Estimating Gramians of large-scale time-varying systems.

Chahlaoui, Younes and Van Dooren, Paul

2002

MIMS EPrint: **2008.21**

Manchester Institute for Mathematical Sciences

School of Mathematics

The University of Manchester

ESTIMATING GRAMIANS OF LARGE-SCALE TIME-VARYING SYSTEMS

Younes Chahlaoui, * **Paul Van Dooren** *

* *Université catholique de Louvain, Belgium*

Abstract: In this paper we present a Smith-like updating technique to estimate a low rank approximation of the Gramians of a time-varying system. We obtain error estimates of our approximation and also explain how to use this for model reduction of time-varying systems.

Keywords: Model reduction, time-varying systems, Gramians

1. INTRODUCTION

Model reduction of large-scale dynamical systems has become very popular in the last decade. The idea is to construct a lower order model that approximates well the behavior of the larger dynamical model. There are many ways to achieve this : an often used technique is that of balanced truncation, introduced by Moore (Moore, 1981), who showed asymptotic stability of the associated reduced-order model for a linear time-invariant system. This approach has the advantage that the states which are difficult to control are also difficult to observe, so that one just truncates those "weak" states to produce a good reduced order model. In (Imae *et al.*, 1992) and (Shokoohi *et al.*, 1983) a generalization to time-varying systems is presented, but without any algorithmic details. Glover (Glover, 1984) established certain error bounds between original plants and reduced-order models in terms of Hankel singular values. Recently, many papers study iterative methods utilizing Krylov-based projection (moment-matching, rational interpolation, Arnoldi, Padé, Lanczos, see (Van Dooren, 2000) for a survey). All the ideas developed for linear systems are based

on the dominant spaces of Gramians (energy functions for -in and outgoing signals), which are the solutions of Lyapunov or Stein equations. A lot of work is still needed to efficiently compute these solutions (or their dominant spaces) when the system matrices are large and sparse. In fact, most direct methods ignore sparsity in the Lyapunov or Stein equations and are not very attractive for parallelization. Their use is therefore limited if large, sparse equations have to be solved. The complexity of these methods is roughly $\Theta(N^3)$ floating point operations and they require about $\Theta(N^2)$ words of memory.

2. TIME-VARYING SYSTEMS

Discrete-time case
We first consider a discrete-time system

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k \\ \quad\; y_k = C_k x_k \end{cases} \tag{1}$$

with input $u_k \in \mathbb{R}^m$, state $x_k \in \mathbb{R}^N$ and output $y_k \in \mathbb{R}^p$, and $m, p << N$.
The input sequence is assumed to be square-summable, i.e $u_k \in l_2^m$ (this is a Hilbert space with inner product $\langle x, y \rangle_{l_2^m} \doteq \sum_{-\infty}^{\infty} x_k^T y_k$), and we assume that $\{A_k\}_0^{\infty}$, $\{B_k\}_0^{\infty}$, and $\{C_k\}_0^{\infty}$ are bounded sequences of matrices with appropriate dimensions. (A sequence of matrices $\{M_k\}$ is said to be bounded if there exists a constant $M \in \mathbb{R}$

Using the recurrence (1) over several time steps, one obtains the state at step k in function of past inputs over the interval $[k_0, k]$:

$$x_k = \Phi(k, k_0)x_{k_0} + \sum_{i=k_0}^{k-1} \Phi(k, i+1)B_i u_i$$

where $\Phi(k, k_0) \doteq A_{k-1} \dots A_{k_0}$ is the discrete transition matrix over time period $[k_0, k]$. The transition matrix has the following properties :

$$\begin{cases} \Phi(k, k_0) = \Phi(k, k_1)\Phi(k_1, k_0), \ k_0 \le k_1 \le k \\ \Phi(k, k) = I_N \quad \forall k \end{cases}$$

We will assume the time-varying system to be asymptotically stable, which means that

$$\|\Phi(k, k_0)\| \le m \cdot a^{(k-k_0)}, \text{ with } m > 0, 0 < a < 1.$$

Under such conditions one can define the Gramians over any interval $[k_0, k_f]$ as follows :

$$\mathcal{G}_c(k) = \sum_{i=k_0}^{k-1} \Phi(k, i+1)B_i B_i^T \Phi^T(k, i+1),$$

$$\mathcal{G}_o(k) = \sum_{i=k}^{k_f} \Phi^T(i, k)C_i^T C_i \Phi(i, k).$$

(Notice that the asymptotic stability is needed when $k_0 = -\infty$ or $k_f = +\infty$.)

