general error bounds (1) than in the previously mentioned references. In particular we will allow for time-varying Gramians, which is not treated in [13] and [14]. As special cases we will recover the known bounds for both time-invariant and time-varying systems. Furthermore, the method will give new results on input–output stability of the reduced models.

The ability to vary the state-space dimension over time is not only of interest for technical reasons. In for example stiff problems, such as chemical reactions, it is frequent that in the initial phase, many complex reactions take place and that the dynamics then slows down. It is then reasonable to have a model with

many states in the initial phase and then switch to a low-order model after some time. The analysis presented will help to decide when to switch the number of states and also how much loss in accuracy a certain choice might give.

D. Organization

The organization of the paper is as follows. In Sections II and III, notation for discrete and continuous-time systems will be introduced, along with two lemmas on observability and reachability. The lemmas will form the basis of the following analysis. In Section IV, we will define what a balanced system is and how we, with the help of the lemmas, can attain simple upper error bounds. In Section V, input–output stability of all truncated models is proved. In Section VI, a lower error bound for truncated models is given. In Section VII, an example of how balanced model truncation works in practice is given. In Appendix II, it is shown how sampling of a continuous-time system can be combined with model truncation.

II. DISCRETE-TIME SYSTEMS

As some aspects of the calculations are simpler for discrete-time systems, we will start at that end. It should, however, be pointed out that everything presented here will later also be done for continuous-time systems.

A. Preliminaries and Notation

The linear systems G that we consider are assumed to have a finite-dimensional state-space realization

$$G: \begin{cases} x(k+1) = A(k)x(k) + B(k)u(k) & x(0) = 0 \\ y(k) = C(k)x(k) + D(k)u(k) \end{cases} \quad (2)$$

with m inputs and p outputs. It will be useful to utilize time-varying state-space dimension as commented in the introduction. It is known that minimal realizations of linear systems in general have this property; see [17]. However, it will also be a useful technical tool for reducing the order of systems where the state-space dimension originally is constant over time. Let the state-space dimension at time k be $n(k)$. The signals and matrices then have the dimensions

$$
\begin{aligned}
A(k) &\in \boldsymbol{R}^{n(k+1)\times n(k)} \\
B(k) &\in \boldsymbol{R}^{n(k+1)\times m} \\
x(k) &\in \boldsymbol{R}^{n(k)} \\
C(k) &\in \boldsymbol{R}^{p\times n(k)} \\
D(k) &\in \boldsymbol{R}^{p\times m} \\
y(k) &\in \boldsymbol{R}^{p} \\
u(k) &\in \boldsymbol{R}^{m}.
\end{aligned}
$$

We will assume that all the matrices are real, bounded, and defined for $k \in [0, T]$. Sometimes we will have $T = +\infty$, and then the system is assumed to be stable. Notice that as the model order may vary with k, $A(k)$ is not necessarily a square matrix but rather rectangular. We could also let the number of inputs and outputs vary over time, but we avoid this for the sake of notational simplicity.

The signals will belong to the Hilbert space $\ell_2[0, T]$. We will utilize the weighted Euclidean norm as defined by

$$|x(k)|_P^2 = x^T(k)P(k)x(k)$$

with a positive–definite matrix $P(k) \in \boldsymbol{R}^{n(k)\times n(k)}$, and also the weighted ℓ_2-norm

$$\|x\|_P^2 = \sum_{k=0}^{T} |x(k)|_P^2. \quad (3)$$

Discrete-time signals x over a time interval $[0, \infty)$ belong to $\ell_2^n[0, \infty)$ iff the norm (3) is finite for $P(k) = I$ with $T = +\infty$. If we want to emphasize that the norm is taken over the interval $[0, T]$, we will write $\|x\|_{P,[0,T]}$, but the interval will normally be clear from the context. Linear systems as defined in (2) can be identified with a linear operator $G : \ell_2^m[0, T] \to \ell_2^p[0, T]$. The operator is *bounded* iff the induced norm

$$\|G\| = \sup_{\|u\|\leq 1} \|Gu\|$$

is bounded. Often we will make an upper estimate of $\|G\|$ by finding a constant $C(G) > 0$ such that

$$\|y\| \leq C(G) \cdot \|u\|$$

for all admissible u.

The system we would like to obtain, $\hat{G}$, will be called a reduced-order system. It will have the state-space dimension $\hat{n}(k)$ where $\hat{n}(k) \leq n(k)$ for all k. We will construct $\hat{G}$ from a truncation of the realization of G. The following partitions will be used:

$$
\begin{aligned}
A(k) &= \begin{bmatrix} A_{11}(k) & A_{12}(k) \\ A_{21}(k) & A_{22}(k) \end{bmatrix} & A_{11}(k) &\in \boldsymbol{R}^{\hat{n}(k+1)\times \hat{n}(k)} \\
B(k) &= \begin{bmatrix} B_1(k) \\ B_2(k) \end{bmatrix} & B_1(k) &\in \boldsymbol{R}^{\hat{n}(k+1)\times m} \\
C(k) &= \begin{bmatrix} C_1(k) & C_2(k) \end{bmatrix} & C_1(k) &\in \boldsymbol{R}^{p\times \hat{n}(k)} \\
x^T(k) &= \begin{bmatrix} x_1^T(k) & x_2^T(k) \end{bmatrix} & x_1(k) &\in \boldsymbol{R}^{\hat{n}(k)}.
\end{aligned}
$$

If the realization (2) is chosen such that the states $x_2(k)$ are "small" in some sense, a reasonable reduced-order candidate is obtained by truncating the corresponding states

$$\hat{G}: \begin{cases} \hat{x}(k+1) = A_{11}(k)\hat{x}(k) + B_1(k)u(k) & \hat{x}(0) = 0 \\ \hat{y}(k) = C_1(k)\hat{x}(k) + D(k)u(k) & \hat{x}(k) \in \boldsymbol{R}^{\hat{n}(k)}. \end{cases} \quad (4)$$

The auxiliary signal

$$\hat{z}(k+1) = A_{21}(k)\hat{x}(k) + B_2(k)u(k) \quad (5)$$

will naturally show up later. It is not needed to evaluate the map $\hat{G}$. $\hat{z}(k)$ has dimension $\boldsymbol{R}^{n(k)-\hat{n}(k)}$ and is defined when truncation has occurred, i.e., $\hat{n}(k) < n(k)$. As $\hat{z}$ is not necessarily defined for all k, it will be useful to collect the time points where it does exist in a set $\mathcal{T}$

$$\mathcal{T} = \{k: \hat{z}(k) \text{ exists}\}. \quad (6)$$

Furthermore, let us define $\hat{z}(0) = 0$ if $\hat{n}(0) < n(0)$.

If the systems G and $\hat{G}$ are supposed to have a similar input–output behavior when the above truncation scheme is used, it is important that the coordinate system in the realization of G is well chosen. As we will see, such coordinate

systems exist in many cases. A change in coordinate system, $x(k) = T(k)\hat{x}(k)$, for invertible $T(k)$, will transform the realization according to

$$\{A(k), B(k), C(k), D(k)\}$$
$$\xrightarrow{T(k)} \{\tilde{A}(k), \tilde{B}(k), \tilde{C}(k), \tilde{D}(k)\}$$
$$= \{T^{-1}(k+1)A(k)T(k),$$
$$T^{-1}(k+1)B(k), C(k)T(k), D(k)\}. \qquad (7)$$

B. The Observability Lyapunov Inequality

Consider the *Lyapunov observability inequality*

$$A^T(k)Q(k+1)A(k) + C^T(k)C(k) \leq Q(k), \qquad k \in [0, T]. \qquad (8)$$

$Q(k)$ is often called an *observability Gramian*. The positive–semidefinite solutions $Q(k)$, $k = 0 \ldots T + 1$, bound the amount of energy there will be in the output for a given initial state $x(0)$ of the system G with zero input

$$|x(T+1)|^2_Q + \|y\|^2_{[0,T]} \leq |x(0)|^2_Q.$$

The inequality can, however, also be used to calculate the ℓ_2-norm of the difference in the outputs from G and $\hat{G}$ when both systems are driven by the same input signal. To see this, assume there is a positive–semidefinite solution $Q(k)$ to (8) with the block-diagonal structure

$$Q(k) = \begin{bmatrix} Q_1(k) & 0 \\ 0 & q(k) \cdot I_{n(k)-\hat{n}(k)} \end{bmatrix} \in \mathbf{R}^{n(k) \times n(k)} \qquad (9)$$

for $k = 0 \ldots T + 1$ and $q(k)$ scalar. Then rewrite (8) for each k in the following way:

$$\begin{bmatrix} A(k) \\ I \end{bmatrix}^T \begin{bmatrix} Q(k+1) & 0 \\ 0 & -Q(k) \end{bmatrix} \begin{bmatrix} A(k) \\ I \end{bmatrix}$$
$$+ C^T(k)C(k) \leq 0. \qquad (10)$$

If we apply the same input signal u to (2) and (4) we obtain the trajectories x and $\hat{x}$. Use the trajectories to calculate the difference

$$\begin{bmatrix} x_1(k) - \hat{x}(k) \\ x_2(k) \end{bmatrix} \in \mathbf{R}^{n(k)}.$$

Multiply (10) for each k from the right with the difference and from the left with its transpose. We then obtain

$$\begin{bmatrix} x_1(k+1) - \hat{x}(k+1) \\ x_2(k+1) - \hat{z}(k+1) \\ x_1(k) - \hat{x}(k) \\ x_2(k) \end{bmatrix}^T \begin{bmatrix} Q(k+1) & 0 \\ 0 & -Q(k) \end{bmatrix}$$
$$\times \begin{bmatrix} x_1(k+1) - \hat{x}(k+1) \\ x_2(k+1) - \hat{z}(k+1) \\ x_1(k) - \hat{x}(k) \\ x_2(k) \end{bmatrix} + |y(k) - \hat{y}(k)|^2 \leq 0$$

which is the same as

$$\Delta \left\| \begin{bmatrix} x_1(k) - \hat{x}(k) \\ x_2(k) \end{bmatrix} \right\|^2_Q - 2q(k+1)\hat{z}^T(k+1)x_2(k+1)$$
$$+ |\hat{z}(k+1)|^2_q + |y(k) - \hat{y}(k)|^2 \leq 0 \qquad (11)$$

using the structure (9) of $Q(k)$. The forward difference operator Δ is defined as

$$\Delta r(k) = r(k+1) - r(k)$$

on a scalar sequence $\{r(k)\}$. Now we can state the following lemma:

