2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Model reduction of linear systems using extended balanced truncation

Henrik Sandberg

***Abstract*— An extension to balanced truncation is presented. Balanced truncation is a standard method for model reduction and it has many good properties, such as preservation of model stability and a priori error bounds. Balanced truncation is done using controllability and observability Gramians. The Gramians can be found by solving a set of linear matrix inequalities. In this paper, we show that these linear matrix inequalities can be extended so that the number of decision variables are at least doubled. This leads to the concept of extended Gramians. It is shown that all the good properties of balanced truncation also hold for extended balanced truncation. It turns out that extended balanced truncation is especially useful when there are additional structure constraints in the model to be reduced. This is useful in frequency-weighted reduction, for example.**

I. Introduction

Model reduction is about systematic simplification of complex models. This is a topic that has received a great deal of interest at least since the early 1980's, but still hard problems remain to be solved. In particular, most available methods that are guaranteed to deliver good approximations only apply to linear systems, and these methods are usually not structure preserving. In this paper, we propose an extension to balanced truncation that increases the chance of preserving structures.

Balanced truncation is a standard method to reduce the order of linear systems. The method was introduced in [1] and it was later shown in [2] how stability of models is preserved, and in [3], [4] that simple a priori error bounds hold. These results were all shown for continuous-time models. In discrete time, the corresponding results were shown in [5]. For balanced truncation, the controllability and observability Gramians are needed. They are obtained from solving two Lyapunov equations. In [6], it was shown that if the Lyapunov equations are replaced by linear matrix inequalities, the error bounds in [5] can be improved. In this paper, we build upon this idea and introduce more general (extended) inequalities. The extended inequalities are more likely than the normal inequalities to have Gramians of block-diagonal structure. Block-diagonal Gramians are important for controller reduction [7], for uncertain systems [8], and for reduction of structured and interconnected systems [9]–[11].

Extending linear matrix inequalities the way that it is done in this paper was first proposed in [12], [13]. In [12], [13], the applications were stability analysis and controller synthesis for parameter-dependent models. The same type of extension was also used in [14], [15] to construct distributed estimators. Here we use the idea for model reduction.

The organization of the paper is as follows: In Section II, we introduce the model structure and define what we mean by model truncation. In Section III, the normal and extended Gramians are defined, and their equivalence is discussed. In Section IV, it is shown how the extended Gramians can be used to prove approximation error bounds and preserve stability for truncated models. In Section V, it is shown how one can balanced the extended Gramians, which is required for the error bounds to apply. Finally, in Section VI extended and regular balanced truncation is applied to two different models, and the results are discussed.

II. Preliminaries

We consider linear finite-dimensional discrete-time systems $\mathcal{G}$, with realization

$$\mathcal{G}\begin{cases} x(k+1) = Ax(k) + Bu(k), & x(k) \in \mathbb{R}^n, u(k) \in \mathbb{R}^m, \\ \quad y(k) = Cx(k) + Du(k), & y(k) \in \mathbb{R}^p. \end{cases}$$

For simplicity, we leave out the time index k in the notation in the following, and use the notation $x^+ := x(k+1), x := x(k), u := u(k)$ etc. Thus, the model is written

$$\mathcal{G}\begin{cases} x^+ = Ax + Bu, & x \in \mathbb{R}^n, u \in \mathbb{R}^m, \\ \;y = Cx + Du, & y \in \mathbb{R}^p, \end{cases}$$

in the following. The transfer function $\mathcal{G}(z)$ is defined by

$$\mathcal{G}(z) = D + C(zI - A)^{-1}B =: \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right],$$

and the H_∞-norm by

$$\|\mathcal{G}\|_\infty := \sup_{z\in\mathbb{C}\setminus\bar{\mathbb{D}}} |\mathcal{G}(z)|.$$

With $\|u\|_{[k_1,k_2]}$ we mean

$$\|u\|_{[k_1,k_2]} = \sqrt{\sum_{k=k_1}^{k_2} u(k)^T u(k)}$$

and $\|u\| := \|u\|_{[0,\infty]}$. Square-summable sequences, $\|u\| < \infty$, belong to the Hilbert space ℓ_2. We know that $\|\mathcal{G}\|_\infty = \sup_{(u\neq 0)\in\ell_2} \frac{\|y\|}{\|u\|}$.

