INT. J. CONTROL, 1989, VOL. 50, NO. 4, 1379–1405

Singular perturbation approximation of balanced systems

YI LIU† and BRIAN D. O. ANDERSON†

This paper relates the singular perturbation approximation technique for model reduction to the direct truncation technique if the system model to be reduced is stable, minimal and internally balanced. It shows that these two methods constitute two fully compatible model-reduction techniques for a continuous-time system, and both methods yield a stable, minimal and internally balanced reduced-order system with the same L_∞-norm error bound on the reduction. Although the upper bound for both reductions is the same, the direct truncation method tends to have smaller errors at high frequencies and larger errors at low frequencies, while the singular perturbation approximation method will display the opposite character. It also shows that a certain bilinear mapping not only preserves the balanced structure between a continuous-time system and an associated discrete-time system, but also preserves the slow singular perturbation approximation structure. Hence the continuous-time results on the singular perturbation approximation of balanced systems are easily extended to the discrete-time case. Examples are used to show the compatibility and the differences in the two reduction techniques for a balanced system.

1. Introduction

Recent control literature shows that an important role is played by the balanced-realization truncation (weighted and unweighted) order-reduction techniques in model and controller reduction procedures. For the open-loop model-reduction problem, a technique of truncation of a balanced realization due to Moore (1982) offers some advantages over some conventional model reduction techniques (for an overview of many of these reduction schemes see Jamshidi 1982). It is well known now that direct balanced-realization truncation retains stability in the reduced-order model (under a very weak condition) (Pernebo and Silverman 1982). This reduced-order system is also balanced and minimal (in the continous-time case) (Pernebo and Silverman 1982). Further, there is an easily calculable frequency error bound available (Glover 1984, Enns 1984 a, b). There are, however, some drawbacks to this reduction technique. For instance, the reduction technique has a mismatch of the DC gains of the high-order model and the reduced-order model. As pointed out by Sreeram and Agathoklis (1989), the technique tends to give a good approximation of the impulse response but has a large steady-state error for the step response. In fact, if the order reduction is 1, the frequency domain error between the full-order model and the reduced-order model will achieve its upper bound (twice the smallest Hankel singular value of the full-order model) at DC Enns (1984 a). In contrast, at very high frequencies, the error of the reduction tends to zero. This is to say that the reduction error of the directly balanced trunation method of Moore tends to zero at very high frequencies but is in general non-zero at very low frequencies. The latter property at least is quite contradictory to what one would often seek in model and controller

Received 8 March 1989.

† Department of Systems Engineering, Research School of Physical Sciences, The Australian National University, G.P.O. Box 4, ACT 2601, Australia.

reduction. In most situations one would expect a nice reduction procedure to retain as much as possible the low- and medium-frequency properties of the high-order transfer function. Certainly, the introduction of frequency weighting into the balanced-truncation method will overcome part of the problem. However, the price for this is loss of the easily calculable frequency error bound (see Enns 1984 a, b). Also, one still faces the problem of mismatch of the DC gains in the reduction. This motivated us to look for alternatives. Inspired by some interesting work of Fernando and Nicholson (1982 a, b, 1983 a, b, c), we shall show that the singular perturbation approximation of a balanced realization is a good candidate for model-reduction problems on the grounds of the provable properties of the frequency-domain error.

In the next section we present a brief review of the properties of directly truncating a balanced system and the singular perturbation order-reduction technique. Then in § 3 we shall show that by applying the singular perturbation approximation technique to a balanced realization (instead of truncation of the realization), one can obtain a reduced-order system that matches the DC gain and still has an easily calculable frequency error bound (as for the direct truncation). Section 4 will extend the results to discrete-time systems. Some examples will be used in § 5 to illustrate the above model-reduction methods. Section 6 concludes the paper.

2. Preliminaries

Let us consider an nth-order, linear, time-invariant and asymptotically stable system $G(s)$ with a minimal realization

$$G(s) = C(sI_n - A)^{-1}B + E \tag{2.1}$$

The controllability gramian and the observability gramian of the system are defined as follows:

$$P = \int_0^\infty \exp(At)\, BB^T \exp(A^T t)\, dt \qquad (2.2\,a)$$

$$Q = \int_0^\infty \exp(A^T t)\, C^T C \exp(At)\, dt \qquad (2.2\,b)$$

It is well known that these gramians satisfy the following Lyapunov equations:

$$AP + PA^T + BB^T = 0 \qquad (2.3\,a)$$

$$A^T Q + QA + C^T C = 0 \qquad (2.3\,b)$$

A realization (A, B, C, E) of the system $G(s)$ is said to be 'internally balanced' if

$$P = Q = \Sigma = \text{diag}\,\{\sigma_1, \sigma_2, \ldots, \sigma_n\} \tag{2.4}$$

where $\sigma_i \geqslant \sigma_{i+1}$, $i = 1, 2, \ldots, n-1$.

Now partition the balanced system (A, B, C, E) and the gramian Σ conformally as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \quad C_2], \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \tag{2.5}$$

where A_{11} and Σ_1 are $r \times r$ $(r < n)$ matrices.

Moore (1982) suggested that the subsystem (A_{11}, B_1, C_1) should be a good approximation of the balanced system (A, B, C) if $\sigma_r \gg \sigma_{r+1}$. We call this rth-order system a 'direct-truncation' (DT) approximation of the balanced system. Several well-known results concerning the approximation are available.

Lemma 2.1 (Pernebo and Silverman 1982)

The subsystems (A_{ii}, B_i, C_i) $(i = 1, 2)$ are internally balanced with gramian Σ_i $(i = 1, 2)$.

Lemma 2.2 (Pernebo and Silverman 1982)

The matrices A_{ii} $(i = 1, 2)$ are asymptotically stable, i.e. $\operatorname{Re}(\lambda_k\{A_{ii}\}) < 0$ for all k $(i = 1, 2)$ if Σ_1 and Σ_2 have no diagonal entries in common. Further, the subsystem (A_{11}, B_1, C_1) is controllable and observable.