Lemma 1 (Observability): If there is a solution $Q(k)$ with the structure (9) to the Lyapunov inequality (8) on the interval $[0, T+1]$, then the solutions of (2) and (4) satisfy the following.
  i)

$$\left\| \begin{bmatrix} x_1(T+1) - \hat{x}(T+1) \\ x_2(T+1) \end{bmatrix} \right\|^2_Q + \|y - \hat{y}\|^2_{[0,T]}$$
$$+ \sum_{k \in \mathcal{T}} \left(|\hat{z}(k)|^2_q - 2q(k)\hat{z}^T(k)x_2(k) \right) \leq 0 \qquad (12)$$

  where equality holds if (8) was solved with equality.
  ii) For every nonincreasing positive scalar sequence $\{a(k)\}^T_{k=0}$, we have

$$\|y - \hat{y}\|^2_{a,[0,T]} - \sum_{k \in \mathcal{T}} a(k-1)2q(k)\hat{z}^T(k)x_2(k) \leq 0. \qquad (13)$$

Proof:
  i) Sum the inequalities (11) over the interval $k = 0 \ldots T$ and notice the cancelling terms.
  ii) Multiply (11) with $a(k)$ for each k, and sum over $k = 0 \ldots T$. For nonincreasing $a(k)$ the partially cancelling terms become nonnegative numbers. The sum over $\mathcal{T}$ is the only sign-indefinite term, which leads to the inequality (13). ∎

As seen if $\mathcal{T} = \emptyset$ the difference in output is zero, as $G = \hat{G}$. All terms in (12) are necessarily nonnegative except the terms $\hat{z}^T(k)x_2(k)$. These terms are the price we pay for truncating states. One might think that if the numbers $q(k)$ are small for $k \in \mathcal{T}$, then $\|y - \hat{y}\|$ will be small. Indeed, if the states $x_2(k)$ are unobservable there is a solution $Q(k)$ such that $q(k) = 0$ and $\|y - \hat{y}\| = 0$. Thus a small $q(k)$ could indicate that k should be included in the set $\mathcal{T}$ and that the corresponding states $x_2(k)$ should be truncated. However, we should remember that $q(k)$ is only a weight. A sufficient condition for a small $\|y - \hat{y}\|^2$ is that $|x_2(k)|^2_q$ is small for all k in $\mathcal{T}$. This can be seen by completing the squares in the sum (12). Then we see that $\|y - \hat{y}\|^2$ is bounded by $\sum_{k \in \mathcal{T}} |x_2(k)|^2_q$. However, this is not a bound of the type (1). To obtain such a bound we will make a dual analysis, which is the topic of Section II-C.

C. Reachability Lyapunov Inequality

Here, it will be seen how far away the states in G and $\hat{G}$ can be forced with the input signal u. The following inequality will be called the *Lyapunov reachability inequality*:

$$A(k)P(k)A^T(k) + B(k)B^T(k) \leq P(k+1), \qquad k \in [0, T]. \qquad (14)$$

$P(k)$ is often called a *reachability Gramian*. Assume there is a positive–definite block-diagonal solution to (14)

$$P(k) = \begin{bmatrix} P_1(k) & 0 \\ 0 & p(k) \cdot I_{n(k)-\hat{n}(k)} \end{bmatrix} \in \mathbf{R}^{n(k) \times n(k)} \quad (15)$$

with $k = 0 \ldots T+1$ and $p(k)$ scalar. Notice that (14) is equivalent to

$$\begin{bmatrix} A(k) & B(k) \\ I & 0 \end{bmatrix}^T \begin{bmatrix} P^{-1}(k+1) & 0 \\ 0 & -P^{-1}(k) \end{bmatrix}$$
$$\times \begin{bmatrix} A(k) & B(k) \\ I & 0 \end{bmatrix} \leq \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}. \quad (16)$$

Now, assume we apply the same input signal u to G and $\hat{G}$. We then obtain the system trajectories x and $\hat{x}$. Multiply (16) for each k with

$$\begin{bmatrix} x_1(k) + \hat{x}(k) \\ x_2(k) \\ 2u(k) \end{bmatrix} \in \mathbf{R}^{n(k)+m}$$

from the right and with its transpose from the left. This gives

$$\Delta \left\| \begin{bmatrix} x_1(k) + \hat{x}(k) \\ x_2(k) \end{bmatrix} \right\|_{P^{-1}}^2 + 2p^{-1}(k+1)\hat{z}^T(k+1)x_2(k+1)$$
$$+ |\hat{z}(k+1)|_{p^{-1}}^2 \leq 4|u(k)|^2 \quad (17)$$

if the structure of $P(k)$ is used. Now, the following lemma can be stated.

Lemma 2 (Reachability): If there is a solution $P(k)$ to the inequality (14) with the structure (15) on the interval $[0, T+1]$ then the solutions to (2) and (4) satisfy the following.

i)

$$\left\| \begin{bmatrix} x_1(T+1) + \hat{x}(T+1) \\ x_2(T+1) \end{bmatrix} \right\|_{P^{-1}}^2$$
$$+ \sum_{k \in \mathcal{T}} \left(|\hat{z}(k)|_{p^{-1}}^2 + 2p^{-1}(k)\hat{z}^T(k)x_2(k) \right) \leq 4\|u\|_{[0,T]}^2. \quad (18)$$

ii) For every positive nonincreasing scalar sequence $\{b(k)\}_{k=0}^T$

$$\sum_{k \in \mathcal{T}} b(k-1)2p^{-1}(k)\hat{z}^T(k)x_2(k) \leq 4\|u\|_{b,[0,T]}^2. \quad (19)$$

Proof: As in Lemma 1. Use (17) instead of (11). ∎

The lemma gives boundaries on the reachable set in the state–space for fixed amounts of input energy. Notice that when $\mathcal{T} = \emptyset$ (18) reduces to the well-known result

$$|x(T+1)|_{P^{-1}}^2 \leq \|u\|_{[0,T]}^2 + |x(0)|_{P^{-1}}^2$$

as $x(k) = \hat{x}(k)$ for all k. Also notice that the sum in (19) potentially can cancel the sum in (13), namely if

$$a(k-1)q(k) = b(k-1)p^{-1}(k) \quad (20)$$

for all $k \in \mathcal{T}$. We have obtained a bound on the terms $\hat{z}^T(k)x_2(k)$ and this will be utilized Section IV.

As we will utilize the truncation recursively in the following it is convenient that the realization of $\hat{G}$, $\{A_{11}(k), B_1(k), C_1(k), D(k)\}$, fulfills the Lyapunov inequalities (8) and (14), with $Q_1(k)$ and $P_1(k)$ respectively. This can be seen from straightforward calculations.

III. CONTINUOUS-TIME SYSTEMS

The previous ideas in discrete time goes through in continuous time without much alternation. However, we have to be somewhat careful when the number of states change over time.

A. Preliminaries and Notation

The linear operator G will now operate on the Hilbert space $L_2[0, T]$, that is $G : L_2^m[0, T] \to L_2^p[0, T]$. A measurable signal x belongs to $L_2^n[0, T]$ iff the norm

$$\|x\|_P^2 = \int_0^T |x(t)|_P^2 dt$$

is finite for $P(t) = I$. The norm $\|G\|$ is the standard induced norm. We assume there is a finite-dimensional realization of G

$$G: \begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) & x(0) = 0 \\ y(t) = C(t)x(t) + D(t)u(t). \end{cases} \quad (21)$$

The matrices and signals have the same dimensions as in discrete time, we will for now assume that the state dimension is $n(t) = n$ and is constant over time. We will assume that the matrices are continuous and bounded over time in all their entries. With these conditions existence and uniqueness of solutions to (21) is guaranteed, see for example [18]. When the infinite time-horizon case is studied the system is assumed to be stable.

If we use the same matrix partitions as before we can define the $\hat{n}$th-order reduced-order system $\hat{G}$

$$\hat{G}: \begin{cases} \dot{\hat{x}}(t) = A_{11}(t)\hat{x}(t) + B_1(t)u(t) & \hat{x}(0) = 0 \\ \hat{y}(t) = C_1(t)\hat{x}(t) + D(t)u(t). \end{cases} \quad (22)$$

The auxiliary error signal $\hat{z} \in \mathbf{R}^{n-\hat{n}}$ becomes

$$\hat{z}(t) = A_{21}(t)\hat{x}(t) + B_2(t)u(t). \quad (23)$$

As we assume constant state dimension for now, the set $\mathcal{T}$ is the interval $[0, T]$.

Coordinate transformations $x(t) = T(t)\tilde{x}(t)$ with a continuously differentiable $T(t)$, nonsingular for all t, gives

$$\{A(t), B(t), C(t), D(t)\}$$
$$\xrightarrow{T(t)} \{\tilde{A}(t), \tilde{B}(t), \tilde{C}(t), \tilde{D}(t)\}$$
$$= \{T^{-1}(t)[A(t)T(t) - \dot{T}(t)], T^{-1}(t)B(t)$$
$$C(t)T(t), D(t)\} \quad (24)$$

so that the input–output map is invariant.