To solve the model-reduction problem we should find a new linear system $\hat{\mathcal{G}}$ with $r < n$ states such that $\|\mathcal{G} - \hat{\mathcal{G}}\|_\infty$

H. Sandberg is at Royal Institute of Technology (KTH), School of Electrical Engineering, Automatic Control, SE-100 44 Stockholm, Sweden. hsan@ee.kth.se

This work was supported by the Swedish Foundation for Strategic Research.

is small. This is done by truncating the realization, in this paper. We use the partition

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},\ B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix},\ A_{11} \in \mathbb{R}^{r\times r}, B_1 \in \mathbb{R}^{r\times m},$$
$$C = \begin{bmatrix} C_1 & C_2 \end{bmatrix},\ C_1 \in \mathbb{R}^{p\times r}.$$

A candidate approximation $\hat{\mathcal{G}}$ is given by

$$\hat{\mathcal{G}} \begin{cases} \hat{x}_1^+ = A_{11}\hat{x} + B_1 u, & \hat{x}_1 \in \mathbb{R}^r, u \in \mathbb{R}^m, \\ \hat{y} = C_1\hat{x} + Du, & \hat{y} \in \mathbb{R}^p. \end{cases}$$

We will compare the state $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $(x_1 \in \mathbb{R}^r, x_2 \in \mathbb{R}^s)$ of $\mathcal{G}$ to the state $\hat{x}_1$ of $\hat{\mathcal{G}}$, and the output y of $\mathcal{G}$ to the output $\hat{y}$ of $\hat{\mathcal{G}}$, when both models are excited with same input u. We will also use the signal $\hat{z}_2$ defined by

$$\hat{z}_2^+ = A_{21}\hat{x}_1 + B_2 u \tag{1}$$

later in the paper. One can interpret $\hat{z}_2$ as an estimate of what x_2 is, given that we have the model $\hat{\mathcal{G}}$.

The main benefit with balanced truncation is that it shows how to choose good coordinates x and approximation order r such that $\|\mathcal{G} - \hat{\mathcal{G}}\|_\infty$ is guaranteed to be small. As we shall see, extended balanced truncation has the same benefits, and give some additional degrees of freedom.

III. Extended Lyapunov Inequalities

Controllability and observability Lyapunov inequalities,

$$P - APA^T - BB^T > 0, \quad P > 0, \tag{2}$$
$$Q - A^TQA - C^TC > 0, \quad Q > 0, \tag{3}$$

with symmetric Gramians $P = P^T, Q = Q^T \in \mathbb{R}^{n\times n}$ have solutions if, and only if, $\mathcal{G}$ is asymptotically stable ($\rho(A) < 1$). The Gramians contain information about controllability and observability of the realization and are used to choose coordinate system for balanced truncation [5], [16]. In this paper, we instead use the *extended controllability and observability Lyapunov inequalities*

$$\begin{bmatrix} P & AF & B \\ F^TA^T & F + F^T - P & 0 \\ B^T & 0 & I \end{bmatrix} > 0, \tag{4}$$
$$\begin{bmatrix} G + G^T - Q & GA & 0 \\ A^TG^T & Q & C^T \\ 0 & C & I \end{bmatrix} > 0, \tag{5}$$

with *extended controllability Gramian* (P, F) and *extended observability Gramian* (Q, G). Here $P = P^T, Q = Q^T \in \mathbb{R}^{n\times n}$ are symmetric and $F, G \in \mathbb{R}^{n\times n}$. A key result in this paper is that (4)–(5) and (2)–(3) in fact are equivalent, see Theorem 1. The idea of extending inequalities in this way was first presented by Oliviera *et al.* in [12], [13]. Similar equivalences were also stated there. The applications in these papers were stability analysis and controller synthesis for parameter-dependent models. Here the focus is on model reduction.

Theorem 1: The inequalities (2)–(3) have solutions $P = P^T, Q = Q^T$ if, and only if, the inequalities (4)–(5) have solutions $P = P^T, Q = Q^T$ and F, G.

Proof: The controllability case is proved in Theorem 1 in [13]. We repeat the idea of the proof here (in the observability case) since it helps to understand the roles of P, Q and F, G.

(*Necessity*) If (2)–(3) holds, choose $F = P$ and $G = Q$. In the observability case, using Schur complements [16] we obtain

$$\begin{bmatrix} Q & QA & 0 \\ A^TQ & Q & C^T \\ 0 & C & I \end{bmatrix} > 0 \Leftrightarrow \begin{bmatrix} Q & QA \\ A^TQ & Q - C^TC \end{bmatrix} > 0$$
$$\Leftrightarrow \begin{bmatrix} Q & 0 \\ 0 & Q - A^TQA - C^TC \end{bmatrix} > 0.$$

(*Sufficiency*) Assume (4)–(5) have solutions. In the observability case we have that $G + G^T > Q > 0$, and $(G - Q)Q^{-1}(G^T - Q) = GQ^{-1}G^T - G - G^T + Q \geq 0$ and thus $GQ^{-1}G^T \geq G + G^T - Q$. From (5) it follows

$$\begin{bmatrix} GQ^{-1}G^T & GA & 0 \\ A^TG^T & Q & C^T \\ 0 & C & I \end{bmatrix} > 0.$$

If we multiply this inequality with $\operatorname{diag}\{G^{-T}Q, I, I\}$ from the right and with the transpose from the left, we obtain

$$\begin{bmatrix} Q & QA & 0 \\ A^TQ & Q & C^T \\ 0 & C & I \end{bmatrix} > 0.$$

■

Remark 1: If the extended inequalities (4)–(5) are solved, then the P, Q components of the extended Gramians can always be used as Gramians in the traditional sense to solve (2)–(3). For example, for reachability and observability analysis. The F, G components contain other information. Their role and use will be shown in the later sections.