These Gramians can also be obtained from the Stein recurrence formulas:

$$\mathcal{G}_c(k+1) = A_k \mathcal{G}_c(k)A_k^T + B_k B_k^T, \qquad (2)$$

and

$$\mathcal{G}_o(k) = A_k^T \mathcal{G}_o(k+1)A_k + C_k^T C_k, \qquad (3)$$

with respective initial conditions

$$\mathcal{G}_c(k_0) = \mathbf{0}, \quad \mathcal{G}_o(k_f + 1) = \mathbf{0}.$$

Continuous-time case
The continuous-time case is very similar. Consider the linear finite-dimensional dynamical system represented by:

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ y(t) = C(t)x(t) \end{cases} \qquad (4)$$

with input $u(t) \in \mathbb{R}^m$, state $x(t) \in \mathbb{R}^N$ and output $y(t) \in \mathbb{R}^p$, and $p, m << N$. Here $u(t)$ is assumed to be piecewise continuous, and $A(t)$, $B(t)$, and $C(t)$ are matrices of bounded, continuously differentiable functions with appropriate dimensions. Moreover, the input sequence is assumed to be square-summable, i.e $u(t) \in l_2^m$ with an inner product that is now defined as

$$\langle x, y \rangle_{l_2^m} \doteq \int_{-\infty}^{+\infty} x(t)^T y(t) dt .$$

The state as function of past inputs can be derived from (4) and is given by :

$$\int^t$$

where $\Phi(t, t_0)$ is the transition matrix associated with the homogeneous equation $\dot{x}(t) = A(t)x(t)$. For an asymptotically stable system we also have

$$\|\Phi(t, t_0)\| \le m \cdot e^{-a(t-t_0)}, \text{ with } m > 0, a > 0.$$

We can then define the controllability and observability Gramians over any interval $[t_o, t_f]$ by:

$$\mathcal{G}_c(t) = \int_{t_0}^{t} \Phi(t, s)B(s)B^T(s)\Phi^T(t, s)ds$$

$$\mathcal{G}_o(t) = \int_{t}^{t_f} \Phi^T(s, t)C^T(s)C(s)\Phi(s, t)ds$$

(Notice again that asymptotic stability is needed when $t_0 = -\infty$ or $t_f = +\infty$.)

These Gramians also satisfy the Lyapunov differential equations:

$$\frac{\partial \mathcal{G}_c(t)}{\partial t} = A(t)\mathcal{G}_c(t) + \mathcal{G}_c(t)A^T(t) + B(t)B^T(t) \quad (5)$$

and

$$\frac{\partial \mathcal{G}_o(t)}{\partial t} = -A^T(t)\mathcal{G}_o(t) - \mathcal{G}_o(t)A(t) - C^T(t)C(t) \tag{6}$$

with respective initial conditions

$$\mathcal{G}_c(t_0) = \mathbf{0}, \quad \mathcal{G}_o(t_f) = \mathbf{0}.$$

All results in continuous-time and discrete-system are very similar, except that discrete transition matrices and summations are replaced by continuous transition matrices and integrations, and that the sequence $[k_0, k_f]$ is replaced by the interval $[t_0, t_f]$.

Conversion to discrete-time
The numerical scheme proposed in later sections is entirely based on discrete-time models. We show that it can be applied to continuous-time systems as well by using an appropriate discretization.