B. Observability Lyapunov Inequality

The *observability Lyapunov inequality* takes the form

$$Q(t)A(t) + A^T(t)Q(t) + \dot{Q}(t) + C^T(t)C(t) \leq 0 \quad (25)$$

in continuous time. We can perform the same analysis as in Section II-B by noting that (25) can be written as

$$\begin{bmatrix} A(t) \\ I \end{bmatrix}^T \begin{bmatrix} 0 & Q(t) \\ Q(t) & \dot{Q}(t) \end{bmatrix} \begin{bmatrix} A(t) \\ I \end{bmatrix} + C^T(t)C(t) \leq 0. \quad (26)$$

As in Section II-B we get:

Lemma 3 (Observability): If there is a solution $Q(t)$ with the structure (9) to the Lyapunov inequality (25) on the interval $[0, T]$, then the solutions of (21) and (22) satisfy the following.

i)

$$\left\| \begin{bmatrix} x_1(T) - \hat{x}(T) \\ x_2(T) \end{bmatrix} \right\|_Q^2 + \|y - \hat{y}\|^2$$
$$- \int_0^T 2q(t)\hat{z}^T(t)x_2(t)dt \le 0 \quad (27)$$

where equality holds if (25) was solved with equality.

ii) For every nonincreasing positive continuous scalar $a(t)$ we have

$$\|y - \hat{y}\|_a^2 - \int_0^T a(t)2q(t)\hat{z}^T(t)x_2(t)dt \le 0. \quad (28)$$

Proof: As Lemma 1 but use (26) instead of (10). Replace summation with integration. ∎

C. Reachability Lyapunov Inequality

The *reachability Lyaponov inequality* takes the form

$$A(t)P(t) + P(t)A^T(t) - \dot{P}(t) + B(t)B^T(t) \le 0 \quad (29)$$

in continuous time. If there is a positive–definite solution $P(t)$, (29) is equivalent to

$$\begin{bmatrix} A(t) & B(t) \\ I & 0 \end{bmatrix}^T \begin{bmatrix} 0 & P^{-1}(t) \\ P^{-1}(t) & \frac{d}{dt}P^{-1}(t) \end{bmatrix}$$
$$\times \begin{bmatrix} A(t) & B(t) \\ I & 0 \end{bmatrix} \le \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}. \quad (30)$$

The analog to Lemma 2 becomes the following.

Lemma 4 (Reachability): If there is a solution $P(t)$ to the inequality (29) with the structure (15) on the interval $[0, T]$ then the solutions to (21) and (22) satisfy the following.

i)

$$\left\| \begin{bmatrix} x_1(T) + \hat{x}(T) \\ x_2(T) \end{bmatrix} \right\|_{P^{-1}}^2$$
$$+ \int_0^T 2p^{-1}(t)\hat{z}^T(t)x_2(t)dt \le 4\|u\|^2. \quad (31)$$

ii) For every positive nonincreasing continuous scalar $b(t)$

$$\int_0^T b(t)2p^{-1}(t)\hat{z}^T(t)x_2(t)dt \le 4\|u\|_b^2. \quad (32)$$

Proof: As Lemma 2. Use (30) instead of (16). Replace summation with integration. ∎

D. Continuous-Time Systems With Time-Varying State Dimension

It is possible to analyze systems where the state dimension varies over time, i.e., $n(t)$ takes integer values but changes with time. This will be useful in Section IV as we then do not need to distinguish between discrete- and continuous-time systems.

Assume that G has n states and that $\hat{G}$ has $\hat{n}_1$ states until time t^-, and then switches to $\hat{n}_2$ states at t^+, i.e., an instant switch. The question is what to do with new states and also with the ones that disappear. Furthermore, are Lemmas 3 and 4 still valid?

From t^- to t^+, the control signal u will not have time to influence the states as the input energy becomes zero on this interval of zero measure. The dynamics of the original system G becomes

$$x(t^+) = A^J x(t^-) \quad A^J = I_n$$

i.e., nothing happens with the states. The truncated realizations $A_{11}^J \in \mathbf{R}^{\hat{n}_2 \times \hat{n}_1}$ become

$$A_{11}^J = \begin{bmatrix} I_{\hat{n}_1} \\ 0 \end{bmatrix} \quad \hat{n}_2 > \hat{n}_1 \quad \text{or}$$
$$A_{11}^J = [I_{\hat{n}_1} \quad 0] \quad \hat{n}_2 < \hat{n}_1. \quad (33)$$

So, new states should just be initialized to zero. If there are continuous solutions $Q(t)$ and $P(t)$ to the inequalities (25) and (29) we can readily use them as solutions to the discrete-time Lyapunov equations for the jump

$$\left(A^J\right)^T Q\left(t^+\right) A^J = Q\left(t^-\right) \quad \text{and}$$
$$A^J P\left(t^-\right) \left(A^J\right)^T = P\left(t^+\right) \quad (34)$$

which are fulfilled with $Q\left(t^-\right) = Q\left(t^+\right) = Q(t)$ and $P\left(t^-\right) = P\left(t^+\right) = P(t)$. For each jump, we therefore get the following addition to Lemma 3:

$$\left\| \begin{bmatrix} x_1\left(t^+\right) - \hat{x}\left(t^+\right) \\ x_2\left(t^+\right) \end{bmatrix} \right\|_Q^2 - \left\| \begin{bmatrix} x_1\left(t^-\right) - \hat{x}\left(t^-\right) \\ x_2\left(t^-\right) \end{bmatrix} \right\|_Q^2$$
$$+ |\hat{z}\left(t^+\right)|_q^2 - 2q\left(t^+\right)\hat{z}^T\left(t^+\right) x_2\left(t^+\right) = 0$$

with $\hat{x}(t^-) \in \mathbf{R}^{\hat{n}_1}$ and $\hat{x}(t^+) \in \mathbf{R}^{\hat{n}_2}$. The two first terms get canceled by the boundary terms of the integrals from the constant state modes before and after the switch in the lemma. So, the only real contribution is the two last terms. For Lemma 4, the additions become

$$\left\| \begin{bmatrix} x_1\left(t^+\right) + \hat{x}\left(t^+\right) \\ x_2\left(t^+\right) \end{bmatrix} \right\|_{P^{-1}}^2 - \left\| \begin{bmatrix} x_1\left(t^-\right) + \hat{x}\left(t^-\right) \\ x_2\left(t^-\right) \end{bmatrix} \right\|_{P^{-1}}^2$$
$$+ |\hat{z}\left(t^+\right)|_{p^{-1}}^2 + 2p^{-1}\left(t^+\right)\hat{z}^T\left(t^+\right) x_2\left(t^+\right) \le 0.$$

Again, the only real contribution is the two last terms. The remaining sign-indefinite terms

$$q\left(t^+\right)\hat{z}^T\left(t^+\right) x_2\left(t^+\right) \text{ and } p^{-1}\left(t^+\right)\hat{z}^T\left(t^+\right) x_2\left(t^+\right)$$

can be canceled by proper choice of $a(t)$ and $b(t)$ as will be discussed in Section IV.

The conclusion is that if the jump transitions (33) are used there is no real change to the results in Lemmas 3 and 4 and the set $\mathcal{T}$ can be defined exactly as in the discrete-time case, (6), and we may replace the integrals $\int_0^T$ in Lemmas 3 and 4 by $\int_{\mathcal{T}}$.

Remark 1 (Discontinuities in $\hat{y}$): With the proposed scheme we see that when new states are added, i.e., $\hat{n}_2 > \hat{n}_1$, $\hat{y}$ will be continuous at the switching instant as the new states are initialized to zero. Moreover $\hat{z}(t^+)$ is zero. In the other case when $\hat{n}_2 < \hat{n}_1$, $\hat{y}$ can be discontinuous at the switching instant as states are thrown away, and then $|\hat{z}(t^+)|^2 > 0$.

Remark 2 (Discontinuous State Transformations): The technique here can also be used when one, at some time instant, would like to make an instantaneous state transformation, i.e., $T(t)$ is discontinuous. Then the jump transition matrix A^J becomes the solution to

$$T\left(t^+\right) A^J = T\left(t^-\right)$$

and all the calculations in this section can be redone with this jump matrix A^J. The corresponding Lyapunov equations to (34) become

$$\left(A^J\right)^T \tilde{Q}^+\left(t^+\right) A^J = \tilde{Q}^-\left(t^-\right) \text{ and}$$
$$A^J \tilde{P}^-\left(t^-\right)\left(A^J\right)^T = \tilde{P}^+\left(t^+\right).$$

$\tilde{\cdot}^-$ and $\tilde{\cdot}^+$ denote matrices given in the coordinate systems $T\left(t^-\right)$ and $T\left(t^+\right)$, respectively. How the solutions $P(t)$ and $Q(t)$ are transformed is discussed in Section IV, (37).

IV. BALANCED REALIZATIONS AND ERROR BOUNDS

Sections II and III rely heavily on the ability to obtain block-diagonal solutions to the inequalities (8), (14), (25), and (29), respectively. Often this is possible to obtain. In particular, if there are any solutions $P(\cdot) > 0$ and $Q(\cdot) > 0$ for all time instants in some realization of G, then there exists a *balanced realization* of G where the Lyapunov inequalities take the form

$$\tilde{A}^T(k)\Sigma(k+1)\tilde{A}(k) - \Sigma(k) + \tilde{C}^T(k)\tilde{C}(k) \leq 0$$
$$\tilde{A}(k)\Sigma(k)\tilde{A}^T(k) - \Sigma(k+1) + \tilde{B}(k)\tilde{B}^T(k) \leq 0 \quad (35)$$

in discrete time, and in continuous time with some extra regularity conditions

$$\Sigma(t)\tilde{A}(t) + \tilde{A}^T(t)\Sigma(t) + \dot{\Sigma}(t) + \tilde{C}^T(t)\tilde{C}(t) \leq 0$$
$$\tilde{A}(t)\Sigma(t) + \Sigma(t)\tilde{A}^T(t) - \dot{\Sigma}(t) + \tilde{B}(t)\tilde{B}^T(t) \leq 0 \quad (36)$$

with the diagonal solution (balanced Gramians)

$$\Sigma(\cdot) = \tilde{P}(\cdot) = \tilde{Q}(\cdot) = \text{diag}\{\sigma_1(\cdot), \sigma_2(\cdot), \ldots, \sigma_n(\cdot)\} > 0.$$

A linear system G with a realization fulfilling (35) or (36) with a Gramian Σ is called a *balanced system*. σ_i will be denoted as the singular value corresponding to the state x_i in a particular balanced system. Notice that it is always possible to permute the singular values in Σ. Normally one chooses to put the elements in descending order so that

$$\sigma_1(\cdot) \geq \sigma_2(\cdot) \geq \cdots \geq \sigma_n(\cdot) > 0.$$

As the singular values change in size over time it may be that the ordering must be changed at some time instants to maintain the aforementioned order. This can be done with an instantaneous coordinate transformation (permutation), see Remark 2 in Section III-D. However, as we will see, the ordering is not critical to our discussion. But in general it makes good sense to put small singular values last in the Σ-matrix. By defining a balanced realization with inequalities instead of equalities it becomes nonunique, and the singular values are nonunique. This was introduced in [12] and [19], and has several appealing properties including the possibility of tighter error bounds and that every truncated realization remains balanced.