Lemma 2.3 (Glover 1984, Enns 1984 a, b)

There is an upper bound on the approximation error

$$\begin{aligned}\|C(j\omega I - A)^{-1}B - C_1(j\omega I - A_{11})^{-1}B_1\|_\infty &\leqslant 2(\sigma_{r+1} + \dots + \sigma_n)\\ &= 2\operatorname{tr}(\Sigma_2)\end{aligned} \tag{2.6}$$

where the infinity norm is defined as

$$\|X(j\omega)\|_\infty \triangleq \sup_{\omega\in\mathbb{R}} \bar{\sigma}\{X(j\omega)\} = \sup_{\omega\in\mathbb{R}} \{X(j\omega)X^{\mathrm{T}}(-j\omega)\}$$

Now let us focus on the application of the singular perturbation technique to the order reduction of a linear time-invariant system. The interested reader is referred to Kokotovic *et al.* (1976) and Saksena *et al.* (1984) for overviews of the technique.

Consider a linear time-invariant system

$$\begin{aligned}\begin{bmatrix}\dot{x}\\ \dot{z}\end{bmatrix} &= \begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}x\\ z\end{bmatrix} + \begin{bmatrix}B_1\\ B_2\end{bmatrix}u, \quad \begin{bmatrix}x(0)\\ z(0)\end{bmatrix} = \begin{bmatrix}x^0\\ z^0\end{bmatrix}\\ y &= [C_1 \quad C_2]\begin{bmatrix}x\\ z\end{bmatrix} + Eu, \quad x\in\mathbb{R}^r, \quad z\in\mathbb{R}^{n-r} \quad (r<n)\end{aligned} \tag{2.7}$$

where $u(t)\in\mathbb{R}^l$ is a control vector and $y(t)\in\mathbb{R}^m$ is an output vector. When z is a fast and stable variable (see Kokotovic *et al.* 1976, Saksena *et al.* 1984), (2.7) can be approximated by setting $\dot{z}$ equal to zero. More precisely, if

$$\min_i |\lambda_i\{A_{22}\}| > \max_j |\lambda_j\{A_{11} - A_{12}A_{22}^{-1}A_{21}\}| \tag{2.8}$$

and

$$\operatorname{Re}(\lambda_i\{A_{22}\}) < 0 \quad \forall i \tag{2.9}$$

we replace (2.7) by

$$\left.\begin{aligned}\dot{x}_s &= A_{11}x_s + A_{12}z_s + B_1u, \quad x_s(0) = x^0\\ 0 &= A_{21}x_s + A_{22}z_s + B_2u\\ y_s &= C_1x_s + C_2z_s + Eu\end{aligned}\right\} \tag{2.10}$$

whence

$$\left.\begin{aligned}\dot{x}_s &= \bar{A}x_s + \bar{B}u_s, \quad x_s(0) = x^0\\ y_s &= \bar{C}x_s + \bar{E}u_s\end{aligned}\right\} \tag{2.11}$$

where

$$\left.\begin{aligned}\bar{A} &= A_{11} - A_{12}A_{22}^{-1}A_{21}\\ \bar{B} &= B_1 - A_{12}A_{22}^{-1}B_2\\ \bar{C} &= C_1 - C_2A_{22}^{-1}A_{21}\\ \bar{E} &= E - C_2A_{22}^{-1}B_2\end{aligned}\right\} \tag{2.12}$$

It is generally accepted that (2.11) is a good approximation to (2.7) except near $t = 0$ (when a very rapid change in $z(\cdot)$ occurs in a so-called boundary layer). In this paper we shall examine the approximation where (2.7) is balanced, but where (2.8) does *not* necessarily hold; we shall demonstrate that the frequency-domain error between the transfer function matrices of (2.7) and (2.11) has attractive properties.

It is very interesting to note that singular perturbation approximations can be developed in the frequency domain (Fernando and Nicholson 1982 b). For a system given as in (2.1) (not necessarily balanced), using the partition of the system of (2.5), the transfer function can be written in the form

$$G(s) = [C_1 \quad C_2]\begin{bmatrix} sI_r - A_{11} & -A_{12}\\ -A_{21} & sI_{n-r} - A_{22}\end{bmatrix}^{-1}\begin{bmatrix}B_1\\ B_2\end{bmatrix} + E \tag{2.13}$$

We can now decompose additively the transfer function $G(s)$ as

$$G(s) = G_1(s) + G_2(s) \tag{2.14}$$

where

$$\left.\begin{aligned}G_1(s) &= \bar{C}(s)(sI_r - \bar{A}(s))^{-1}\bar{B}(s) + E\\ \bar{A}(s) &= A_{11} + A_{12}(sI_{n-r} - A_{22})^{-1}A_{21}\\ \bar{B}(s) &= B_1 + A_{12}(sI_{n-r} - A_{22})^{-1}B_2\\ \bar{C}(s) &= C_1 + C_2(sI_{n-r} - A_{22})^{-1}A_{21}\\ G_2(s) &= C_2(sI_{n-r} - A_{22})^{-1}B_2\end{aligned}\right\} \tag{2.15}$$

We now have the following result.

Lemma 2.4 (Fernando and Nicholson 1982 b)

Consider a continuous-time, linear, time-invariant and stable system (A, B, C, E) defined as in (2.1). Partition the system conformally as in (2.5) and then additively decompose the transfer function $G(s) = G_1(s) + G_2(s)$, where $G_1(s)$ and $G_2(s)$ are defined as in (2.15). Now if the subsystem $G_2(s)$ is stable and 'fast' (i.e. its states have very fast transient dynamics and decay rapidly to certain 'steady values') in the neighbourhood of a given frequency $s = \sigma_0$ then, by ignoring the dynamics of this fast subsystem, the system with transfer function $G(s)$ can be approximated by the reduced-order system with transfer function

$$\bar{G}(s) = \bar{C}(\sigma_0)[sI_r - \bar{A}(\sigma_0)]^{-1}\bar{B}(\sigma_0) + \bar{E}(\sigma_0) \tag{2.16}$$

where

$$\bar{E}(\sigma_0) = E + C_2(\sigma_0 I_{n-r} - A_{22})^{-1} B_2$$

and $\bar{A}(s)$, $\bar{B}(s)$ and $\bar{C}(s)$ are defined as in (2.15).

This model order-reduction method is termed the *generalized (slow) singular perturbation approximation* at frequency $s = \sigma_0$ (Fernando and Nicholson 1982 b).

The meaning of the modification of the feedthrough term in (2.16) is very clear. It will ensure that the magnitude of the reduced-order system (2.16) matches the magnitude of the high-order system (2.1) at the frequency $s = \sigma_0$, i.e. the reduction error becomes zero at $s = \sigma_0$. If $\sigma_0 = 0$ then we can match the DC gains in the reduction.