Let us apply Euler's explicit scheme to the solution of the ordinary differential equations (5,6), then for step size h we obtain :

$$[\mathcal{G}_c((k+1)h) - \mathcal{G}_c(kh)]/h =$$
$$A(kh)\mathcal{G}_c(kh) + \mathcal{G}_c(kh)A^T(kh) + B(kh)B(kh)^T,$$

and

$$-[\mathcal{G}_o((k+1)h) - \mathcal{G}_o(kh)]/h =$$
$$A(kh)^T \mathcal{G}_o((k+1)h) + \mathcal{G}_o((k+1)h)A(kh) + C(kh)^T C(kh)$$

with respective initial conditions

$$\mathcal{G}_c(t_o) = \mathbf{0}, \quad \mathcal{G}_o(t_f) = \mathbf{0}.$$

One easily checks that up to $O(h^2)$ this is equivalent to solving the recurrences (2,3) with the correspondences

$$A_k \doteq (I + hA(kh)), \ B_k \doteq \sqrt{h}B(kh),$$

$$C_k \doteq \sqrt{h}C(kh).$$

We are thus approximating the continuous-time Gramians at time steps $\mathcal{G}_c(t_o + kh)$ and $\mathcal{G}_o(t_o$

3. BALANCED TRUNCATION

The method of balanced truncation of linear systems is well established for model reduction. For linear time-invariant systems, the approach requires only matrix computations, and has been successfully used in control systems design. A-priori error bounds in the induced 2-norm can be given for the error between the original and the reduced system (Zhou *et al.*, 1995).

For linear time-varying systems, the procedure is analogous to the linear time-invariant case but requires the solution of difference or differential equations in order to obtain the Gramians. From (Shokoohi *et al.*, 1983), any uniformly controllable and observable linear system has a uniformly balanced realization, i.e., the product of the Gramians has an eigenvalue decomposition of the form:

$$T(t)^{-1}\mathcal{G}_c(t)\mathcal{G}_o(t)T(t) = \tilde{\mathcal{G}}_c(t)\tilde{\mathcal{G}}_o(t) = \Sigma^2(t)$$

where $T(t)$ is a Lyapunov transformation (bounded, with bounded inverse and continuous derivative), $\Sigma(t)$ is diagonal and $\tilde{\mathcal{G}}_c(t)$, $\tilde{\mathcal{G}}_o(t)$ are the Gramians of the system $\{\tilde{A}(t), \tilde{B}(t), \tilde{C}(t)\}$ obtained via the state space transformation $T(t)$. This new system is uniformly balanced and can then be truncated to obtain the reduced order model.

The results for the discrete-time case are very similar. There exists a state space transformation T_k (uniformly bounded with uniformly bounded inverse) such that the transformed Gramians satisfy

$$T_k^{-1}\mathcal{G}_c(k)\mathcal{G}_o(k)T_k = \tilde{\mathcal{G}}_c(k)\tilde{\mathcal{G}}_o(k) = \Sigma^2(k).$$

One also shows that if there is a uniform gap between "large" and "small" singular values, the constructed reduced model is asymptotically stable, and uniformly controllable and observable.

As pointed out earlier, it is more convenient to work in the discrete-time setting. Rather than computing the complete transformation T_k, one can try to estimate only the first few columns of T_k, i.e. a matrix $X_k \in \mathbb{R}^{N \times n}$ whose columns span the "dominant" eigenvectors of the product $\mathcal{G}_c(k)\mathcal{G}_o(k)$.

Similarly, one will need a matrix $Y_k \in \mathbb{R}^{N \times n}$ whose columns span the "dominant" eigenvectors of the product $\mathcal{G}_o(k)\mathcal{G}_c(k)$.

One then obtains the reduced model for the system $\{A_k, B_k, C_k\}$ as $\{Y_k^T A_k X_k, Y_k^T B_k, C_k X_k\}$, provided the bases were normalized using $Y_k^T X_k = I_n$ (see (Van Dooren, 2000) for more details). The rest of the paper now focuses on the case were the original system is large and sparse. This is typically the case if it originates from a spatial discretization of partial differential equations. We refer to (Verlaan and Heemink, 1997) for such a typical time-varying example.

When large scale systems are sparse it is impor-

ciency. Iterative methods are very suitable for this and are often easy to parallelize as well. Balanced truncation requires the solution of large-scale Lyapunov or Stein equations in order to obtain the Gramians. Since these are dense matrices this will typically costs $\Theta(N^3)$ floating point operations and $\Theta(N^2)$ words of memory (Penzl, 2000). But when model reduction is to be used, the Gramians have often rapidly decaying eigenvalues (Antoulas *et al.*, 2001*b*), which suggests to approximate the Gramians at each step by a low-rank factorization. We show below how to obtain such approximations and at the same time exploit the sparsity of the model $\{A_k, B_k, C_k\}$. Although all material below should be applied to both Gramians $\mathcal{G}_c(k)$ and $\mathcal{G}_o(k)$ we will focus on the controllability Gramian only.