If the Lyapunov inequalities (2)–(3) are solved, then the Gramians P, Q can always be used to construct extended Gramians (P, P) and (Q, Q) to solve (4)–(5). However, this will not give us anything new. The idea is to *not* choose $F = P, G = Q$ and utilize the extra degrees of freedom.

IV. Model truncation using extended Gramians

We start out this section by making the bold assumption that the F, G components of the extended Gramians have the block-diagonal structure

$$F = \begin{bmatrix} F_1 & 0 \\ 0 & I_s \cdot \sigma \end{bmatrix} > 0, \quad G = \begin{bmatrix} G_1 & 0 \\ 0 & I_s \cdot \sigma \end{bmatrix} > 0. \tag{6}$$

In Section V, it is shown how it is possible to fulfill this assumption. No special structure is assumed for the P, Q components, apart from them being symmetric positive-definite matrices.

Let us first study the extended observability Lyapunov inequality (5). After an application of the Schur lemma [16] we obtain the inequality

$$\begin{bmatrix} G+G^T-Q & GA \\ A^TG^T & Q-C^TC \end{bmatrix} > 0. \tag{7}$$

If we assume the structure (6) for G, and multiply (7) from the right and left with properly chosen state vectors, we obtain,

$$\begin{bmatrix} -\begin{pmatrix} x_1-\hat{x}_1 \\ x_2 \end{pmatrix}^+ \\ \begin{pmatrix} x_1-\hat{x}_1 \\ x_2 \end{pmatrix} \end{bmatrix}^T \begin{bmatrix} G+G^T-Q & GA \\ A^TG^T & Q-C^TC \end{bmatrix} \times \begin{bmatrix} -\begin{pmatrix} x_1-\hat{x}_1 \\ x_2 \end{pmatrix}^+ \\ \begin{pmatrix} x_1-\hat{x}_1 \\ x_2 \end{pmatrix} \end{bmatrix} \geq 0$$

$$\iff$$

$$\begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix}^T Q \begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix}^{+T} Q \begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix}^+ + 2x_2^{+T}\hat{z}_2^+\sigma - |y-\hat{y}|^2 \geq 0, \tag{8}$$

where $\hat{z}_2$ is defined in (1), and we have used the identities

$$A\begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1-\hat{x}_1 \\ x_2-\hat{z}_2 \end{bmatrix}^+, \quad C\begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix} = y-\hat{y},$$

$$\begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix}^{+T} G \begin{bmatrix} x_1-\hat{x}_1 \\ x_2-\hat{z}_2 \end{bmatrix}^+ = \begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix}^{+T} G \begin{bmatrix} x_1-\hat{x}_1 \\ x_2 \end{bmatrix}^+ - x_2^{+T}\hat{z}_2^+\sigma.$$

Similarly, if we define

$$\tilde{F} := F^{-1} = \begin{bmatrix} F_1^{-1} & 0 \\ 0 & I_s\cdot\sigma^{-1} \end{bmatrix},$$
$$\tilde{P} := \tilde{F}^T P \tilde{F},$$

we can multiply the extended controllability Lyapunov inequality (4) from the left with $\operatorname{diag}\{\tilde{F},\tilde{F},I\}^T$ and from the right with $\operatorname{diag}\{\tilde{F},\tilde{F},I\}$, and we obtain the equivalent inequality

$$\begin{bmatrix} \tilde{P} & \tilde{F}^TA & \tilde{F}^TB \\ A^T\tilde{F} & \tilde{F}+\tilde{F}^T-\tilde{P} & 0 \\ B^T\tilde{F} & 0 & I \end{bmatrix} > 0. \tag{9}$$

Using the structure of $\tilde{F}$, we have

$$\begin{bmatrix} -\begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix}^+ \\ \begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix} \\ 2u \end{bmatrix}^T \begin{bmatrix} \tilde{P} & \tilde{F}^TA & \tilde{F}^TB \\ A^T\tilde{F} & \tilde{F}+\tilde{F}^T-\tilde{P} & 0 \\ B^T\tilde{F} & 0 & I \end{bmatrix} \times \begin{bmatrix} -\begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix}^+ \\ \begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix} \\ 2u \end{bmatrix} \geq 0$$