It should be pointed out that, theoretically speaking, the frequency σ_0 can be any complex number. However, this would cause the transfer function $\bar{G}(s)$ of the reduced-order system to become a complex-coefficient transfer function. This is not attractive at all. Hence one should concentrate on a real frequency σ_0 in the generalized singular perturbation approximation.

We now consider two extreme cases of the generalized singular perturbation reduction method.

(i) $\sigma_0 = 0$: we then have the reduced-order model as

$$\bar{G}(s) = \bar{C}(0)[sI_r - \bar{A}(0)]^{-1} \bar{B}(0) + \bar{E}(0)$$

where

$$\left.\begin{aligned} \bar{A}(0) &= A_{11} - A_{12}A_{22}^{-1}A_{21} \\ \bar{B}(0) &= B_1 - A_{12}A_{22}^{-1}B_2 \\ \bar{C}(0) &= C_1 - C_2A_{22}^{-1}A_{21} \\ \bar{E}(0) &= E - C_2A_{22}^{-1}B_2 \end{aligned}\right\} \tag{2.17}$$

Hence we obtain the familiar singular perturbation approximation as in (2.12).

(ii) $\sigma_0 \to \infty$: we obtain

$$\bar{A}(\sigma_0) \to A_{11}, \quad \bar{B}(\sigma_0) \to B_1, \quad \bar{C}(\sigma_0) \to C_1, \quad \bar{E}(\sigma_0) \to E \quad \text{as } \sigma_0 \to \infty$$

Hence this case corresponds to the direct-truncation reduction method. (However, up to this point in this paper we have not assumed that we are working with a balanced realization.)

We now consider the singular perturbation approximation method in the discrete-time case. As in the continuous-time case, for the discrete-time system defined by

$$x(k+1) = \Phi x(k) + \Gamma u(k) \qquad y(k) = Hx(k) + E_d u(k) \tag{2.18}$$

with the system partitioned conformally as

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix}, \quad H = [H_1 \quad H_2] \tag{2.19}$$

the transfer function $G(z)$ can be decomposed additively as

$$G(z) = H(zI_n - \Phi)^{-1}\Gamma + E_d = G_1(z) + G_2(z) \tag{2.20}$$

where

$$\left.\begin{aligned} G_1(z) &= \bar{H}(z)[zI_r - \bar{\Phi}(z)]^{-1}\bar{\Gamma}(z) + E_d \\ \bar{\Phi}(z) &= \Phi_{11} + \Phi_{12}(zI_{n-r} - \Phi_{22})^{-1}\Phi_{21} \\ \bar{\Gamma}(z) &= \Gamma_1 + \Phi_{12}(zI_{n-r} - \Phi_{22})^{-1}\Gamma_2 \\ \bar{H}(z) &= H_1 + H_2(zI_{n-r} - \Phi_{22})^{-1}\Phi_{21} \\ G_2(z) &= H_2(zI_{n-r} - \Phi_{22})^{-1}\Gamma_2 \end{aligned}\right\} \tag{2.21}$$

We now have the following.

Lemma 2.5 (Fernando and Nicholson 1983 b)

Consider a discrete-time, linear, time-invariant and stable system (Φ, Γ, H, E_d) defined as in (2.18) with transfer-function matrix $G(z)$ as in (2.20). Partition the system conformally as in (2.19) and additively decompose $G(z) = G_1(z) + G_2(z)$, where $G_1(z)$ and $G_2(z)$ are defined as in (2.21). Now if the subsystem $G_2(z)$ is stable and 'fast' near the frequency $z = z_0$ we then obtain by neglecting the dynamics of this fast subsystem a low-order approximation of $G(z)$ with transfer function

$$\bar{G}(z) = \bar{H}(z_0)[zI_r - \bar{\Phi}(z_0)]^{-1}\bar{\Gamma}(z_0) + \bar{E}_d(z_0) \tag{2.22}$$

where

$$\bar{E}_d(z_0) = E_d + H_2(z_0 I_{n-r} - \Phi_{22})^{-1}\Gamma_2 \tag{2.23}$$

and $\bar{\Phi}(z)$ and $\bar{H}(z)$ are defined as in (2.21) and $0 < |z_0| \leq 1$.

Two extreme cases can be considered.

(i) $z_0 = 1$: we obtain the slow singular perturbation approximation defined as follows:

$$\left.\begin{aligned} \bar{\Phi}(1) &= \Phi_{11} + \Phi_{12}(I_{n-r} - \Phi_{22})^{-1}\Phi_{21} \\ \bar{\Gamma}(1) &= \Gamma_1 + \Phi_{12}(I_{n-r} - \Phi_{22})^{-1}\Gamma_2 \\ \bar{H}(1) &= H_1 + H_2(I_{n-r} - \Phi_{22})^{-1}\Phi_{21} \\ \bar{E}_d(1) &= E_d + H_2(I_{n-r} - \Phi_{22})^{-1}\Gamma_2 \end{aligned}\right\} \tag{2.24}$$

(ii) If $z_0 \to 0$ and Φ_{22}^{-1} exists then we have

$$\left.\begin{aligned} \bar{\Phi}(0) &= \Phi_{11} - \Phi_{12}\Phi_{22}^{-1}\Phi_{21} \\ \bar{\Gamma}(0) &= \Gamma_1 - \Phi_{12}\Phi_{22}^{-1}\Gamma_2 \\ \bar{H}(0) &= H_1 - H_2\Phi_{22}^{-1}\Phi_{21} \\ \bar{E}_d(0) &= E_d - H_2\Phi_{22}^{-1}\Gamma_2 \end{aligned}\right\} \tag{2.25}$$

We have seen that the singular perturbation approximation method for model reduction can be developed in the time domain as well as the frequency domain. The generalized slow singular perturbation approximation when we take $\sigma_0 = 0$ ($z_0 = 1$ for the discrete-time case) will match the DC gains in the reduction. Now the question arises as to whether, if we apply this reduction technique instead of direct truncation to a balanced system, we will secure the advantages of both reduction methods. In other words, we need to know whether the reduction method can have very good behaviour at low frequencies while it still has a very easily calculable frequency error bound. We shall address these problems in the next two sections.