Smith method
A popular approach for the time-invariant case is the Smith method. This method computes recursive approximations to the solution of the Stein equation:

$$\mathcal{G}_c(k+1) = A\mathcal{G}_c(k)A^T + BB^T$$

which yields, for $\mathcal{G}_c(0) = \mathbf{0}$, the iterates :

$$\mathcal{G}_c(k) = \sum_{i=1}^{k} \Phi(k,i)BB^T\Phi(k,i)^T$$

converging to the solution G of

$$G = AGA^T + BB^T.$$

The Smith method converges linearly but has a drawback that it computes the solution in dense form and hence requires $\Theta(N^2)$ storage. To take advantage of the ("approximate") low-rank property of the Gramian, one has to use the so called low-rank Smith method.

Low-rank Smith methods
The key idea of the low-rank method is to substitute the iterates by their symmetric factorization :

$$\mathcal{G}_c(k) = S_k S_k^T.$$

This is always possible because $\mathcal{G}_c(k)$ can be shown to be symmetric and positive semi-definite. This method still has the drawback that the number of columns of S_k grows linearly at each iteration. It is suggested in (Antoulas *et al.*, 2001*a*) to use instead a low-rank approximation

$$\mathcal{G}_c(k) \approx \tilde{S}_k \tilde{S}_k^T, \quad \tilde{S}_k \in \mathbb{R}^{N \times n},$$

at each step of the recursion. It is precisely such a scheme that we develop in the next section for the time-varying case.

4. LOW-RANK GRAMIANS

We consider the following time-varying equation:

The main goal of a low-rank version is to approximate the square root of $\mathcal{G}_c(k)$ by a rank n_k approximation at each iteration (typically n_k is constant).

Let $\tilde{S}_0 \in \mathbb{R}^{N \times n}$ be an initialization satisfying:

$$S_0 S_0^T = \tilde{S}_0 \tilde{S}_0^T + E_0 E_0^T$$

(we can always choose, e.g. $\tilde{S}_0 = \mathbf{0}$, which would imply $E_0 = S_0$), then the kth low-rank approximation $\tilde{S}_k$ is obtained as follows :

$$\begin{aligned}
\hat{S}_k &= \left[\, A_{k-1}\tilde{S}_{k-1}\,\middle|\,B_{k-1}\,\right] = U\Sigma V^T \\
\tilde{S}_k &\doteq \hat{S}_k V(:,1:n) \\
E_k &\doteq \hat{S}_k V(:,n+1:n+m)
\end{aligned} \qquad (8)$$

where $U \in \mathbb{R}^{N \times (n+m)}$, $\Sigma \in \mathbb{R}^{(n+m) \times (n+m)}$ and $V \in \mathbb{R}^{(n+m) \times (n+m)}$ are the Short Singular Values Decomposition (SSVD) matrices of $\hat{S}_k$.

Let us define $\mathcal{P}_k$ as the low-rank approximation of the Gramian $\mathcal{G}_c(k)$, i.e. $\mathcal{P}_k = \tilde{S}_k \tilde{S}_k^T$. From the iteration (8) it follows that $\mathcal{P}_k$ is the best rank n approximation to $\hat{S}_k \hat{S}_k^T$. But this is not sufficient since we want to compare $\mathcal{G}_c(k) = S_k S_k^T$ with $\mathcal{P}_k = \tilde{S}_k \tilde{S}_k^T$. This is analyzed below.

Convergence properties
Define $S_k = \left[\, B_{k-1}\,\middle|\,A_{k-1}B_{k-2}\,\middle|\ldots\middle|\,\Phi(k,0)S_0\,\right]$ then clearly $\mathcal{G}_c(k) = S_k S_k^T$. We now show the following result :

Theorem 1.
At each iteration, there exists an orthogonal matrix $V_k \in \mathbb{R}^{(n+m) \times (n+m)}$ satisfying :

$$S_k V_k = \left[\, \tilde{S}_k\,\middle|\,E_k\,\middle|\,A_{k-1}E_{k-1}\,\middle|\ldots\middle|\,\Phi(k,0)E_0\,\right]$$

where E_k is the neglected part at iteration k (8).