$$\iff$$

$$\begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix}^T (\tilde{F}+\tilde{F}^T-\tilde{P}) \begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix}^{+T} (\tilde{F}+\tilde{F}^T-\tilde{P}) \begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix}^+ - 2x_2^{+T}\hat{z}_2^+\sigma^{-1} + 4|u|^2 \geq 0, \tag{10}$$

where we have used the identities

$$A\begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix} + 2Bu = \begin{bmatrix} x_1+\hat{x}_1 \\ x_2+\hat{z}_2 \end{bmatrix}^+,$$

$$\begin{bmatrix} x_1+\hat{x}_1 \\ x_2 \end{bmatrix}^{+T} \tilde{F} \begin{bmatrix} x_1+\hat{x}_1 \\ x_2+\hat{z}_2 \end{bmatrix}^+ = \begin{bmatrix} x_1+\hat{x}_1 \\ x_2 \end{bmatrix}^{+T} \tilde{F} \begin{bmatrix} x_1+\hat{x}_1 \\ x_2 \end{bmatrix}^+ + x_2^{+T}\hat{z}_2^+\sigma^{-1}.$$

If we assume that $x(0)=0$ and $\hat{x}(0)=0$ and sum the inequalities (8) and (10) over the time interval $[0,T]$ we have that (note the canceling terms)

$$2\sigma\sum_{k=1}^{T+1}(x_2^T\hat{z}_2)(k) \geq \left|\begin{pmatrix} x_1-\hat{x}_1 \\ x_2 \end{pmatrix}(T+1)\right|_Q^2 + \|y-\hat{y}\|_{[0,T]}^2 \tag{11}$$

and

$$-2\sigma^{-1}\sum_{k=1}^{T+1}(x_2^T\hat{z}_2)(k) + 4\|u\|_{[0,T]}^2 \geq \left|\begin{pmatrix} x_1+\hat{x}_1 \\ x_2 \end{pmatrix}(T+1)\right|_{\tilde{F}+\tilde{F}^T-\tilde{P},}^2 \tag{12}$$

where $|x|_Q^2$ means x^TQx. Using these inequalities we have the following lemma that bounds the input-output approximation error.

Lemma 1: Assume that $\mathcal{G}$ and $\hat{\mathcal{G}}$ initially are at rest, and that the F,G components of the extended Gramians have the structure (6). Then for all inputs $u\in\ell_2$ it holds that $\|y-\hat{y}\|\leq 2\sigma\|u\|$, that is

$$\|\mathcal{G}-\hat{\mathcal{G}}\|_\infty \leq 2\sigma.$$

Proof: Multiply inequality (12) with σ^2, and add it to inequality (11). Notice that the sums containing the sign-indefinite terms $x_2^T\hat{z}_2$ cancel. All the remaining terms are positive and the result follows as $T\to\infty$. ■

Hence, if we truncate states and have block-diagonal F,G components in the extended Gramians, then the input-output approximation error is easily bounded. As we see next, asymptotic stability is also preserved and we can apply the results recursively.

Lemma 2: Assume that $\mathcal{G}$ has extended Gramians in the form

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}, \quad P_{22}\in\mathbb{R}^{s\times s}, \quad F = \begin{bmatrix} F_1 & 0 \\ 0 & I_s\cdot\sigma \end{bmatrix},$$
$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix}, \quad Q_{22}\in\mathbb{R}^{s\times s}, \quad G = \begin{bmatrix} G_1 & 0 \\ 0 & I_s\cdot\sigma \end{bmatrix}.$$

Then the truncated system $\hat{\mathcal{G}}$ has extended Gramians (P_{11}, F_1) and (Q_{11}, G_1). Furthermore, $\hat{\mathcal{G}}$ is asymptotically stable ($\rho(A_{11}) < 1$).

Proof: We only show the result in the controllability case. Multiply the extended controllability inequality (4) from the right with

$$\begin{bmatrix} \begin{pmatrix} I_r \\ 0_{s\times r} \end{pmatrix} & 0_{n\times r} & 0_{n\times m} \\ 0_{n\times r} & \begin{pmatrix} I_r \\ 0_{s\times r} \end{pmatrix} & 0_{n\times m} \\ 0_{m\times r} & 0_{m\times r} & I_m \end{bmatrix}$$

and with the transpose from the left. Then we obtain

$$\begin{bmatrix} P_{11} & A_{11}F_1 & B_1 \\ F^T A_{11}^T & F_1 + F_1^T - P_{11} & 0 \\ B_1^T & 0 & I \end{bmatrix} > 0.$$

That is, (P_{11}, F_1) is an extended Gramian for $\hat{\mathcal{G}}$. Since P_{11} is a normal Gramian satisfying (2) for $\hat{\mathcal{G}}$, asymptotic stability follows. ∎

V. Balanced extended Gramians

In this section, we justify the assumption (6) about block-diagonal F, G components in the extended Gramians. This leads to the concept of *balanced extended Gramians*.