If we have solutions $Q(\cdot)$ and $P(\cdot)$ in a given coordinate system we can obtain the needed coordinate transformation $T(\cdot)$ to obtain a balanced realization. This is the topic of many papers in discrete time; see, for example, [16], [20], and the references therein. In continuous time, we need regularity conditions on the realization to guarantee the existence of a well-behaved balancing transformation. In [10], for instance, analyticity of the realization is assumed. In [9] and [21], uniform observability

and controllability is assumed. How in practice to obtain $T(t)$ in continuous time is not obvious, as we need $T(t)$ and also $\dot{T}(t)$ on an interval. Pointwise we can always obtain a $T(\cdot)$ as we will see. An approximate approach to obtain $T(t)$ over an interval is presented in [22].

We will not go into much detail here, as this is done in the references previously mentioned, let us just notice that under the coordinate transformation (7) and (24) the solutions to the Lyapunov inequalities transform as

$$\tilde{Q}(\cdot) = T^T(\cdot)Q(\cdot)T(\cdot)$$
$$\tilde{P}(\cdot) = T^{-1}(\cdot)P(\cdot)T^{-T}(\cdot) \quad (37)$$

so that the eigenvalues of their product is invariant. Therefore, we can calculate the singular values for a realization with Gramians P and Q as

$$\sigma_i^2(\cdot) = \lambda_i(P(\cdot)Q(\cdot)) = \lambda_i(\tilde{P}(\cdot)\tilde{Q}(\cdot))$$

at each time-instant and also obtain a balancing coordinate system $T(\cdot)$.

As a first step toward error-bounds for truncated balanced realizations let us note that from Lemmas 1 and 2 and Lemmas 3 and 4 we get the following bound:

Proposition 1 (Cancelling Condition): If the nonincreasing weights $a(\cdot)$ and $b(\cdot)$ are chosen so that for all time-instants k or t in $\mathcal{T}$

$$a(k-1)q(k) = b(k-1)p^{-1}(k) \quad \text{(Discrete time)}$$
$$a(t)q(t) = b(t)p^{-1}(t) \quad \text{(Continuous time)} \quad (38)$$

then

$$\|y - \hat{y}\|_a \leq 2\|u\|_b. \quad (39)$$

Proof: Add Lemma 1 ii) with Lemma 2 ii) and notice that the sign-indefinite terms are canceled if $a(k)$ and $b(k)$ fulfill the previous condition. Analogous in continuous time. ∎

A. Monotonically Balanced Systems

We will proceed by formulating an error bound for truncated balanced realizations which looks familiar to the well-known time-invariant result in [6] and [7]. We will first look at balanced systems where the singular values are monotonic in time, as this is the simplest nontime-invariant case. It is useful to group equal singular values together as this makes the error bound sharper. If there are $N(\cdot)$ unique singular values use the notation

$$\Sigma(\cdot) = \text{diag}\{\sigma_1(\cdot)I_{s_1}, \ldots, \sigma_N(\cdot)I_{s_N}\}$$

where $s_1(\cdot) + \cdots + s_N(\cdot) = n(\cdot)$. Now, the following result is easily obtained.

Theorem 1 (Monotonically Balanced Systems): Suppose the system G has a balanced realization on the interval $[0, T]$ with $\Sigma(\cdot) = \text{diag}\{\Sigma_1(\cdot), \Sigma_2(\cdot)\}$

$$\Sigma_1(\cdot) = \text{diag}\{\sigma_1(\cdot)I_{s_1}, \ldots, \sigma_r(\cdot)I_{sr}\}$$
$$\Sigma_2(\cdot) = \text{diag}\{\sigma_{r+1}(\cdot)I_{s(r+1)}, \ldots, \sigma_N(\cdot)I_{sN}\}$$

where each singular value $\sigma_i(\cdot)$, $i = r + 1 \ldots N$ is either nonincreasing or nondecreasing over time.

The truncated $(s_1 + \cdots + s_r)$-order system $\hat{G}$ is then balanced by $\Sigma_1(\cdot)$ and

$$\left\| G - \hat{G} \right\| \leq 2 \sum_{i=r+1}^{N} \sup_{t \in [0,T]} \sigma_i(t). \tag{40}$$

Proof: Start by removing the states with the singular value σ_N, and call this truncated system $\hat{G}_1$. Thus put $p = q = \sigma_N$. By assumption there are two possibilities: σ_N is nonincreasing or nondecreasing. First, consider the nonincreasing case. Then, choose $b(t) = \sigma_N^2(t)$ and $a(t) = 1$ in Proposition 1 ($\mathcal{T} = [0,T]$, use $a(k-1)$ and $b(k-1)$ in discrete time) and notice that the cancelling condition is fulfilled. In the nondecreasing case, choose $a = \sigma_N^{-2}$ and $b = 1$. It follows that

$$\|y - \hat{y}_1\| \leq 2 \|u\|_{\sigma_N^2}, \text{ or}$$
$$\|y - \hat{y}_1\|_{\sigma_N^{-2}} \leq 2 \|u\|$$

which leads to $\left\| G - \hat{G}_1 \right\| \leq 2 \sup_t \sigma_N(t)$. Next notice that $\hat{G}_1$ is still balanced with the rest of Σ, Σ_{N-1}. We proceed iteratively and remove σ_{N-1} from $\hat{G}_1$, and repeat the scheme until the system $\hat{G} = \hat{G}_{N-r}$ is reached. Finally use the triangular inequality:

$$\left\| G - \hat{G} \right\| = \left\| G - \hat{G}_1 + \hat{G}_1 + \cdots + \hat{G}_{N-r-1} - \hat{G} \right\|$$
$$\leq 2 \sum_{i=r+1}^{N} \sup_{t \in [0,T]} \sigma_i(t).$$

∎

Remark 3 (Time-Invariant $\sigma(\cdot)$): For time-invariant asymptotically stable systems we can find time-invariant solutions $\Sigma(\cdot) = \Sigma$ to (35) and (36), which become algebraic Lyapunov inequalities. We then recover the well-known error bound for time-invariant systems. Also for time-varying systems we may find time-invariant solutions. If we look for solutions to the LMIs (35) and (36) with the constraint $\Sigma_2(\cdot) = \Sigma_2$ using standard semi-definite programming techniques, we obtain the error bound first shown in [13], [14]. In [14] it was shown that there always exists a solution $\Sigma(\cdot) = \Sigma$ to (35), so that problem is always feasible. However, if the time horizon $[0,T]$ of the problem is large, the LMIs are of high dimension and become computationally expensive to solve.

B. Nonmonotonically Balanced Systems

For many systems we expect the balanced Gramians $\Sigma(\cdot)$ to be nonmonotonic in time. We might try to resolve this by changing the boundary conditions to the Lyapunov equations until a monotonic solution is found, and then use Theorem 1. Alternatively we may search for time-invariant solutions, as commented in Remark 3. In any case, we would still like to have a bound for nonmonotonic solutions, and this will be derived in this section. The following definition will be useful:

Definition 1 (The Max–Min Ratio of σ): Let the singular value $\sigma(\cdot)$ be defined on the interval $\mathcal{T} = [t_0, t_f]$, and let it have M local *maximums* for $t > t_0$, located at $t_0 < t_1^{\max} < \cdots < t_M^{\max} \leq t_f$. Then there will be M local *minimums* so that

$$t_0 \leq t_1^{\min} < t_1^{\max} < \cdots < t_M^{\min} < t_M^{\max} \leq t_f$$

where $\sigma(t_i^{\min})$ is the local minimum immediately before $\sigma(t_i^{\max})$ for $i = 1 \ldots M$. The max–min ratio of σ is defined as

$$S_{\mathcal{T}}(\sigma) = \sigma(t_0) \prod_{i=1}^{M} \frac{\sigma(t_i^{\max})}{\sigma(t_i^{\min})}, \quad M > 0$$
$$S_{\mathcal{T}}(\sigma) = \sigma(t_0), \quad M = 0.$$

Now, we can formulate a general error-bound that applies both to monotonically and nonmonotonically balanced systems.

Theorem 2 (General Error Bound): Let $l(i)$ be any function that is defined for $i = 1 \ldots L$ and takes integer values in the range $1 \ldots n$, where n is the number of states in G.

The error between the balanced system G and its truncated balanced realization $\hat{G}$, where the states $x_{l(i)}$ have been truncated on the time intervals $\mathcal{T}_i$, $i = 1 \ldots L$, is bounded by

$$\left\| G - \hat{G} \right\| \leq 2 \sum_{i=1}^{L} S_{\mathcal{T}_i}(\sigma_{l(i)}) \tag{41}$$

and $\hat{G}$ is balanced.

If the singular value for some other state x_k, $k \neq l(i)$, coincides with one in the sum (41), then x_k can be truncated over the same interval without inducing extra error.