3. Properties of singular perturbation approximation of balanced systems—continuous-time case

Before we state the properties of the singular perturbation approximation (SPA) of internally balanced systems, we first establish some more properties of balanced systems.

Let us consider a continuous-time, linear, time-invariant and stable system $G(s) = C(sI_n - A)^{-1}B + E$, with (A, B, C, E) being minimal and balanced with gramian $\Sigma = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$, $\sigma_i \geqslant \sigma_{i+1} > 0$, $i = 1, 2, \ldots, n-1$, i.e. we have

$$A\Sigma + \Sigma A^{\mathrm{T}} + BB^{\mathrm{T}} = 0 \tag{3.1 a}$$

$$A^{\mathrm{T}}\Sigma + \Sigma A + C^{\mathrm{T}}C = 0 \tag{3.1 b}$$

Now define (using the partition (2.5))

$$\left.\begin{aligned} \bar{A} &= A_{11} - A_{12}A_{22}^{-1}A_{21} \\ \bar{B} &= B_1 - A_{12}A_{22}^{-1}B_2 \\ \bar{C} &= C_1 - C_2A_{22}^{-1}A_{21} \\ \bar{E} &= E - C_2A_{22}^{-1}B_2 \end{aligned}\right\} \tag{3.2}$$

Then we know from the last section that the reduced-order system $(\bar{A}, \bar{B}, \bar{C}, \bar{E})$ is the slow singular perturbation approximation of the balanced system (A, B, C, E) if A_{22} is stable and its eigenvalues are fast. However, we shall not henceforth require the property that the eigenvalues are fast. This means that the approximation may not enjoy all the time-domain properties usually associated with a singular perturbation approximation. As we shall see, however, it still has useful properties.

Now define the 'reciprocal system' $(\hat{A}, \hat{B}, \hat{C})$ of (A, B, C) by (Fernando and Nicholson 1983 a)

$$\hat{A} \triangleq A^{-1}, \quad \hat{B} \triangleq A^{-1}B, \quad \hat{C} \triangleq CA^{-1} \tag{3.3}$$

assuming that A is non-singular. We have the following.

Lemma 3.1 (Fernando and Nicholson 1983 a)

Let (A, B, C) be the minimal and internally balanced realization with gramian Σ of a linear, time-invariant and stable system. Then the reciprocal system $(\hat{A}, \hat{B}, \hat{C})$ (defined as in (3.3)) is also stable and internally balanced with gramian Σ.

Partition the system $(\hat{A}, \hat{B}, \hat{C})$ and the gramian Σ conformally as

$$\hat{A} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \end{bmatrix}, \quad \hat{C} = [\hat{C}_1 \quad \hat{C}_2], \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \tag{3.4}$$

Then we have the following (by Lemmas 2.1 and 2.2).

Lemma 3.2

Let the hypothesis of Lemma 3.1 hold and let the reciprocal system $(\hat{A}, \hat{B}, \hat{C})$ be partitioned as in (3.4). Then the subsystem $(\hat{A}_{ii}, \hat{B}_i, \hat{C}_i)$ $(i = 1, 2)$ is also internally balanced with gramian Σ_i $(i = 1, 2)$.

Lemma 3.3

Let the hypothesis of Lemma 3.2 hold. Then the subsystem matrix $\hat{A}_{ii}$ $(i = 1, 2)$ is asymptotically stable if Σ_1 and Σ_2 have no common diagonal element. Further, the subsystem $(\hat{A}_{ii}, \hat{B}_i, \hat{C}_i)$ $(i = 1, 2)$ is controllable and observable.

We also have the following frequency error bound by Lemma 2.3.

Lemma 3.4

Let the hypothesis of Lemma 3.2 hold. Then we have

$$\|\hat{C}(j\omega I_n - \hat{A})^{-1}\hat{B} - \hat{C}_1(j\omega I_r - \hat{A}_{11})^{-1}\hat{B}_1\|_\infty \leqslant 2\,\mathrm{tr}\,(\Sigma_2) \tag{3.5}$$

Now, if we apply the singular perturbation approximation technique to the internally balanced system (A, B, C, E) instead of the direct truncation method, what kind of properties can we expect for this reduction? To answer this question, the key is to exploit the 'reciprocal system' of (A, B, C, E).

Notice from (3.3) and (3.4) and the block-matrix inversion lemma (Kailath 1980) that

$$\bar{A} = \hat{A}_{11}^{-1}, \quad \bar{B} = \hat{A}_{11}^{-1}\hat{B}_1, \quad \bar{C} = \hat{C}_1\hat{A}_{11}^{-1} \tag{3.6}$$

i.e. the system $(\bar{A}, \bar{B}, \bar{C})$ (defined as in (3.2)) is the 'reciprocal system' of $(\hat{A}_{11}, \hat{B}_1, \hat{C}_1)$. Hence we immediately obtain some properties of the singular perturbation approximation reduction method.

Lemma 3.5 (Fernando and Nicholson 1983 c)

Let (A, B, C, E) be the minimal and internally balanced realization with the gramian Σ of a continuous-time, linear, time-invariant and asymptotically stable system. Partition the system conformally as in (2.5). Then the slow singular perturbation approximation $(\bar{A}, \bar{B}, \bar{C}, \bar{E})$ defined in (3.2) of the system (A, B, C, E) is also internally balanced with gramian Σ_1 (Σ_1 is defined as in (2.5)).

Theorem 3.1

Let the hypothesis of Lemma 3.5 hold true. Then the singular perturbation reduced-order system $(\bar{A}, \bar{B}, \bar{C}, \bar{E})$ defined as in (3.2) of the system (A, B, C, E) is also asymptotically stable and controllable and observable if Σ_1 and Σ_2 (as defined in (2.5)) have no common diagonal element.

The proof of this result is straightforward by noting Lemma 3.3 and the relation (3.6). Next we have one of the main results of this paper.

Theorem 3.2

Let the hypothesis of Lemma 3.5 hold true. Then there is a frequency error bound available for the singular perturbation approximation $(\bar{A}, \bar{B}, \bar{C}, \bar{E})$ defined in (3.2) of the stable and internally balanced system (A, B, C, E):

$$\|C(j\omega I_n - A)^{-1}B + E - \bar{C}(j\omega I_r - \bar{A})^{-1}\bar{B} - \bar{E}\|_\infty \leqslant 2\,\mathrm{tr}\,(\Sigma_2) \tag{3.7}$$

where Σ_2 is defined as in (3.4).