Proof
At each step, there exists an orthogonal matrix $\tilde{V}_{k+1}$ such that $\left[\, B_k\,\middle|\,A_k\tilde{S}_k\,\right]\tilde{V}_{k+1} = \left[\, \tilde{S}_{k+1}\,\middle|\,E_{k+1}\,\right]$. For $k=0$ we have $S_0 = \left[\, \tilde{S}_0\,\middle|\,E_0\,\right]$.
We prove the general result by induction. Suppose that there exists an orthogonal matrix V_k such that

$$S_k V_k = \left[\, \tilde{S}_k\,\middle|\,E_k\,\middle|\,A_{k-1}E_{k-1}\,\middle|\ldots\middle|\,\Phi(k,0)E_0\,\right].$$

Since $S_{k+1} = \left[\, B_k\,\middle|\,A_k S_k\,\right]$, we choose

$$V_{k+1} = \begin{bmatrix} I_m & 0 \\ 0 & V_k \end{bmatrix}\begin{bmatrix} \tilde{V}_{k+1} & 0 \\ 0 & I_l \end{bmatrix}, \quad (l=(k+1)m)$$

from which it follows that

$$\begin{aligned}
S_{k+1}V_{k+1} &= \left[\, B_k\,\middle|\,A_k S_k\,\right]\begin{bmatrix} I_m & 0 \\ 0 & V_k \end{bmatrix}\begin{bmatrix} \tilde{V}_{k+1} & 0 \\ 0 & I_l \end{bmatrix} \\
&= \left[\, B_k\,\middle|\,A_k S_k V_k\,\right]\begin{bmatrix} \tilde{V}_{k+1} & 0 \\ 0 & I_l \end{bmatrix} \\
&= \left[\, B_k\,\middle|\,A_k\tilde{S}_k\,\middle|\,A_k E_k\,\middle|\ldots\middle|\,\Phi(k+1,0)E_0\,\right]\begin{bmatrix} \tilde{V}_{k+1} & 0 \\ 0 & I_l \end{bmatrix} \\
&= \left[\, \tilde{S}_{k+1}\,\middle|\,E_{k+1}\,\middle|\,A_k E_k\,\middle|\ldots\middle|\,\Phi(k+1,0)E_0\,\right].
\end{aligned}$$

We can use this to compare the solution $\mathcal{G}_c(k)$ of (7) with $\mathcal{P}_k$ as follows

$$\mathcal{G}_c(k) = S_k S_k^T =$$

$$\tilde{S}_k \tilde{S}_k^T + E_k E_k^T + \sum_{i=0}^{k-1} \Phi(k,i)E_i E_i^T \Phi(k,i)^T.$$

Since $\Phi(k,k) = I_N$ it follows that

$$\mathcal{G}_c(k) = \mathcal{P}_k + \sum_{i=0}^{k} \Phi(k,i)E_i E_i^T \Phi(k,i)^T. \qquad (9)$$

Taking norms we obtain

$$\|\mathcal{G}_c(k) - \mathcal{P}_k\|_2 \leq \sum_{i=0}^{k} \|\Phi(k,i)\|_2^2 \|E_i\|_2^2$$

and if we define $\eta = \max_{0 \leq i \leq \infty} \|E_i\|_2$, we obtain the simple inequality

$$\|\mathcal{G}_c(k) - \mathcal{P}_k\|_2 \leq \eta^2 \sum_{i=0}^{k} \|\Phi(k,i)\|_2^2. \qquad (10)$$

Since we have supposed that our system is asymptotically stable, we have:

$$\|\Phi(k,k_0)\| \leq m \cdot a^{(k-k_0)}, \text{ with } m > 0, 0 < a < 1.$$

and (10) becomes :

$$\|\mathcal{G}_c(k) - \mathcal{P}_k\|_2 \leq \eta^2 m^2 \sum_{i=0}^{k} a^{2i} \leq \frac{\eta^2 m^2}{1-a^2}.$$

When $k \to \infty$ this bounds the difference between $\mathcal{P}_\infty$ and $\mathcal{G}_c(\infty)$.