Let us consider coordinate transformations $\bar{x} = Tx$, T invertible. We know that the realization transforms as

$$\bar{A} = TAT^{-1},\ \bar{B} = TB,\ \bar{C} = CT^{-1},\ \bar{D} = D.$$

How the extended Gramians transform is shown in the next lemma.

Lemma 3: Under coordinate transformations $\bar{x} = Tx$, the extended Gramians transform as

$$\begin{aligned} \bar{P} &= TPT^T, & \bar{F} &= TFT^T, \\ \bar{Q} &= T^{-T}QT^{-1}, & \bar{G} &= T^{-T}GT^{-1}. \end{aligned}$$

Proof: Replace A, B, C in (4)–(5) with $T^{-1}\bar{A}T$, $T^{-1}\bar{B}$, $\bar{C}T$. If (4) is multiplied with $\operatorname{diag}\{T^T, T^T, I\}$ from the right, and with the transpose from the left, then we can identify $\bar{P}$ and $\bar{F}$. A similar technique is used to prove the observability case. ∎

The F and G components transform just as the normal Gramians P and Q. In particular, the eigenvalues

$$\begin{aligned} \lambda_i(PQ) &= \lambda_i(\bar{P}\bar{Q}), \\ \lambda_i(FG) &= \lambda_i(\bar{F}\bar{G}), \end{aligned}$$

are invariant under coordinate transformations. The numbers $\sigma_i = \sqrt{\lambda_i(PQ)}$ are often called *Hankel singular values*. It is well known that there is a coordinate transformation T that makes $\bar{P}$ and $\bar{Q}$ equal and diagonal [1], [16]. Since F and G transform in the same way, one could hope that $\bar{F}$ and $\bar{G}$ can also be made equal and diagonal. However, if F and G are *not* symmetric, such a coordinate transformation may not exist. But we can always sacrifice some degrees of freedom in F and G and make them symmetric.

Lemma 4: Let the F, G components of the extended Gramians be symmetric. Then there exist a coordinate transformation $\bar{x} = Tx$ such that

$$\bar{F} = \bar{G} = \Sigma_e = \operatorname{diag}\{\sigma_{e,1}, \ldots, \sigma_{e,n}\},$$

where $\sigma_{e,i} := \sqrt{\lambda_i(FG)} > 0$ are the *extended Hankel singular values* of $\mathcal{G}$.

Proof: The transformation T is constructed as in Theorem 7.5 in [16] using the substitutions $P = F$ and $Q = G$. ∎

The extended Gramians are called *balanced extended Gramians* if the F, G components have the form Σ_e. The main theorem of the paper can now be stated.

Theorem 2: Suppose that

$$\mathcal{G}(z) = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right], \quad A_{11} \in \mathbb{R}^{r\times r},$$

is asymptotically stable and has balanced extended Gramians (P, Σ_e) and (Q, Σ_e) where $\Sigma_e = \operatorname{diag}\{\Sigma_{e,1}, \Sigma_{e,2}\}$ and

$$\begin{aligned} \Sigma_{e,1} &= \operatorname{diag}\{\sigma_{e,1}, \ldots, \sigma_{e,r}\}, \\ \Sigma_{e,2} &= \operatorname{diag}\{\sigma_{e,r+1}, \ldots, \sigma_{e,n}\}. \end{aligned}$$

Then the truncated system

$$\hat{\mathcal{G}}(z) = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right],$$

is asymptotically stable, has balanced extended Gramians $(P_{11}, \Sigma_{e,1})$ and $(Q_{11}, \Sigma_{e,1})$, and

$$\|\mathcal{G} - \hat{\mathcal{G}}\|_\infty \le 2 \sum_{i=r+1}^{n} \sigma_{e,i}.$$

Proof: The theorem follows by iteratively applying Lemma 1 and Lemma 2 to the balanced realization. The error bound then follows from the triangle inequality. ∎

Remark 2: It is the F, G components of the extended Gramians that provide the input-output error bound in Theorem 2. Obviously, if we use $F = P$ and $G = Q$, the result reduces to standard balanced truncation, see Theorem 7.11 in [16]. The point is that the normal Gramians P, Q do not appear in the result here other than as extra decision variables in the extended inequalities. Thus we could expect that the reduced models coming from extended balanced truncation are at least as good as those coming from regular balanced truncation.