Proof

Define a 1–1 mapping on $\mathbb{C}$: $s' = 1/s$; then

$$\begin{aligned}
C(sI_n - A)^{-1}B = C\left(\frac{1}{s'}I_n - A\right)^{-1}B &= CA^{-1}s'(A^{-1} - s'I_n)^{-1}B \\
&= -CA^{-1}[(s'I_n - A^{-1}) + A^{-1}](s'I_n - A^{-1})^{-1}B \\
&= -CA^{-1}B - CA^{-1}(s'I_n - A^{-1})^{-1}A^{-1}B \\
&= -CA^{-1}B - \hat{C}(s'I_n - \hat{A})^{-1}\hat{B}
\end{aligned} \tag{3.8}$$

where $(\hat{A}, \hat{B}, \hat{C})$ is defined as in (3.3). Now, from Lemma 3.1, we know that $(\hat{A}, \hat{B}, \hat{C})$ is balanced with gramian Σ. Perform a direct-truncation approximation of $(\hat{A}, \hat{B}, \hat{C})$ as $(\hat{A}_{11}, \hat{B}_1, \hat{C}_1)$ defined in (3.4). Following the above derivation (3.8) and using a mapping $s' \mapsto s$, we have

$$\hat{C}_1(s'I_r - \hat{A}_{11})^{-1}\hat{B}_1 = -\hat{C}_1\hat{A}_{11}^{-1}\hat{B}_1 - \hat{C}_1\hat{A}_{11}^{-1}(sI_r - \hat{A}_{11}^{-1})^{-1}\hat{A}_{11}^{-1}\hat{B}_1$$

Using the relation (3.6), we obtain

$$\hat{C}_1(s'I_r - \hat{A}_{11})^{-1}\hat{B}_1 = -\bar{C}\bar{A}^{-1}\bar{B} - \bar{C}(sI_r - \bar{A})^{-1}\bar{B} \tag{3.9}$$

Further, it is easy to verify that

$$CA^{-1}B = \bar{C}\bar{A}^{-1}\bar{B} + C_2A_{22}^{-1}B_2 \tag{3.10}$$

On the $j\omega$-axis, the mapping $s \mapsto s'$ is equivalent to the mapping $j\omega \mapsto j\omega'$. Therefore we have

$$\begin{aligned}
&\|C(j\omega I_n - A)^{-1}B + E - \bar{C}(j\omega I_r - \bar{A})^{-1}\bar{B} - \bar{E}\|_\infty \\
&\quad = \|-CA^{-1}B - \hat{C}(j\omega' I_n - \hat{A})^{-1}\hat{B} + \hat{C}_1(j\omega' I_r - \hat{A}_{11})^{-1}\hat{B}_1 \\
&\qquad + \bar{C}\bar{A}^{-1}\bar{B} + C_2A_{22}^{-1}B_2\|_\infty && \text{(by (3.8) and (3.9))} \\
&\quad = \|\hat{C}(j\omega' I_n - \hat{A})^{-1}\hat{B} - \hat{C}_1(j\omega' I_r - \hat{A}_{11})^{-1}\hat{B}_1\|_\infty && \text{(by (3.10))} \\
&\quad \leqslant 2\,\mathrm{tr}\,(\Sigma_2) && \text{(by Lemma 3.4)}
\end{aligned}$$

□

The significance of this theorem is very clear. To use the singular perturbation approximation technique (3.2) to reduce the order of the internally balanced system (A, B, C, E), it is *not* necessary that the time-scale separation properties (2.9) should hold true. In fact, as long as the balanced system (A, B, C, E) has a weakly controllable and weakly observable subsystem, i.e. the sum of the singular values corresponding to this weak subsystem, $\mathrm{tr}\,(\Sigma_2)$, is small, the result of Theorem 3.2 guarantees a good reduction from the singular perturbation approximation technique in the sense that the reduction error will be over bounded by a small quantity $2\,\mathrm{tr}\,(\Sigma_2)$.

It is interesting to note that if the high-order balanced system is strictly proper, i.e.

$E = 0$, then in (3.2), $\bar{E} = -C_2 A_{22}^{-1} B_2$. Hence if the system to be reduced is strictly proper, its singular perturbation approximation will usually be a proper but not strictly proper reduced-order model. If one insists on a strictly proper singular perturbation approximation as in Fernando and Nicholson (1982 a, b, 1983 c) and Santiago and Jamshidi (1986), one must expect a larger frequency error bound.

Corollary 3.1

Assume the same hypotheses as Theorem 3.2; then

$$\| C(j\omega I_n - A)^{-1} B - \bar{C}(j\omega I_r - \bar{A})^{-1} \bar{B} \|_\infty \leq 4 \operatorname{tr}(\Sigma_2) \tag{3.11}$$

Proof

In Theorem 3.2, (3.7), let $\omega \to \infty$; then

$$\| C(j\omega I_n - A)^{-1} B + E - \bar{C}(j\omega I_r - \bar{A})^{-1} \bar{B} - \bar{E} \|_\infty \to \| C_2 A_{22}^{-1} B_2 \|_\infty \leq 2 \operatorname{tr}(\Sigma_2) \tag{3.12}$$

Hence

$$\begin{aligned}
&\| C(j\omega I_n - A)^{-1} B - \bar{C}(j\omega I_r - \bar{A})^{-1} \bar{B} \|_\infty \\
&\quad = \| C(j\omega I_n - A)^{-1} B + E - \bar{C}(j\omega I_r - \bar{A})^{-1} \bar{B} - (E - C_2 A_{22}^{-1} B_2) - C_2 A_{22}^{-1} B_2 \|_\infty \\
&\quad \leq \| C(j\omega I_n - A)^{-1} B + E - \bar{C}(j\omega I_r - \bar{A})^{-1} \bar{B} - \bar{E} \|_\infty + \| C_2 A_{22}^{-1} B_2 \|_\infty \\
&\quad \leq 4 \operatorname{tr}(\Sigma_2)
\end{aligned}$$

□

We have now seen that if we use direct truncation on a stable and balanced system to obtain a low-order approximation, we can have a very good reduction error (near zero) at very high frequencies, but not such a good one at low frequencies. If we use the singular perturbation approximation technique on a stable and balanced system to find the low-order model, we have the reverse conclusion. Hence we can imagine that a mixed use of these two reduction techniques on a stable balanced system will 'average' the behaviour of both methods at high and low frequencies. We have the following.