Proof: Start to truncate all states with the singular value $\sigma_{l(1)}$ over $\mathcal{T}_1$. Permute the states so that we can use Proposition 1. Then $p(\cdot) = q(\cdot) = \sigma_{l(1)}(\cdot)$. We need to find nonincreasing a and b such that

$$a(\cdot) \sigma_{l(1)}^2(\cdot) = b(\cdot).$$

If $\sigma_{l(1)}$ is initially nonincreasing put $b(t) = \sigma_{l(1)}^2(t)$ and $a(t) = 1$ (use $b(k-1)$ and $a(k-1)$ in discrete time). If $\sigma_{l(1)}$ reaches a local minimum at $t_1^{\min} < t_f$ define $b(t) = \sigma_{l(1)}^2(t_1^{\min})$ and $a(t) = \sigma_{l(1)}^2(t_1^{\min})/\sigma_{l(1)}^2(t)$ for $t > t_1^{\min}$. A local maximum will be reached, either at the end of the interval or before, so $t_1^{\max}$ exists. We can continue to define $a(t)$ and $b(t)$ as before, i.e., one is always constant and the other decreasing. When the whole interval $\mathcal{T}_1$ is covered we have from Proposition 1

$$a(t_f) \|y - \hat{y}_1\|^2 = \inf_{t \in \mathcal{T}_1} a(t) \|y - \hat{y}_1\|^2$$
$$\leq 4 \sup_{t \in \mathcal{T}_1} b(t) \|u\|^2 = 4b(t_0) \|u\|^2$$

and, therefore

$$\left\| G - \hat{G}_1 \right\| \leq 2 \sqrt{\frac{b(t_0)}{a(t_f)}} = 2 S_{\mathcal{T}_1}(\sigma_{l(1)}).$$

If $\sigma_{l(1)}$ is initially nondecreasing an analogous treatment is applicable.

Finally, we can continue recursively with $i = 2 \ldots L$ and use the triangular inequality to obtain the final result, just as in Theorem 1. ∎

Remark 4 (Large Max–Min Ratios): The max–min ratio may in some cases be an unnecessarily conservative bound. This is the case when $\sigma(t_i^{\max})/\sigma(t_i^{\min})$ is a large number. Then it is advisable to split the interval $\mathcal{T}$ into two intervals: $\mathcal{T}_1 = [t_0, t_i^{\min}]$ and $\mathcal{T}_2 = [t_i^{\min}, t_f]$, and truncate the state in two steps. We can always divide every time interval $\mathcal{T}$ into smaller ones so that
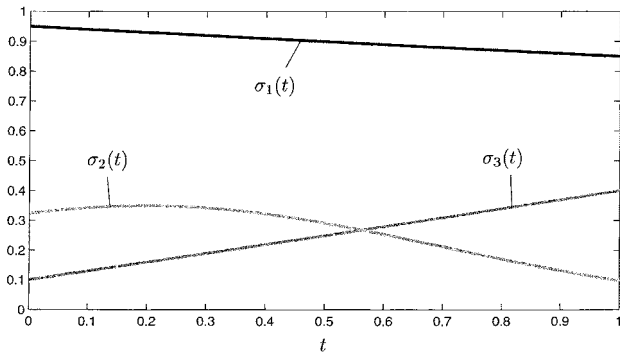
Fig. 1. Singular values for Example 2. σ_2 and σ_3 are truncated in the example.

the singular value is monotonic in each subinterval, and remove them recursively.

Example 1 (The Monotonic Case): Theorem 1 follows from Theorem 2. Notice that for monotonic singular values $\sigma(\cdot)$, $S_{\mathcal{T}}(\sigma) = \sup_{\mathcal{T}} \sigma(\cdot)$. So, we have

$$l(1) = r+1 \quad \mathcal{T}_1 = [0, T] \quad S_{\mathcal{T}_1}(\sigma_{r+1}) = \sup_{t \in \mathcal{T}_1} \sigma_{r+1}(t)$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$l(L) = N \quad \mathcal{T}_L = [0, T] \quad S_{\mathcal{T}_L}(\sigma_N) = \sup_{t \in \mathcal{T}_L} \sigma_N(t).$$

Example 2 (Continuous-Time System): Assume we have a third-order balanced continuous-time system G over the time interval $[0, 1]$. The realization has the dimensions

$$A(t) \in \mathbf{R}^{3 \times 3} \quad B(t) \in \mathbf{R}^{3 \times 1} \quad C(t) \in \mathbf{R}^{1 \times 3}$$

and the balanced Gramian is $\Sigma(t) = \mathrm{diag}\{\sigma_1(t), \sigma_2(t), \sigma_3(t)\}$, so that σ_i is the singular value of state x_i. The singular values are plotted in Fig. 1. If we truncate the state x_3 over $[0, 1]$ we obtain the system $\hat{G}_1$. As σ_3 is monotonic, we can use Theorem 1

$$\|G - \hat{G}_1\| \le 2 \sup_t \sigma_3(t) = 0.8.$$

Alternatively, we use Theorem 2 and get the same value

$$\|G - \hat{G}_1\| \le 2 S_{[0,1]}(\sigma_3) = 2\sigma_3(0) \frac{\sigma_3(1)}{\sigma_3(0)} = 0.8.$$

If we then want to truncate x_2 over $[0, 1]$ from $\hat{G}_1$, to get $\hat{G}_2$, we have the bound

$$\left\| \hat{G}_1 - \hat{G}_2 \right\| \le 2 S_{[0,1]}(\sigma_2) = 2\sigma_2(0) \frac{\sigma_2(0.2)}{\sigma_2(0)} = 0.7$$

as the only maximum is $\sigma_2(0.2)$, and the minimum immediately before is $\sigma_2(0)$. Therefore, the error between the first-order system $\hat{G}_2$ and G is bounded by

$$\left\| G - \hat{G}_2 \right\| \le 0.8 + 0.7 = 1.5.$$

As noted in Remark 4 it is important how the intervals $\mathcal{T}_i$ are chosen and how much the singular values varies over that interval. It may very well be that we need to let the state dimension vary in order for the error to be smaller than some chosen threshold. Finally, notice that the max–min ratio is just a bound resulting from a particular choice of $a(\cdot)$ and $b(\cdot)$. There are other choices and bounds; see [23] for an entirely different but more complex choice.

C. Periodic Systems

Periodic systems are very important special cases of time-varying systems. For instance, we obtain such a system when a nonlinear system is linearized around a limit cycle. Periodic systems have realizations where

$$A(\cdot) = A(\cdot + T) \quad B(\cdot) = B(\cdot + T)$$
$$C(\cdot) = C(\cdot + T) \quad D(\cdot) = D(\cdot + T)$$

for some time period T. These systems have received much attention in the literature, see, for instance, [1], [24], and the references therein. For stable balanced periodic systems we can find periodic Gramians

$$\Sigma(\cdot) = \Sigma(\cdot + T)$$

which solves (35) and (36) with equality; see [16] and [25]. These solutions are clearly not monotonic. A problem with applying Theorem 2 directly to these solutions is that for each new period included in $\mathcal{T}_i$, the bound grows. Still we would like to let $t_f \to \infty$ for many periodic systems. In [13], [15], and [16], a bound for balanced discrete time-periodic systems is presented. We can also derive this bound:

Corollary 1 (Balanced Discrete Time-Periodic Systems): If the balanced system G has a Gramian $\Sigma(k+T) = \Sigma(k)$ for all k and some T, and $\Sigma(k)$ is partitioned as in Theorem 1, then its truncation $\hat{G}$ is balanced with $\Sigma_1(k)$ and

$$\left\| G - \hat{G} \right\| \le 2 \sum_{k=1}^{T} \sum_{i=r(k)+1}^{N(k)} \sigma_i(k) \qquad (42)$$

over the infinite horizon $[0, \infty)$.

Proof: Use Proposition 1. Remove first the states with the singular value $\sigma_N(1)$. As the system is periodic we can simultaneously remove these states at $1, 1+T, 1+2T, \ldots$. So, $\mathcal{T} = \{1, 1+T, 1+2T, \ldots\}$. The constant values $a(k) = 1$ and $b(k) = \sigma_N^2(1)$ for all k fulfills the cancelling condition. Continue then recursively over the whole period and use then the triangular inequality. ∎

This might seem to be a satisfactory error bound. However, if the period is long (T large) this bound gets large very quickly if states are removed over the whole period. In particular, if we sample a continuous-time periodic system then the bound gets less useful the faster we have sampled the system. In the limit case, when we use the result directly on a continuous-time system, the bound is always infinity. More on sampling is given in Appendix II. A better technique to obtain a bound in this case may be to utilize the inequalities in (35) and (36) and to look for *time-invariant* diagonal solutions Σ; see Remark 3.

V. INPUT–OUTPUT STABILITY OF TRUNCATED SYSTEMS

One of the advantages of the analysis so far is that it has not been necessary to worry about stability. The only thing we need is a diagonal solution $\Sigma(\cdot)$ over some interval $[0, T]$. We could for instance reduce an unstable plant over a finite interval and still get error bounds. Many balanced truncation schemes in the literature requires asymptotic stability of the plants. Still, in order for our methodology to be good, a truncated realization

of a stable system G should also be stable in some sense. That this is indeed the case will be shown here in the continuous-time case. The discrete-time case is analogous. Assume from now on that there exist constants

$$0 < \underline{\sigma}I \le \Sigma(t) \le \overline{\sigma}I < \infty \quad (43)$$

for $0 \le t < \infty$. From the reachability Lyapunov inequality, we have

$$x^T(t)\Sigma^{-1}(t)x(t) \le \|u\|^2_{[0,t]}$$
$$\hat{x}^T(t)\Sigma_1^{-1}(t)\hat{x}(t) \le \|u\|^2_{[0,t]}. \quad (44)$$

So for all $u \in L_2^m[0,\infty)$ both x and $\hat{x}$ will be bounded. Even if the states in both G and $\hat{G}$ are bounded, it is not clear that if G is input–output stable (finite L_2-gain), that $\hat{G}$ will be input-output stable. But with the results from Sections II–IV we have the following theorem.

Theorem 3 (Input–Output Stability): If the balanced system G is input–output stable and there are constants $\overline{\sigma}$ and $\underline{\sigma}$ satisfying (43), then all states x are bounded for all u in $L_2^m[0,\infty)$, and every truncated system $\hat{G}$ is also input–output stable and the states $\hat{x}$ are bounded.

Proof: See Appendix I. ∎

This result might seem contradictory to the result in [5], which says that we get guaranteed asymptotic stability on $\hat{G}$ if Σ_1 and Σ_2 have no entries in common. But in the theorem above we concentrate on input–output stability. To see the effects consider, the following example from [26].