VI. Example

Here we apply balanced truncation and extended balanced truncation to two different linear systems. The Gramians are computed using SeDuMi [17] and YALMIP [18]. For balanced truncation, we solve

$$\begin{aligned} &\min \operatorname{Tr} P \quad \text{subject to (2)} \\ &\min \operatorname{Tr} Q \quad \text{subject to (3)} \end{aligned}$$

TABLE I

COMPARISON OF REGULAR ($\hat{\mathcal{G}}_b$) AND EXTENDED ($\hat{\mathcal{G}}_{eb}$) BALANCED TRUNCATION IN EXAMPLE 1.

r	$\|\mathcal{G}-\hat{\mathcal{G}}_b\|_\infty$	Upper bound	$\|\mathcal{G}-\hat{\mathcal{G}}_{eb}\|_\infty$	Upper bound
1	$3.84\cdot 10^{-1}$	$5.42\cdot 10^{-1}$	$3.85\cdot 10^{-1}$	$5.04\cdot 10^{-1}$
2	$5.63\cdot 10^{-2}$	$7.55\cdot 10^{-2}$	$5.58\cdot 10^{-2}$	$7.12\cdot 10^{-2}$
3	$3.18\cdot 10^{-3}$	$4.41\cdot 10^{-3}$	$3.04\cdot 10^{-3}$	$4.71\cdot 10^{-3}$
4	$5.72\cdot 10^{-5}$	$8.19\cdot 10^{-5}$	$5.52\cdot 10^{-5}$	$1.46\cdot 10^{-4}$

and for extended balanced truncation we solve

$$\begin{aligned}&\min \operatorname{Tr} F \quad \text{subject to } (4) \text{ and } F=F^T\\ &\min \operatorname{Tr} G \quad \text{subject to } (5) \text{ and } G=G^T.\end{aligned}$$

The conditions $F=F^T$ and $G=G^T$ are there to guarantee that the extended Gramians can be balanced, see Lemma 4. We have chosen to minimize the trace of the Gramians for simplicity. One could consider more complicated objective functions since we really want to make the (non-convex) singular values $\sigma_i^2=\lambda_i(PQ)$ and $\sigma_{e,i}^2=\lambda_i(FG)$ small. The minimization of the trace of Gramians can be justified as in [19]. We have that

$$\begin{aligned}\sum_{i=1}^n \sigma_i^2 &= \operatorname{Tr}\,(PQ) \le (\operatorname{Tr}\,P)(\operatorname{Tr}\,Q),\\ \sum_{i=1}^n \sigma_{e,i}^2 &= \operatorname{Tr}\,(FG) \le (\operatorname{Tr}\,F)(\operatorname{Tr}\,G),\end{aligned} \tag{13}$$

see Proposition 1 in [19]. If the traces of the Gramians are small, then the singular values are small. The number of decision variables in extended balanced truncation is twice as large as for balanced truncation (P,F,Q,G vs. P,Q). This increases the computation time, but can also lead to better reduced models. Since the set of admissible F is greater or equal to the set of admissible P (Theorem 1), we know that $\min \operatorname{Tr} F \le \min \operatorname{Tr} P$, and similarly $\min \operatorname{Tr} G \le \min \operatorname{Tr} Q$. Hence the upper bound in (13) is smaller for extended balanced truncation. However, this does not mean that the extended Hankel singular values necessarily are smaller than the regular Hankel singular values, since this is just an upper bound (see Example 1). An interesting problem is to find better objective functions that ensure smaller extended singular values.

Example 1: In [5], balanced truncation is applied to the model

$$\mathcal{G}(z) = \frac{(-1.7328\cdot 10^{-3})(z+1.8381)(z-0.3321)(z-0.2813)(z-0.1667)}{(z-0.7125)^5},$$

and we use the same model here. In Fig. 1, the regular and extended Hankel singular values are shown. We note that $\sigma_{e,i}<\sigma_i$ for $i=1,2,3$. For $i=4,5$, $\sigma_{e,i}$ is slightly larger than σ_i. Even if $\operatorname{Tr} F\le \operatorname{Tr} P$ and $\operatorname{Tr} G \le \operatorname{Tr} Q$ it can happen that particular extended singular values are larger than the regular ones.

The approximation error and the upper bounds for $\hat{\mathcal{G}}_b$ (balanced truncation) and $\hat{\mathcal{G}}_{eb}$ (extended balanced truncation) are shown in Table I for various approximation orders r. In this example, extended balanced truncation gives a smaller

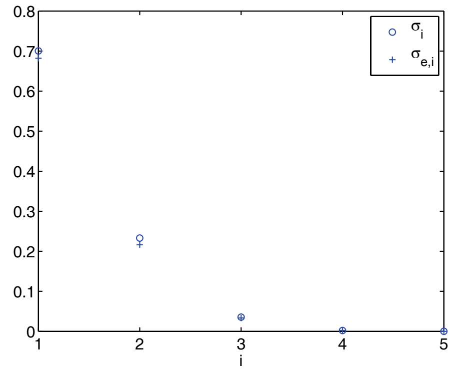

Fig. 1. The Hankel singular values and the extended Hankel singular values in Example 1.

approximation error in all cases, except when $r=1$ when it is 0.3% larger. The error bounds are not always tighter, however.