Theorem 3.3

Assume that the hypothesis of Lemma 3.5 hold true. Then the techniques of direct truncation (DT) and singular perturbation approximation (SPA) can be used in a mixed way to reduce the order of a linear, time-invariant, minimal, stable and internally balanced system. That is to say, the reduction is done by several sequential steps, and, for each step, either the direct truncation or the singular perturbation approximation method can be employed to reduce the order. Further, the final reduced-order model is also minimal, internally balanced and stable. The frequency error bound of the reduction will remain the same as if the reduction has been done by either method in one step.

Proof

This is just a simple consequence of Lemmas 2.1–2.3 in conjunction with Theorems 3.1 and 3.2. For each technique the resulting low-order system is still balanced with the gramian Σ_1, and the frequency error bound of the reduction only depends on Σ_2. So, for each sequential reduction step, we can regard the system obtained from the last reduction step as a freshly balanced system with gramian Σ_1. Hence the conclusion follows. □

The question now arises as to whether the above results can be extended to generalized singular perturbation approximation of a balanced system. For the generalized singular perturbation approximation of (2.16), if the frequency σ_0 is non-zero and a finite real number, unfortunately we are not able to extend the above results. However, if $\sigma_0 = j\xi$, with ξ a real number (note that this will produce a complex-coefficient transfer function), we obtain the following.

Theorem 3.4

Assume that the generalized singular perturbation approximation reduction technique (2.16) has been employed to reduce the order of a stable and balanced system $G(s) = C(sI_n - A)^{-1}B + E$ defined as in Lemma 3.5, with the frequency $\sigma_0 = j\xi$ in (2.16), where ξ is an arbitrary real number. Then we have

$$\|G(j\omega) - \bar{C}(\sigma_0)[j\omega I_r - \bar{A}(\sigma_0)]^{-1}\bar{B}(\sigma_0) - \bar{E}(\sigma_0)\|_\infty \leqslant 2\,\mathrm{tr}\,(\Sigma_2) \tag{3.13}$$

where Σ_2 is defined as in (2.5).

The proof of this theorem is quite lengthy, and is relegated to the Appendix.

There are two particular points to note in this theorem.

(i) Perhaps surprisingly, the frequency $\sigma_0 = j\xi$ does not appear in the error bound.

(ii) As a particular instance of this point, when we take two special cases, i.e. $\sigma_0 = 0$ or $\sigma_0 = j\infty$, we have the results of Theorem 3.2 and Lemma 2.3. Note that the proof is quite different from those in Glover (1984) and Enns (1984 a, b) and from the proof of Theorem 3.2.

In the next section we shall consider the discrete-time case.

4. Properties of singular perturbation approximation of balanced systems—discrete-time case

In the last section we have shown that the singular perturbation approximation of a balanced system is fully compatible with Moore's direct-truncation method in the continuous-time case. We now want to examine whether this is still true for discrete-time systems. To do this, we first establish some properties of the bilinear mapping between the complex s-plane and z-plane.

We start with a continuous-time, linear, time-invariant, minimal and stable system with the transfer function $C(sI_n - A)^{-1}B + E$. Assume that the discrete-time, linear, time-invariant, minimal and stable system $H(zI_n - \Phi)^{-1}\Gamma + E_d$ is obtained from the continuous-time system via the following bilinear mapping:

$$s = \frac{z-1}{z+1} \quad \left(\text{inverse: } z = \frac{1+s}{1-s}\right) \tag{4.1}$$

which maps the left-half complex s-plane, Re $(s) \leqslant 0$, onto the unit circle $|z| \leqslant 1$ in the z-plane. We then have the relations

$$\left.\begin{aligned} A &= (\Phi + I_n)^{-1}(\Phi - I_n) \\ B &= \sqrt{2}\,(\Phi + I_n)^{-1}\Gamma \\ C &= \sqrt{2}\,H(\Phi + I_n)^{-1} \\ E &= E_{\mathrm{d}} - H(\Phi + I_n)^{-1}\Gamma \end{aligned}\right\} \tag{4.2}$$

or, in another form,

$$C(sI_n - A)^{-1}B + E \underset{z=(1+s)/(1-s)}{\overset{s=(z-1)/(z+1)}{\leftrightarrows}} H(zI_n - \Phi)^{-1}\Gamma + E_{\mathrm{d}} \tag{4.3}$$

We then obtain the following.

Lemma 4.1 (Al-Saggaf 1986, Al-Saggaf and Franklin 1986)

The bilinear mapping (4.1) preserves the internally balanced property of the linear time-invariant system, in the sense that if the stable continuous-time system (A, B, C, E) is internally balanced with gramian Σ then the discrete-time system $(\Phi, \Gamma, H, E_{\mathrm{d}})$ obtained by the above bilinear mapping (4.1) and (4.2) is also internally balanced with the same gramian Σ, and vice versa.

Now partition the balanced system $(\Phi, \Gamma, H, E_{\mathrm{d}})$ conformally as in (2.19). Define

$$\left.\begin{aligned} \bar{\Phi} &= \Phi_{11} + \Phi_{12}(I_{n-r} - \Phi_{22})^{-1}\Phi_{21} \\ \bar{\Gamma} &= \Gamma_1 + \Phi_{12}(I_{n-r} - \Phi_{22})^{-1}\Gamma_2 \\ \bar{H} &= H_1 + H_2(I_{n-r} - \Phi_{22})^{-1}\Phi_{21} \\ \bar{E}_{\mathrm{d}} &= E_{\mathrm{d}} + H_2(I_{n-r} - \Phi_{22})^{-1}\Gamma_2 \end{aligned}\right\} \tag{4.4}$$

Then, from § 2, we know that the low-order system $(\bar{\Phi}, \bar{\Gamma}, \bar{H}, \bar{E}_{\mathrm{d}})$ is the slow singular perturbation approximation of the balanced system $(\Phi, \Gamma, H, E_{\mathrm{d}})$ if Φ_{22} is fast and stable. Once more, we can afford to drop the assumption about fast eigenvalues, and simply require stability.

Further, we have the following as a new result.