Example 3 [26]: The continuous-time system with the transfer function

$$\frac{s^2 - s + 2}{s^2 + s + 2} \text{ and realization}$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|c} 0 & -\sqrt{2} & 0 \\ \sqrt{2} & -1 & \sqrt{2} \\ \hline 0 & -\sqrt{2} & 1 \end{array}\right]$$

is balanced with $\Sigma = I$. The $\{\sigma_2\}$-truncated system

$$\left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right] = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & 1 \end{array}\right]$$

is clearly not asymptotically stable but the pole is neither observable nor controllable, and the system is *input–output stable* and $\hat{x}_1$ will be bounded, more precisely 0, for all u. Theorem 3 says that this will always happen when truncating a balanced system. (Obviously, in this case, a better approximation is just to keep $D = 1$, as we then get a zero-order model and the same error bound.)

The result may seem unnecessary as we can truncate states that have equal singular values without extra cost. But the result shows that we do not need to worry about singular values that are equal for some time-instants, we will not lose input–output stability. The example also shows that a truncated system may have a nonminimal realization. The theorem, however, guarantees it is well behaved.

VI. LOWER BOUND ON THE APPROXIMATION ERROR

When doing optimal Hankel-norm approximation of time-invariant systems a lower bound on the Hankel-norm for approximations of different system order (McMillan degree) is obtained; see [7] and [27]. As the Hankel-norm is always smaller than or equal to the induced L_2-norm we also get a bound on the best possible approximation in this norm. We will see that a similar analysis is possible for linear time-varying systems. Let us consider finite-horizon linear systems G in continuous time and the following Lyapunov equations:

$$A^T(t)Q(t) + Q(t)A(t) - \dot{Q}(t) + C^T(t)C(t) = 0 \quad (45)$$
$$A(t)P(t) + P(t)A^T(t) + \dot{P}(t) + B(t)B^T(t) = 0 \quad (46)$$
$$Q(t_f) = 0 \quad P(t_0) = 0 \quad (47)$$
$$t \in (t_0, t_f): Q(t) > 0, \qquad P(t) > 0. \quad (48)$$

Inequalities (48) means that the realization of G is completely reachable and observable. Notice that we here have dropped the Lyapunov inequalities for equalities. This is not a severe restriction. In practice one often solves the equalities as a first step anyhow, as it is less computationally expensive than solving the strict inequalities with semidefinite programming, and because it often gives good enough upper error bounds.

If we can balance the (45)–(48), the balanced Gramian will have the interesting property $\Sigma(t_0) = \Sigma(t_f) = 0$. Balanced finite-horizon systems of this sort were throughly studied in [10]. Among other things it was shown that if $\{A(t), B(t), C(t)\}$ are analytic functions in t, then the coordinate transformation $T(t)$ needed to obtain a balanced realization $\{\tilde{A}(t), \tilde{B}(t), \tilde{C}(t)\}$ exists, and is a Lyapunov transformation in every compact subset of (t_0, t_f). The entries of the balanced realization will tend to infinity at the boundaries t_0 and t_f. For practical computations it seems to be reasonable to embed the interval of interest, $[0, T]$, in a sufficiently large interval $[t_0, t_f]$.

Let us look at the linear system G on the time interval $[t_0, t_f]$, and divide the interval into two parts: $[t_0, \tau]$ and $[\tau, t_f]$. If we have a solution $Q(t)$ to the observability Lyapunov (45) we can compute the norm $\|y\|_{[\tau, t_f]}$ simply if $u(t) = 0$ for $t > \tau$ and $x(\tau)$ is known. Then

$$x^T(\tau)Q(\tau)x(\tau) = \int_\tau^{t_f} |y(t)|^2 dt = \|y\|^2_{[\tau, t_f]}.$$

Analogously, we have results for the reachability (46) and from linear optimal control theory. There is a minimum control signal $u^*(t)$ (in L_2-sense) that takes the state from $x(t_0) = 0$ to any $x(\tau)$ that fulfills

$$x^T(\tau)P^{-1}(\tau)x(\tau) = \int_{t_0}^\tau |u^*(t)|^2 dt = \|u^*\|^2_{[t_0, \tau]}$$

see, for example, [27]. Now, define the Hankel-norm $\|G\|_{H,\tau}$ at time τ and calculate it as

$$\|G\|^2_{H,\tau} = \sup_{u \neq 0} \frac{\|y\|^2_{[\tau, t_f]}}{\|u\|^2_{[t_0, \tau]}}$$
$$= \sup_{x(\tau)} \frac{x^T(\tau)Q(\tau)x(\tau)}{x^T(\tau)P^{-1}(\tau)x(\tau)} = \overline{\sigma}(P(\tau)Q(\tau))$$

where $u(t) = 0$ for $t > \tau$. As $P(\tau) > 0$ and $Q(\tau) > 0$ we can find a balancing coordinate transformation at time τ from (37), so we have $\bar{\sigma}(P(\tau)Q(\tau)) = \bar{\sigma}(\Sigma^2(\tau)) = \sigma_1^2(\tau)$, because the Hankel-norm is invariant under coordinate transformations. Also, notice that $\|G\|_{H,t_0} = \|G\|_{H,t_f} = 0$ and that

$$\|G\|_{H,\tau} \leq \|G\| = \sup_{u \neq 0} \frac{\|y\|_{[t_0,t_f]}}{\|u\|_{[t_0,t_f]}} \qquad (49)$$

for all τ in $[t_0, t_f]$. Next, define the Hankel-operator of G at time τ, $\Gamma_{G,\tau}$, as the past to future restriction of G

$$\Gamma_{G,\tau} \colon L_2^m[t_0, \tau] \xrightarrow{G} L_2^p[\tau, t_f].$$

We see that $\|\Gamma_{G,\tau}\| = \|G\|_{H,\tau}$. The operator $\Gamma_{G,\tau}$ has finite rank, namely n. We here only consider constant-state-dimensional systems as full-order systems G often come in this form. For each choice of τ there is a singular-value decomposition (Schmidt decomposition; see [27] and [28]) of $\Gamma_{G,\tau}$

$$\Gamma_{G,\tau}u = \sum_{i=1}^{n} \sigma_i(\tau)\langle u, v_i \rangle w_i$$

where $\{v_i\}_1^n$ is a set of orthonormal functions in $L_2^m[t_0, \tau]$ and $\{w_i\}_1^n$ is a set of orthonormal functions in $L_2^p[\tau, t_f]$. $\sigma_i(\tau) = \lambda_i^{1/2}(P(\tau)Q(\tau))$ are the singular values. $\langle \cdot, \cdot \rangle$ is the standard scalar product on $L_2^m[t_0, \tau]$

$$\langle u, v \rangle = \int_{t_0}^{\tau} u^T(s)v(s)ds.$$

We can now state the following theorem:

Theorem 4 (Lower Error-Bound): Suppose G is a linear system with a finite-horizon nth-order realization with Gramians that fulfill (45)–(48). Let the singular values be ordered so that $\sigma_1(t) \geq \cdots \geq \sigma_n(t) > 0$ for each t. Then for any linear system $\hat{G}$ of order $r < n$ it holds that

$$\left\|G - \hat{G}\right\|_{H,\tau} \geq \sigma_{r+1}(\tau) \qquad (50)$$

for all $\tau \in [t_0, t_f]$. Furthermore

$$\left\|G - \hat{G}\right\| \geq \max_t \sigma_{r+1}(t). \qquad (51)$$

Proof: The operator $\Gamma_{\hat{G},\tau}$ has rank r. If we use the Schmidt vectors v_i from $\Gamma_{G,\tau}$ as basis there exist numbers $\alpha_i \neq 0$ such that the signal $v = \sum_{i=1}^{r+1} \alpha_i v_i$ gives $\Gamma_{\hat{G},\tau}v = 0$. Now

$$\left\|(\Gamma_{G,\tau} - \Gamma_{\hat{G},\tau})v\right\|^2 = \|\Gamma_{G,\tau}v\|^2 = \left\|\sum_{i=1}^{r+1} \alpha_i \sigma_i(\tau)w_i\right\|^2$$

$$= \sum_{i=1}^{r+1} \alpha_i^2 \sigma_i^2(\tau) \geq \sigma_{r+1}^2(\tau) \sum_{i=1}^{r+1} \alpha_i^2$$

$$= \sigma_{r+1}^2(\tau) \|v\|^2.$$

This gives (50), and (49), then gives (51). ∎

If we make a one-step truncation ($L = 1$) of a finite-horizon balanced system G and truncate the states with the singular value $\sigma_N(t)$, we get

$$\max_t \sigma_N(t) \leq \left\|G - \hat{G}\right\| \leq C \cdot \max_t \sigma_N(t)$$

where $C \geq 2$ depends upon the monotonicity conditions as discussed in Theorem 2. Therefore we can often expect a very good approximation in this type of one-step reductions. For multistep reductions ($L > 1$), the approximation may be much less close to an optimal approximation, just as for standard balanced truncation for time-invariant systems. However, notice that we have not proven that there exists an approximation that really obtains the lower bound, so we do not know exactly how far away the optimum is. In [14], a sufficient and necessary condition for the existence of a system $\hat{G}$ of order r for a given approximation error γ is given. The condition is however nonconvex and hard to check. The discussion here only justifies the balanced truncation procedure when the lower and upper bounds are close to each other.

VII. EXAMPLE: REDUCTION OF DIESEL EXHAUST CATALYST MODEL

Until now there has been no computations that show that the suggested methods really give rise to good low-order approximations in practice. In fact, there has been a fair amount of theoretical work done in the literature on time-varying balancing, but the authors have not found many real examples. Here we will give a brief overview of the results for an example, just to show that the computations are feasible.

We will look at a model taken from [29]. This is a model of a diesel exhaust catalyst. In one end of the catalyst the exhausts from the diesel engine comes in. The exhausts are blended with some extra diesel fuel (HC). The amount of added diesel fuel is the control input in this example. In the catalyst the exhausts and the diesel react and at the other end the concentration of NO_x will have decayed.