Example 2: In this example, normal and extended balanced truncation is applied to a frequency-weighted model reduction problem. This is a simple example of a structured model reduction problem, see [10]. The model we would like to reduce is a resonant 16-th order system,

$$\mathcal{G}(s) = \sum_{k=1}^{8} \frac{\omega_k^2}{s^2+2\xi\omega_k s+\omega_k^2},$$

where $\omega_k \in \{1,2,10,20,30,40,45,50\}$ and $\xi=0.1$. The frequency weight $\mathcal{W}(s)$ is here a band-pass filter, $\mathcal{W}(s)=s/(s/5+1)^2$. To get discrete-time models, $\mathcal{G}(s)$ and $\mathcal{W}(s)$ are discretized using "matched" sampling in Matlab with the sampling period 0.05. With slight abuse of notation, the sampled models are denoted $\mathcal{G}(z)$ and $\mathcal{W}(z)$. The problem is now to find a reduced order system $\hat{\mathcal{G}}$ such that $\|\mathcal{W}(\mathcal{G}-\hat{\mathcal{G}})\|_\infty$ is small. We realize the weighted system as

$$\mathcal{W}(z)\mathcal{G}(z) = \left[\begin{array}{cc|c} A_\mathcal{G} & 0 & B_\mathcal{G}\\ B_\mathcal{W}C_\mathcal{G} & A_\mathcal{W} & 0\\ \hline 0 & C_\mathcal{W} & 0\end{array}\right] =: \left[\begin{array}{c|c} A & B\\ \hline C & 0\end{array}\right],$$

where $(A_\mathcal{G},B_\mathcal{G},C_\mathcal{G})$ and $(A_\mathcal{W},B_\mathcal{W},C_\mathcal{W})$ are realizations of $\mathcal{G}(z)$ and $\mathcal{W}(z)$, respectively. One way to solve the problem is to use a discrete-time version of the method in [11] and search for Gramians with small trace that satisfy (2)–(3) with the structure $P=\operatorname{diag}\{P_\mathcal{G},P_\mathcal{W}\}, Q=\operatorname{diag}\{Q_\mathcal{G},Q_\mathcal{W}\}$, where $P_\mathcal{G},Q_\mathcal{G}\in\mathbb{R}^{16\times16}$ and $P_\mathcal{W},Q_\mathcal{W}\in\mathbb{R}^{2\times2}$. The components $P_\mathcal{G},Q_\mathcal{G}$ are then used to balance the realization of $\mathcal{G}(z)$ which then can be truncated as before, and the singular values are computed as $\sigma_i=\sqrt{\lambda_i(P_\mathcal{G}Q_\mathcal{G})}$. The a priori error bound becomes $\|\mathcal{W}(\mathcal{G}-\hat{\mathcal{G}})\|_\infty \le 2\sum_{i=r+1}^n \sigma_i$.

Using extended balanced truncation, we search for Gramians with small trace that satisfy (4)–(5) with the structure $F=\operatorname{diag}\{F_\mathcal{G},F_\mathcal{W}\}, G=\operatorname{diag}\{G_\mathcal{G},G_\mathcal{W}\}$, where $F_\mathcal{G},G_\mathcal{G}\in\mathbb{R}^{16\times16}$ and $F_\mathcal{W},G_\mathcal{W}\in\mathbb{R}^{2\times2}$. The components P and Q are full matrices. The components $F_\mathcal{G},G_\mathcal{G}$ are then used to balance the realization of $\mathcal{G}(z)$, and the extended Hankel sin-

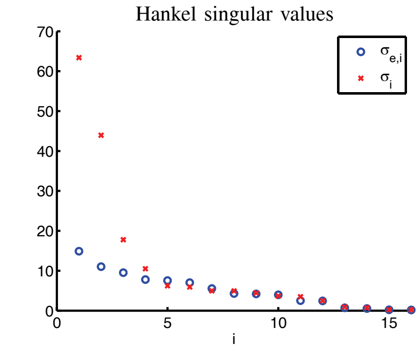

Fig. 2. The normal and extended Hankel singular values in Example 2.

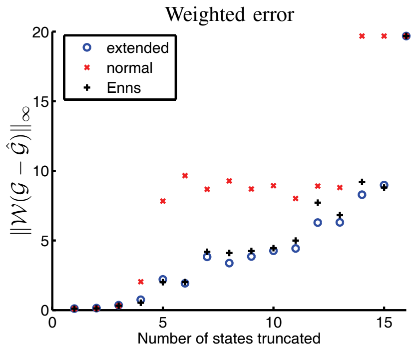

Fig. 3. The weighted approximation error in Example 2 for various approximation orders.

gular values are given by $\sigma_{e,i} = \sqrt{\lambda_i(F_{\mathcal{G}}G_{\mathcal{G}})}$. The a priori error bound becomes $\|\mathcal{W}(\mathcal{G}-\hat{\mathcal{G}})\|_\infty \leq 2\sum_{i=r+1}^{n}\sigma_{e,i}$.