Theorem 4.1

The bilinear mapping (4.1), (4.2) preserves the slow singular perturbation approximation, in the sense that if the linear, time-invariant and stable continuous-continuous-time system (A, B, C, E) are linked by the bilinear mapping (4.1), (4.2) then their slow singular perturbation approximations $(\bar{A}, \bar{B}, \bar{C}, \bar{E})$ defined as in (3.2), and $(\bar{\Phi}, \bar{\Gamma}, \bar{H}, \bar{E}_{\mathrm{d}})$ defined as in (4.4) are also linked by the same bilinear mapping, i.e.

$$\bar{C}(sI_r - \bar{A})^{-1}\bar{B} + \bar{E} \underset{z=(1+s)/(1-s)}{\overset{s=(s-1)(s+1)}{\leftrightarrows}} \bar{H}(zI_r - \bar{\Phi})^{-1}\bar{\Gamma} + \bar{E}_{\mathrm{d}} \tag{4.5}$$

and in particular we have the relations

$$\left.\begin{aligned}\bar{A}&=(\bar{\Phi}+I_r)^{-1}(\bar{\Phi}-I_r)\\ \bar{B}&=\sqrt{2}\,(\bar{\Phi}+I_r)^{-1}\bar{\Gamma}\\ \bar{C}&=\sqrt{2}\,\bar{H}(\bar{\Phi}+I_r)^{-1}\\ \bar{E}&=\bar{E}_{\mathrm{d}}-\bar{H}(\bar{\Phi}+I_r)^{-1}\bar{\Gamma}\end{aligned}\right\}\tag{4.6}$$

Proof

Using the block matrix inversion lemma (Kailath 1980), it is easy to see that

$$\bar{H}(zI_r-\bar{\Phi})^{-1}\bar{\Gamma}+\bar{E}_{\mathrm{d}}=H\left(\begin{bmatrix}zI_r & \\ & I_{n-r}\end{bmatrix}-\Phi\right)^{-1}\Gamma+E_{\mathrm{d}}\tag{4.7}$$

$$\bar{C}(sI_r-\bar{A})^{-1}\bar{B}+\bar{E}=C\left(\begin{bmatrix}sI_r & \\ & 0\end{bmatrix}-A\right)^{-1}B+E\tag{4.8}$$

So what we have to do is to show that if (4.2) holds then the right-hand side of (4.7) is equal to the right-hand side of (4.8) when $s=(z-1)/(z+1)$. Now

$$\begin{aligned}C\left(\begin{bmatrix}sI_r & \\ & 0\end{bmatrix}\right)^{-1}B+E&=C\left(\frac{z-1}{z+1}\begin{bmatrix}I_r & \\ & 0\end{bmatrix}-A\right)^{-1}B+E\\
&=\sqrt{2}\,H(\Phi+I_n)^{-1}\left(\frac{z-1}{z+1}\begin{bmatrix}I_r & \\ & 0\end{bmatrix}-(\Phi+I_n)^{-1}(\Phi-I_n)\right)^{-1}\\
&\quad\times(\Phi+I_n)^{-1}\Gamma\sqrt{2}+E_{\mathrm{d}}-H(\Phi+I_n)^{-1}\Gamma\\
&=H\left\{2(\Phi+I_n)^{-1}\left(\frac{z-1}{z+1}\begin{bmatrix}I_r & \\ & 0\end{bmatrix}-[I_n-2(\Phi+I_n)^{-1}]\right)^{-1}\right.\\
&\quad\times(\Phi+I_n)^{-1}\\
&\quad\left.-(\Phi+I_n)^{-1}\right\}\Gamma+E_{\mathrm{d}}\\
&=H\left\{(\Phi+I_n)^{-1}\left(\frac{z-1}{2(z+1)}\begin{bmatrix}I_r & \\ & 0\end{bmatrix}-\frac{I_n}{2}+(\Phi+I_n)^{-1}\right)^{-1}\right.\\
&\quad(\Phi+I_n)^{-1}\\
&\quad\left.-(\Phi+I_n)^{-1}\right\}\Gamma+E_{\mathrm{d}}\end{aligned}$$

Now, using the matrix-inversion lemma (Kailath 1980),

$$(M+NP^{-1}Q)^{-1}=M^{-1}-M^{-1}N(P+QM^{-1}N)^{-1}QM^{-1}$$

(assuming that M and P are invertible), by setting $M=-(\Phi+I_n)$, $N=I_n$,

$$P=\frac{z-1}{2(z+1)}\begin{bmatrix}I_r & 0\\ 0 & 0\end{bmatrix}-\frac{I_n}{2}$$

and $Q = -I_n$, we then obtain

$$C\left(s\begin{bmatrix} I_r & \\ & 0 \end{bmatrix} - A\right)^{-1} B + E = H\left\{-(\Phi + I_n) - \left(\frac{z-1}{2(z+1)}\begin{bmatrix} I_r & \\ & 0 \end{bmatrix} - \frac{I_n}{2}\right)^{-1}\right\}^{-1}$$

$$\times \Gamma + E_{\mathrm{d}}$$

$$= H\left\{-\Phi - I_n - \left\{\left(\frac{1}{2} - \frac{1}{z+1}\right)\begin{bmatrix} I_r & \\ & 0 \end{bmatrix} - \frac{I_n}{2}\right]^{-1}\right\}^{-1}$$

$$\times \Gamma + E_{\mathrm{d}}$$

$$= H\left\{-\Phi - I_n - \begin{bmatrix} -(z+1)I_r & \\ & -2I_{n-r} \end{bmatrix}\right\}^{-1} \Gamma + E_{\mathrm{d}}$$

$$= H\left(\begin{bmatrix} zI_r & \\ & I_{n-r} \end{bmatrix} - \Phi\right)^{-1} \Gamma + E_{\mathrm{d}}$$

The rest of the proof is straightforward. We omit it. □

From the above two results, we immediately have the following.

Theorem 4.2

Assume that the minimal realization of a linear time-invariant discrete-time system, $(\Phi, \Gamma, H, E_{\mathrm{d}})$ is asymptotically stable and internally balanced with the gramian Σ; then its slow singular perturbation approximation, viz the reduced-order system $(\bar{\Phi}, \bar{\Gamma}, \bar{H}, \bar{E}_{\mathrm{d}})$ defined as in (4.4), is also minimal, internally balanced, asymptotically stable if Σ_1 and Σ_2 have no diagonal element in common, and controllable and observable.