The given model consists of 28 nonlinear stiff differential equations which describe concentrations and temperatures throughout the catalyst. To get a single-input–single-output system we choose the added amount of HC at the inlet as input, and the concentration of NO_2 at the outlet as output. If we are only interested in these aspects of the system, then we can directly drop four of the states in the nonlinear model. To apply the methods of this paper we need to linearize the system. In order to get a time-varying system we linearize the system around a pulsating input signal (three pulses) over a finite horizon, so that the system does not reach steady-state. We then get a time-varying linear system G with 24 states around a nominal trajectory.

To find a balanced realization and the singular values we need to solve two time-varying Lyapunov inequalities. As $n = 24$ this involves rather heavy computations. We choose to first find solutions to the system (45)–(47), with $t_0 = -10$ s and $t_f = 450$ s. The singular values, $\sigma_i(t) = \lambda_i^{1/2}(P(t)Q(t))$, are plotted in Fig. 2. The plot is in logarithmic scale and we notice that one singular value ($\sigma_1(t)$) is dominating. The three pulses in the nominal solution can be seen as three drops in the singular values.
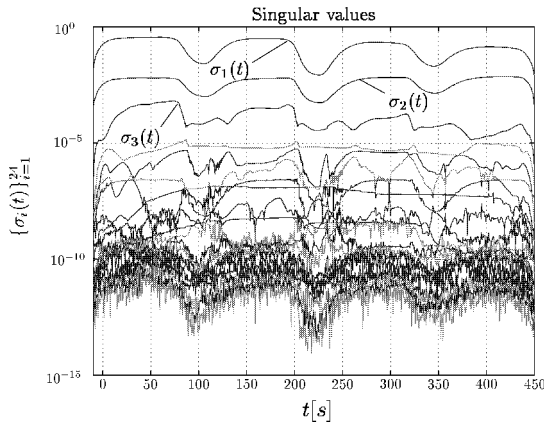
Fig. 2. Singular values for the linear time-varying system G, which approximates the diesel exhaust catalyst over the time interval -10–450 s. One singular value is dominating, which predicts that one state is needed to make the approximation.
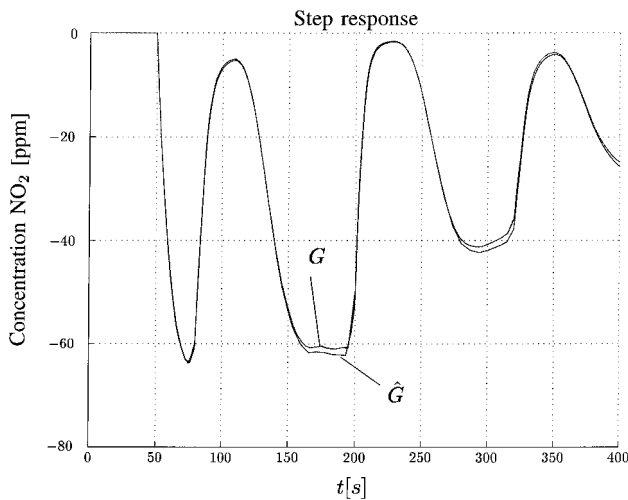

Fig. 3. Step responses for the 24th-order linear time-varying system G and its first-order approximation $\hat{G}$.

To reduce the computation time we have chosen the ODE-solver tolerance (for `ode15s` in MATLAB) so that only the two largest singular values have good accuracy. To find a balanced coordinate system we use the relation (37) on a time grid $\{t_k\}$. As the eigenvectors of $P(t_k)Q(t_k)$ typically give a badly conditioned coordinate transformation, we have chosen to use the numerically sound Schur method developed for time-invariant systems in [30] at each grid point to obtain a well-behaved coordinate transformation $T(t_k)$. Linear interpolation is used between the grid points.

We have shown in this paper that we can truncate states that have a small singular value without inducing large errors. For a first-order approximation ($\hat{n} = 1$), the upper error bound is essentially $2 \cdot S_{[-10,450]}(\sigma_2)$, if we assume that the other much smaller singular values also really only have four maximums. Now, as $\sigma_{2,\max} \approx 6 \cdot 10^{-3}$ and $\sigma_{2,\min} \approx 1 \cdot 10^{-3}$ we get that $S_{[-10,450]}(\sigma_2) \approx (6 \cdot 10^{-3})^4/(1 \cdot 10^{-3})^3 = 1296 \cdot 10^{-3}$. This is an overly conservative bound. Instead, one should divide into time intervals as suggested in Remark 4. So another, and better, bound is given by $2 \cdot (S_{[-10,110]} + S_{[110,225]} + S_{[225,340]} +$

$S_{[340,450]}) \approx 2 \cdot 4 \cdot \sigma_{2,\max} = 48 \cdot 10^{-3}$. As we derived a lower bound in Section VI, we can say

$$6 \cdot 10^{-3} \leq \left\| G - \hat{G} \right\| \leq 48 \cdot 10^{-3} \qquad (52)$$

for the first-order approximation $\hat{G}$. In Fig. 3 we see a step response test for G and $\hat{G}$. The error in this particular case is $7.2 \cdot 10^{-3}$, which shows that a typical error is in the same order of magnitude as the worst-case bounds in (52). Notice that the step responses here are very different from what is obtained from *time-invariant* linear systems. If we instead use a second-order approximation, $\hat{n} = 2$, there is no visible error in the step response test.

We have succeeded in finding a low-order approximation for a nontrivial high-order linear time-varying system. The drawback is that solving for $P(t)$ and $Q(t)$ is computationally heavy, although it is feasible for n of this order of magnitude.

VIII. CONCLUSION

In this paper, we have from basic analysis of the reachability and observability Lyapunov inequalities analyzed the effects of truncation of states for linear systems, in both continuous and discrete time. The analysis also covers the case when the state dimension varies over time. This is valuable as systems may need a different amount of states for different time intervals to be well approximated.

In particular, we have studied balancing of time-varying systems. From the solutions to the two Lyapunov inequalities, the Gramians, we obtain a balanced coordinate system, often well suited for truncation, and singular values. The singular values give an upper bound on the L_2-induced error for truncated models. Furthermore, we obtain a lower error bound also expressed in the singular values. Both bounds are generalizations of well-known results for time-invariant systems.

Stability was not a main issue in the paper, as we can make approximations over a finite time horizon. Nevertheless, we proved that if a full-order system is input–output stable, then every truncated balanced realization of it will also be input–output stable.

Finally, a brief example showed that the methods are possible to use in practice. A 24th-order linear time-varying approximation of a diesel exhaust catalyst was truncated to a first-order system with almost no error.

Future work should include finding sharper error bounds. Especially in the infinite time-horizon case with nonmonotonic singular values. Furthermore, numerical issues should be considered. The method requires knowledge of the Gramians of the system, which restricts the use of the method. The Gramians may be too computationally expensive to obtain for high-order systems.

APPENDIX I

Proof of Theorem 3

We will prove that input–output stability is maintained every time Proposition 1 is used to truncate a system. Under the given assumptions there are constants so that

$$0 < \delta_a \leq a(t) \leq \epsilon_a < \infty \quad 0 < \delta_P \leq P(t) \leq \epsilon_P < \infty$$
$$0 < \delta_b \leq b(t) \leq \epsilon_b < \infty \quad 0 < \delta_Q \leq Q(t) \leq \epsilon_Q < \infty$$

for all t. The calculations will be made in continuous time, but they are very similar in discrete time. Upon adding Lemma 3 ii) and Lemma 4 ii) we obtain

$$\left\| \begin{bmatrix} x_1(T) + \hat{x}(T) \\ x_2(T) \end{bmatrix} \right\|_{bP^{-1}}^2 + \left\| \begin{bmatrix} x_1(T) - \hat{x}(T) \\ x_2(T) \end{bmatrix} \right\|_{aQ}^2$$
$$+ \|y - \hat{y}\|_a^2 \le 4 \|u\|_b^2.$$

Using the inequality $\|x + y\|_a^2 \ge 1/2 \|x\|_a^2 - \|y\|_a^2$, we get

$$\frac{1}{2}|\hat{x}(T)|_{aq+bp^{-1}}^2 + \frac{1}{2} \|\hat{y}\|_a^2 \le 4 \|u\|_b^2 + \|y\|_a^2$$
$$+ |x_1(T)|_{aQ_1+bP_1^{-1}}^2 - |x_2(T)|_{aq+bp^{-1}}^2.$$

If $u \in L_2[0, \infty)$ and G is input–output stable, we know that the terms $\|u\|_b^2$ and $\|y\|_a^2$ are bounded. Because of the relations (44) and $0 < aq + bp^{-1} \le \epsilon_a \epsilon_Q + \epsilon_b \delta_P^{-1} < \infty$ for all t we see that the terms involving $\hat{x}(T)$, $x_1(T)$ and $x_2(T)$ are bounded for all T. Therefore we conclude that $\hat{y} \in L_2[0, \infty)$.

APPENDIX II

Sampled Lyapunov Equations

In this paper, we have treated systems in both discrete and continuous time. Models from physics and engineering often come in the form of differential equations. For control purposes, however, systems have to at some point be transformed into discrete time if implementation on computers is intended. We will see that this discretization can be done at the same time as the model reduction is performed.