In Fig. 2, the Hankel singular values from the two methods mentioned above are plotted. It is noted that the largest extended Hankel singular values, $\sigma_{e,i}$, are much smaller than the largest Hankel singular values σ_i. The singular values can be used to bound the weighted error for various approximation orders r. Approximation errors for the methods are shown in Fig. 3 for various numbers of truncated states. For comparison, Enns' method [3] is also included. As can be seen, extended balanced truncation delivers approximations that are about as good as the ones from Enns' method and often much better than the ones from the method in [11]. It should be noted that even though Enns' method gives as good approximations in this case as the extended method, the extended method also comes with a simple error bound. No simple error bounds exist for Enns' method, which is also not guaranteed to deliver good approximations.

VII. Conclusions

An extension to the balanced truncation method has been presented. All important properties of balanced truncation, such as stability preservation and a priori error bounds, hold in the extended case also. In two examples, it was shown that extended balanced truncation generally gives better approximations. More important, it is more likely that extended balanced truncation can preserve internal structures in the reduced models. This is important in controller reduction and in model reduction of networked systems, and this will be further investigated in future work. Another interesting problem is to find better objective functions of the extended Gramians to be minimized.

Acknowledgment

I would like to thank Prof. Roy Smith for bringing my attention to the papers [12], [13].

References

[1] B. C. Moore, "Principal component analysis in linear systems: controllability, observability, and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, no. 1, pp. 17–32, February 1981.

[2] L. Pernebo and L. Silverman, "Model reduction via balanced state space representation," *IEEE Transactions on Automatic Control*, vol. 27, pp. 382–387, 1982.

[3] D. Enns, "Model reduction with balanced realizations: an error bound and a frequency weighted generalization," in *Proceedings of the IEEE Conference on Decision and Control*, Las Vegas, Nevada, 1984.

[4] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L_∞-error bounds," *International Journal of Control*, vol. 39, pp. 1115–1193, 1984.

[5] U. Al-Saggaf and G. Franklin, "An error-bound for a discrete reduced-order model of a linear multivariable system," *IEEE Transactions on Automatic Control*, vol. 32, no. 9, pp. 815–819, 1987.

[6] D. Hinrichsen and A. Pritchard, "An improved error estimate for reduced-order models of discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 3, pp. 317–320, 1990.

[7] K. Zhou, C. D'Souza, and J. Cloutier, "Structurally balanced controller order reduction with guaranteed closed loop performance," *Systems and Control Letters*, vol. 24, pp. 235–242, 1995.

[8] C. Beck, J. Doyle, and K. Glover, "Model reduction of multidimensional and uncertain systems," *IEEE Transactions on Automatic Control*, vol. 41, no. 10, pp. 1466–1477, October 1996.

[9] L. Li and F. Paganini, "Structured coprime factor model reduction based on LMIs," *Automatica*, vol. 41, no. 1, pp. 145–151, Jan. 2005.

[10] A. Vandendorpe and P. Van Dooren, *Model Order Reduction: Theory, Research Aspects and Applications*, ser. Mathematics in Industry Series. Springer Verlag, 2007, ch. Model reduction of interconnected systems.

[11] H. Sandberg and R. M. Murray, "Model reduction of interconnected linear systems using structured Gramians," in *Proccedings of the IFAC World Congess*, 2008, to appear. Available from `http://www.ee.kth.se/~hsan/ifac2008.pdf`.

[12] M. De Oliviera, J. Bernussou, and J. Geromel, "A new discrete-time robust stability condition," *Systems & Control Letters*, vol. 37, pp. 261–265, 1999.

[13] M. De Oliviera, J. Geromel, and J. Bernussou, "Extended H_2 and H_∞ norm characterizations and controller parametrizations for discrete-time systems," *International Journal of Control*, vol. 75, no. 9, pp. 666–679, 2002.

[14] M. Subbotin and R. Smith, "Design of a distributed decentralized estimator for a formation with fixed communication topology," in *Proceedings of the 2007 American Control Conference*, New York City, New York, July 2007, pp. 5025–5030.

[15] ——, "Distributed estimator design for a formation with Markovian communication topology," in *Proceedings of the 2007 American Control Conference*, New York City, New York, July 2007, pp. 5046–5051.

[16] K. Zhou and J. Doyle, *Essentials of Robust Control*. Upper Saddle River, New Jersey: Prentice Hall, 1998.

[17] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11–12, pp. 625–653, 1999, available from `http://fewcal.kub.nl/sturm/software/sedumi.html`.

[18] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004, available from `http://control.ee.ethz.ch/~joloef/yalmip.php`.

[19] P. Bendotti and C. Beck, "On the role of LFT model reduction methods in robust controller synthesis for a pressurized water reactor," *IEEE Transactions on Control Systems Technology*, vol. 7, no. 2, pp. 248–257, March 1999.