The time-invariant case
For linear-time invariant systems $\{A, B, C\}$, the difference between the approximate low-rank Gramian $\mathcal{P}_k$ and the exact Gramian $\mathcal{G}_c(k)$,

$$\mathcal{E}_k \doteq \mathcal{G}_c(k) - \mathcal{P}_k$$

remains bounded for large k.

Theorem 2.
Let $\mathcal{P}$ be the solution of $\mathcal{P} = A\mathcal{P}A^T + I$ then

$$\|\mathcal{E}_k\|_2 \leq \eta^2 \|\mathcal{P}\|_2 \leq \eta^2 \frac{\kappa(A)^2}{1-\rho(A)^2} \qquad (11)$$

Proof
It follows from (9) that

$$\mathcal{E}_{k+1} = A\mathcal{E}_k A^T + E_k E_k^T.$$

With $\eta = \max_{0 \leq k \leq \infty} \|E_k\|_2$, we can also consider the equation:

$$\mathcal{X}_{k+1} = A\mathcal{X}_k A^T + (\eta^2 I - E_k E_k^T), \quad \mathcal{X}_0 = 0.$$

Its iterates $\mathcal{X}_k$ are clearly positive semi-definite and hence converge to a solution $\mathcal{X}$ which is also positive semi-definite. Moreover by linearity we have

It then follows that

$$\lim_{k\to\infty} \mathcal{E}_k + \mathcal{X}_k = \eta^2 \mathcal{P}$$

and we obtain $\quad \|\mathcal{E}_k\|_2 \le \eta^2 \|\mathcal{P}\|_2$.
The second bound follows from the eigen-decomposition of A. $\qquad\square$

Remarks
− Our bounds are very similar to those obtained in (Antoulas *et al.*, 2001*a*) for the time-invariant case;
− In (11), $\frac{\kappa(A)^2}{1-\rho(A)^2}$ is constant and it is very interesting when $\rho(A) \ll 1$ and $k(A)$ is reasonable;
− As stopping criterion, we can predefine a tolerance ϵ_m and test if $\eta \le \epsilon_m$ for several iterations;
− η can be taken equal to the maximum of $\|E_i\|_2$ for $k_0 \le i \le \infty$, since we can interpret theorems 1 and 2 as starting with step k_0 rather than 0. This is particulary useful if after step k_0 the errors have converged to their minimal value, i.e. the convergence threshold ϵ_m;
− At each iteration, we need to compute only $U_1 = U(:,1:n)$ and η. This requires typically $\Theta(N(n+m)^2)$ flops (Golub and Loan, 1996), which is more reasonable than the complexity given in the introduction (when $N \gg m, n$);
− A reduced model is obtained as :
$\{U_1^T A U_1, U_1^T B, C U_1\}$ where U_1 is the dominant part of the last orthogonal matrix obtained by (8).

The periodic case
A K-periodic system, i.e. a system such that

$$\{A_{K+k}, B_{K+k}, C_{K+k}\} = \{A_k, B_k, C_k\}$$

is equivalent (Tsakalis and Ioannou, 1993) to a "lifted" time-invariant system

$$\begin{cases} \hat{x}(k+1) = \hat{A}\hat{x}(k) + \hat{B}u(k) \\ \hat{y}(k) = \hat{C}\hat{x}(k) + \hat{D}u(k) \end{cases} \quad (12)$$

with appropriately defined system matrices :

$$\hat{A} = \Phi(K,1), \ \hat{B} = \sum_{i=1}^{K} \Phi(K, i+1)B_i,$$

$$\hat{C} = C_K \Phi(K-1,1), \ \hat{D} = C_K \sum_{i=1}^{K} \Phi(K, i+1)B_i.$$

For such systems our analysis can be extended as follows :

Theorem 3.
Let $\mathcal{P}$ be the solution of $\mathcal{P} = \tilde{A}\mathcal{P}\tilde{A}^T + I_{KN}$ where $\mathcal{P} \doteq diag(\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_{K-1}, \mathcal{P}_0)$, and

$$\tilde{A} \doteq \begin{pmatrix} 0 & 0 & \ldots & 0 & A_0 \\ A_1 & 0 & 0 & \ldots & 0 \\ 0 & A_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \ldots & 0 & A_{K-1} & 0 \end{pmatrix}$$

then

Proof
The proof is analogous to that of the time invariant case, and uses the fact that for a periodic system there exists a periodic controllability Gramian (Sreedhar and Dooren, 1994). $\qquad\square$