The first step toward discretization in time is to find a different system representation. We will use so-called *lifting*, see for instance [31]. This transformation is an isomorphic isometry, i.e., the transformation preserves the system structure and norm. We will call the discretization time points $\{t(k)\}$. The inputs and the outputs of the lifted system belong to the signal spaces

$$\bar{u}(k) \in L_2^m[t(k), t(k+1)] \quad \bar{y}(k) \in L_2^p[t(k), t(k+1)].$$

The lifted n-state continuous-time system G is given by

$$\bar{x}(k+1) = \bar{A}(k)\bar{x}(k) + \bar{B}(k)\bar{u}(k)$$
$$\bar{y}(k) = \bar{C}(k)\bar{x}(k) + \bar{D}(k)\bar{u}(k)$$

where

$$\bar{A}(k) = \Phi(t(k+1), t(k)) \tag{53}$$
$$\bar{B}(k)\bar{u}(k) = \int_{t(k)}^{t(k+1)} \Phi(t(k+1), s)B(s)\bar{u}(k; s)ds \tag{54}$$
$$\bar{C}(k; t) = C(t)\Phi(t, t(k)) \tag{55}$$
$$\bar{D}(k)\bar{u}(k)(t) = \int_{t(k)}^{t} C(t)\Phi(t, s)B(s)\bar{u}(k; s)ds + D(t)\bar{u}(k; t) \tag{56}$$

and $\Phi(t, \tau)$ is the fundamental solution of $\dot{x} = A(t)x$. The operators act on the following spaces:

$$\bar{A}(k): \mathbf{R}^n \to \mathbf{R}^n$$
$$\bar{B}(k): L_2^m[t(k), t(k+1)] \to \mathbf{R}^n$$
$$\bar{C}(k): \mathbf{R}^n \to L_2^p[t(k), t(k+1)]$$
$$\bar{D}(k): L_2^m[t(k), t(k+1)] \to L_2^p[t(k), t(k+1)].$$

We will need the adjoint operators. These act on the dual spaces. As all involved spaces are Hilbert spaces we can represent all elements in the dual space with elements in the primal space. The adjoints we need are given by

$$\bar{A}^*(k) = \Phi^T(t(k+1), t(k)) \tag{57}$$
$$\bar{B}^*(k; t) = B^T(t)\Phi^T(t(k+1), t) \tag{58}$$
$$\bar{C}^*(k)\bar{y}(k) = \int_{t(k)}^{t(k+1)} \Phi^T(s, t(k))C^T(s)\bar{y}(k; s)ds. \tag{59}$$

We will now see that if we have solutions to the continuous-time Lyapunov equations, $Q(t)$ and $P(t)$, we can use them at the sampling instants for the lifted system. Consider the observability Lyapunov equation in continuous time for $t \in [t(k), t(k+1)]$

$$Q(t)A(t) + A^T(t)Q(t) + \dot{Q}(t) + C^T(t)C(t) = 0$$

and its solution

$$\Phi^T(t(k+1), t)Q(t(k+1))\Phi(t(k+1), t)$$
$$+ \int_t^{t(k+1)} \Phi^T(s, t)C^T(s)C(s)\Phi(s, t)ds = Q(t). \tag{60}$$

Using the lifting operators putting $t = t(k)$, the solution (60) can be written as a discrete Lyapunov equation with the solution $Q(t(k))$

$$\bar{A}^*(k)Q(t(k+1))\bar{A}(k) + \bar{C}^*(k)\bar{C}(k) = Q(t(k)). \tag{61}$$

We get analogous results for the reachability Lyapunov equation

$$A(t)P(t) + P(t)A^T(t) - \dot{P}(t) + B(t)B^T(t) = 0$$

with the solution

$$\Phi(t, t(k))P(t(k))\Phi^T(t, t(k))$$
$$+ \int_{t(k)}^{t} \Phi(t, s)B(s)B^T(s)\Phi(t, s)ds = P(t) \tag{62}$$

which with the lifting operators becomes

$$\bar{A}(k)P(t(k))\bar{A}^*(k) + \bar{B}(k)\bar{B}^*(k) = P(t(k+1)). \tag{63}$$

So, we can first compute continuous-time solutions $Q(t)$ and $P(t)$, and then choose suitable sampling instants $\{t(k)\}$ and compute the pointwise balancing transformation $T(t(k))$ to balance (61) and (63). Then we can truncate the lifted system and obtain error bounds as we have done before with discrete Lyapunov equations. Finally, finite-dimensional bases should be

chosen to approximate the infinite-dimensional signal spaces. For instance, zero-order hold could be used for the signals $\bar{u}(k)$.

As an alternative, we could for instance first do zero-order hold sampling of the continuous-time system G and then balance the resulting discrete-time system. We would then not obtain the same approximation as above and the error bound will be in induced ℓ_2-sense, not in induced L_2-sense as in the lifting approach.

ACKNOWLEDGMENT

The authors would like to thank B. Westerberg for the simulation model of the diesel exhaust catalyst.

REFERENCES

[1] S. Bittanti and P. Colaneri, "Periodic control," in *Encyclopedia of Electrical and Electronics Engineering*, J. Webster, Ed.   New York: Wiley, 1999, vol. 16, pp. 59–73.

[2] G. Tadmor, "Input/output norms in general linear systems," *Int. J. Control*, vol. 51, no. 4, pp. 911–921, 1990.

[3] C. Mullis and R. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Trans. Circuits Systems*, vol. CAS-23, pp. 551–562, Apr. 1976.

[4] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 17–32, Feb. 1981.

[5] L. Pernebo and L. Silverman, "Model reduction via balanced state space representation," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 382–387, Mar. 1982.

[6] D. Enns, "Model reduction with balanced realizations: An error bound and a frequency weighted generalization," presented at the Conf. Decision Control, Las Vegas, NV, 1984.

[7] K. Glover, "All optimal hankel-norm approximations of linear multivariable systems and their L_∞-error bounds," *Int. J. Control*, vol. 39, pp. 1115–1193, 1984.

[8] U. Al-Saggaf and G. Franklin, "An error-bound for a discrete reduced-order model of a linear multivariable system," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 815–819, Aug. 1987.

[9] S. Shokoohi, L. Silverman, and P. van Dooren, "Linear time-variable systems: Balancing and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 810–822, Aug. 1983.

[10] E. I. Verriest and T. Kailath, "On generalized balanced realizations," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 833–844, Aug. 1983.

[11] G. D. Wood, "Control of parameter-dependent mechanical systems," Ph.D. dissertation, Univ. Cambridge, Cambridge, U.K., 1996.

[12] C. Beck, J. Doyle, and K. Glover, "Model reduction of multidimensional and uncertain systems," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 1466–1477, Oct. 1996.

[13] S. Lall, C. Beck, and G. Dullerud, "Guaranteed error bounds for model reduction of linear time-varying systems," in *Proc. Amer. Control Conf.*, Philadelphia, PA, June 1998, pp. 634–638.

[14] S. Lall and C. Beck, "Error-bounds for balanced model-reduction of linear time-varying systems," *IEEE Trans. Automat. Contr.*, vol. 48, pp. 946–956, June 2003.

[15] S. Longhi and G. Orlando, "Balanced reduction of linear periodic systems," *Kybernetika*, vol. 35, no. 6, pp. 737–751, 1999.

[16] A. Varga, "Balanced truncation model reduction of periodic systems," in *Proc. 39th Conf. Decision Control*, Sydney, Australia, Dec. 2000.

[17] I. Gohberg, M. Kaashoek, and L. Lerer, "Minimality and realization of discrete time-varying systems," *Oper. Theory: Adv. Applicat.*, vol. 56, pp. 261–296, 1992.

[18] W. J. Rugh, *Linear System Theory*, ser. Information and System Sciences.   Upper Saddle River, NJ: Prentice-Hall, 1996.

[19] D. Hinrichsen and A. Pritchard, "An improved error estimate for reduced-order models of discrete-time systems," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 317–320, Feb. 1990.

[20] S. Shokoohi and L. M. Silverman, "Identification and model reduction of time-varying discrete-time systems," *Automatica*, vol. 23, no. 4, pp. 509–521, 1987.

[21] S. Shokoohi, L. Silverman, and P. van Dooren, "Linear time-variable systems: Stability of reduced models," *Automatica*, vol. 20, no. 1, pp. 59–67, 1984.

[22] J. Imae, J. Perkins, and J. Moore, "Toward time-varying balanced realization via riccati equations," *Math. Control, Signals, Syst.*, vol. 5, pp. 313–326, 1992.

[23] H. Sandberg and A. Rantzer, "Balanced model reduction of linear time-varying systems," presented at the 15th IFAC World Congr., Barcelona, Spain, July 2002.

[24] E. Möllerstedt, "Dynamic analysis of harmonics in electrical systems," Ph.D. dissertation, Dept. Automatic Control, Lund Inst. Technol., Lund, Sweden, 2000.

[25] H. Kano and T. Nishimura, "A note on periodic Lyapunov equations," in *35th Conf. Decision Control*, Kobe, Japan, Dec. 1996.

[26] K. Zhou and J. Doyle, *Essentials of Robust Control*.   Upper Saddle River, NJ: Prentice-Hall, 1998.

[27] M. Green and D. J. Limebeer, *Linear Robust Control*, ser. Information and System Sciences.   Upper Saddle River, NJ: Prentice-Hall, 1995.

[28] N. Young, *An Introduction to Hilbert Space*.   Cambridge, U.K.: Cambridge Univ. Press, 1988.

[29] B. Westerberg, C. Künkel, and I. Odenbrand, "Transient modeling of a HC-SCR catalyst for diesel exhaust aftertreatment," *Chem. Eng. J.*, vol. 92, pp. 27–39, 2003.

[30] M. Safonov and R. Chiang, "A schur method for balanced-truncation model reduction," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 729–733, July 1989.

[31] B. Bamieh and J. Pearson, "A general framework for linear periodic systems with applications to H_∞ sampled-data control," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 418–435, Mar. 1992.

Henrik Sandberg (S'02) was born in 1976. He received the M.S. degree in engineering physics from the Lund Institute of Technology, Lund, Sweden, in 1999. He is currently working toward the Ph.D. degree with the Department of Automatic Control at the same university.

He was Guest Graduate Student for a semester in 2001 at the Department of Control and Dynamical Systems at California Institute of Technology, Pasadena. During the first half of 2003, he was a Visiting Student at the Mittag–Leffler Institute, Stockholm, Sweden. His main research interests are in modeling, control, and reduction of linear time-varying systems.

Anders Rantzer (S'90–M'91–SM'97–F'01) was born in 1963. He received the Ph.D. degree in optimization and systems theory from the Royal Institute of Technology (KTH), Stockholm, Sweden in 1991.

After postdoctoral positions at KTH and the University of Minnesota, Minneapolis, he joined the Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, in 1993. In 1999, he was appointed Professor of Automatic Control. His research interests are in modeling, analysis, and synthesis of control systems, with particular attention to uncertainty, nonlinearities, and hybrid phenomena.

Dr. Rantzer was a winner of the 1990 SIAM Student Paper Competition and the 1996 IFAC Congress Young Author Prize. He has served as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and several other journals.