5. NUMERICAL EXAMPLE

Using the multirate sampling data model given in (Tornero *et al.*, 2001), we can construct from any discrete LTI system $\{A, B, C\}$ a periodic time-varying system which can be represented by:

$$\begin{cases} \bar{x}_{k+1} = \bar{A}(k)\bar{x}_k + \bar{B}(k)u_k \\ \bar{y}_k = \bar{C}(k)\bar{x}_k \end{cases}$$

where $\bar{x}_k^T \doteq [x_k^T \ v_k^T] \in \mathbb{R}^{N+m}$ is the enlarged state vector, and the periodic system matrices are given by

$$\bar{A}(k) \doteq \begin{bmatrix} A & B(I - \Delta(k)) \\ 0 & I - \Delta(k) \end{bmatrix}, \ \bar{B}(k) \doteq \begin{bmatrix} B\Delta(k) \\ \Delta(k) \end{bmatrix}$$

$\bar{C}(k) \doteq \bar{\Delta}(k) \begin{bmatrix} C & 0 \end{bmatrix}$. Here $\Delta(.)$ and $\bar{\Delta}(.)$ are N-periodic $m \times m$ and $p \times p$ matrices given by

$$\Delta(k) \doteq diag\{\delta_i(k), i = 1, 2, \ldots, m\}$$
$$\delta_i(k) \doteq \begin{cases} 1, \ k = jN_i, \\ 0, \ k \ne jN_i, \end{cases} \quad j \in \mathbb{Z}^+$$
$$\bar{\Delta}(k) \doteq diag\{\bar{\delta}_i(k), i = 1, 2, \ldots, p\}$$
$$\delta_i(k) \doteq \begin{cases} 1, \ k = jM_i, \\ 0, \ k \ne jM_i, \end{cases} \quad j \in \mathbb{Z}^+,$$

where T is the basic period, output j is updated every M_j periods, input i is sampled every N_i periods and $N = lcm(N_i, M_j)$.

We apply this to the discretized SISO system of the arm of the CD player with a sampling time $T_s = 0.0001$, $N_i = M_i = 2$ and we reduce the order from 122 to 20. For this model we have $\frac{\kappa(\hat{A})^2}{1-\rho(\hat{A})^2} \simeq 10^5$, and $\eta \simeq 10^{-16}$, $\forall k \ge 20$.
In FIG.1 and .2 we show the cosine of the canonical angle between the dominant subspace of two successive iterations $(k-1)$ and k, i.e. $\cos(\angle(\tilde{S}_{k-1}, \tilde{S}_k))$, and the canonical angle with the exact dominant subspace, noted S_∞, of the controllability Gramian of the lifted LTI system (12), i.e. $(\cos(\angle(\tilde{S}_k, S_\infty))$. Those figures show the convergence and the accuracy of our algorithm. It can be seen that we have convergence as soon as the algorithm has eliminated the $n = 20$ vectors of the initial matrix $\tilde{S}_0$. In FIG.3. and .4 we compare frequency responses of the time-invariant lifted systems (12) for starting point 1 FIG.3. and 2 FIG.4.. In each figure we give the amplitude of the frequency response of the original model, the absolute errors in the frequency response after 20 steps and 60 steps, and the absolute errors in the frequency response using the exact dominant

FIG.1. —⊶— $\cos(\sphericalangle(\tilde{S}_k, \tilde{S}_{k-1}))$, —✳— $\cos(\sphericalangle(\tilde{S}_k, S_\infty))$
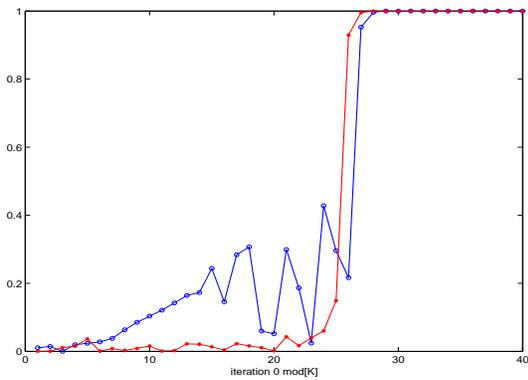

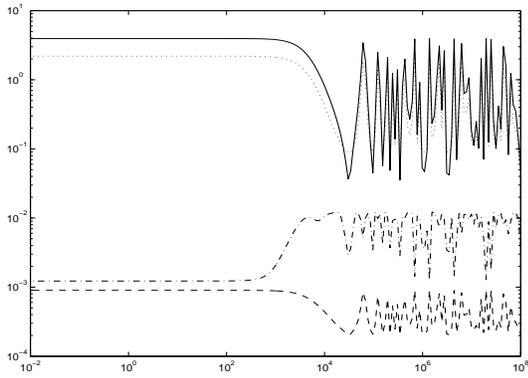FIG.2. —⊶— $\cos(\sphericalangle(\tilde{S}_k, \tilde{S}_{k-1}))$, —✳— $\cos(\sphericalangle(\tilde{S}_k, S_\infty))$

FIG.3: — full model, $\cdots$ approx. errors (20 steps),
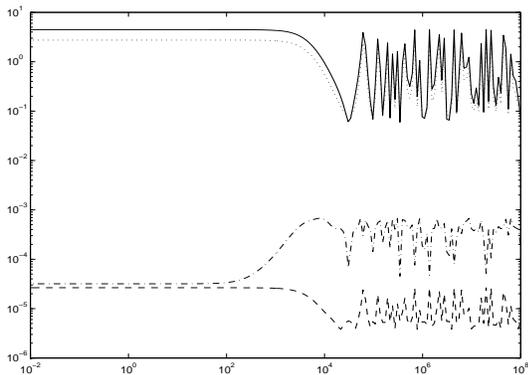-·-· approx. errors (60 steps),
- - approx.errors (exact Gramian)

FIG.4: — full model, $\cdots$ approx. errors (20 steps),

6. REFERENCES

Antoulas, A.C., D.C. Sorenson and S. Gugercin (2001a). A modified low-rank Smith method for large-scale Lyapunov equations. *CAAM Technical Report TR01-10, Rice University.*

Antoulas, A.C., D.C. Sorenson and Y.K. Zhou (2001b). On the decay rate of Hankel singular values and related issues. *CAAM Technical Report TR01-09, Rice University.*

Glover, K. (1984). All optimal Hankel norm approximations of linear multivariable systems and their $\mathcal{L}^\infty$-error bounds. *Internat. J. Control* **39**, 1115–1193.

Golub, G. and C. Van Loan (1996). Matrix computations. *The Johns Hopkins University Press.*

Imae, J., J.E. Perkins and J.B. Moore (1992). Toward time-varying balanced realization via riccati equations. *Math. Control Signals Systems* **5**, 313–326.

Moore, B.C. (1981). Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Automat. Control* **26**, 17–31.

Penzl, T. (2000). A cyclic low-rank Smith method for large sparse Lyapunov equations. *Siam J. Sci. Comput.* **21(4)**, 1404–1418.

Shokoohi, S., L.M. Silverman and P.M. Van Dooren (1983). Linear time-variable systems: Balancing and model reduction. *IEEE Trans. Automat. Control* **28**, 810–822.

Sreedhar, J. and P.M. Van Dooren (1994). A schur approach for solving some periodic matrix equations. systems and networks: Mathematical theory and applications. *Mathematical Research* **77**, 339–362.

Tornero, J., P. Albertos and J. Salt (2001). Periodic optimal control of multirate sampled data systems. *IFAC. Periodic Control Systems. PSYCO2001. Italy* pp. 199–204.

Tsakalis, K.S. and P.A. Ioannou (1993). Linear time-varying systems: Control and adaptation. *Prentice Hall.*

Van Dooren, P. (2000). Gramian based model reduction of large-scale dynamical systems. *in Numerical Analysis 1999* pp. 231–247.

Verlaan, M. and A. Heemink (1997). Tidal flow forecasting using reduced rank square root filters. *Stochastic Hydrology and Hydraulics* **11**, 349–368.

Zhou, K., J.C. Doyle and K. Glover (1995). Robust and optimal control. *Prentice Hall.*