Proof

Use the bilinear mapping (4.1) to transform the discrete-time system $(\Phi, \Gamma, H, E_{\mathrm{d}})$ and its slow singular perturbation approximation $(\bar{\Phi}, \bar{\Gamma}, \bar{H}, \bar{E}_{\mathrm{d}})$ into the continuous-time system (A, B, C, E) and its corresponding slow singular perturbation approximation $(\bar{A}, \bar{B}, \bar{C}, \bar{E})$, then apply the existing results (Lemma 3.5 and Theorem 3.3) to these continuous-time systems. The conclusion follows. It is also possible to prove the result by direct algebraic manipulations as shown by Al-Saggaf (1986). □

Certainly, we also have the following.

Theorem 4.3

Assume the same hypothesis as in Theorem 4.2; then there is a frequency error bound available for the singular perturbation reduction of the balanced system

$$\|H(e^{j\theta}I_n - \Phi)^{-1}\Gamma + E_{\mathrm{d}} - \bar{H}(e^{j\theta}I_r - \bar{\Phi})^{-1}\bar{\Gamma} - \bar{E}_{\mathrm{d}}\|_\infty \leqslant 2\,\mathrm{tr}\,(\Sigma_2)$$

The proof of this result is trivial.

Note that the same claim of this theorem has appeared in Al-Saggaf (1986) and Al-Saggaf and Franklin (1986); however, the proof is different.

As a final remark, it should be pointed out that in the discrete-time case directly truncating a balanced system to reduce the system order is *not* fully compatible with the singular perturbation approximation of the balanced system. The problem is that the reduced-order system obtained by direct truncation of a balanced system is no longer balanced. Hence, in contrast with the continuous-time case (Theorem 3.3), we cannot freely mix the two reduction techniques into one reduction procedure. However, we can state the following.

Theorem 4.4

Assume the same hypothesis as in Theorem 4.2. Then the techniques of direct truncation and the slow singular perturbation approximation can be used in the following way to reduce the order of a discrete-time, linear time-invariant, minimal, stable and internally balanced system. Assume that the reduction is done in two sequential steps. For the first step the slow singular perturbation approximation method is used. In the second step the direct truncation technique is employed to reduce the order. Then the final reduced-order model is also minimal and stable (but not internally balanced). The frequency error bound of the reduction will be the same as if the reduction had been done by either technique in one step.

The proof of this theorem is also trivial.

As for continuous-time systems, we hope this mixed reduction procedure will display behaviour between those of the two reduction techniques from which it is composed.

Figure 1 shows the relations among the reduction methods in this paper.

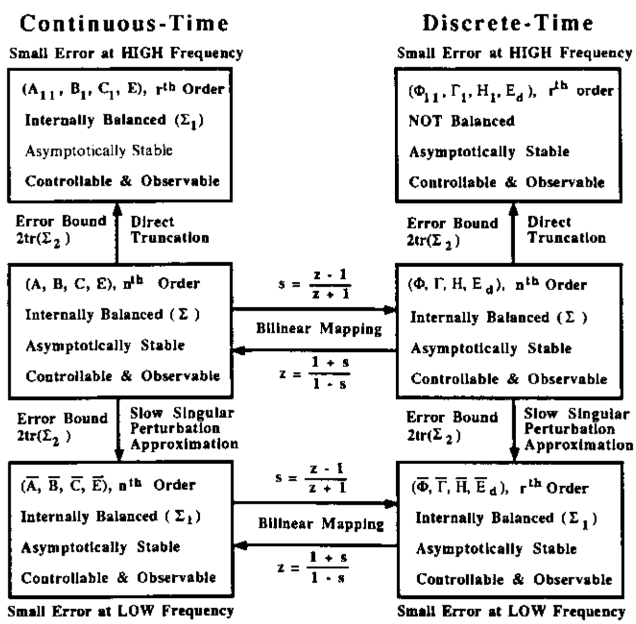

Figure 1. Summary of the model-reduction methods for the balanced system.

In the next section some examples are used to illustrate these reduction methods.

5. Examples

5.1. *Example* 1

To illustrate the two different approaches (and their mixture) in the model order-reduction problem, consider the continuous-time, linear, time-invariant and stable system described by the transfer function

$$G(s) = \frac{s+4}{(s+1)(s+3)(s+5)(s+10)} \tag{5.1}$$

which has appeared in many references (e.g. Moore 1982, Fernando and Nicholson 1983 c, Santiago and Jamshidi 1986). The system can be realized in the following internally balanced format (Moore 1982): $G(s) = C(sI_4 - A)^{-1}B + E$, where

$$A = \begin{bmatrix} -0.43781 & 1.1685 & 0.41426 & 0.05098 \\ -1.1685 & -3.1353 & -2.8352 & -0.32885 \\ 0.41426 & 2.8352 & -12.4753 & -3.2492 \\ -0.05098 & -0.32885 & 3.2492 & -2.9516 \end{bmatrix},$$

$$B = \begin{bmatrix} -0.11814 \\ -0.1307 \\ 0.05634 \\ -0.006875 \end{bmatrix} \tag{5.2}$$

$$C = [-0.11814 \quad 0.1307 \quad 0.05634 \quad -0.006875], \quad E = 0 \tag{5.3}$$

and the balanced gramian matrix Σ is given by

$$\Sigma = \mathrm{diag}\,\{1.5938 \times 10^{-2}, 2.7243 \times 10^{-3}, 1.272 \times 10^{-4}, 8.006 \times 10^{-6}\} \tag{5.4}$$

Now, we want to use Moore's direct truncation, the singular perturbation approximation and a mixture as order-reduction techniques to find a second-order system approximating the above balanced system. We have four different cases.

Case I (DT): Moore's direct-truncation reduction method. Partition the system conformally and then eliminate the weakly controllable and observable subsystem; we obtain the reduced-order system as

$$A_{11} = \begin{bmatrix} -0.43781 & 1.1685 \\ -1.1685 & -3.1353 \end{bmatrix}, \quad B_1 = \begin{bmatrix} -0.11814 \\ -0.1307 \end{bmatrix}$$

$$C_1 = [-0.11814 \quad 0.1307], \quad E = 0$$

$$G_r(s) = C_1(sI_2 - A_{11})^{-1}B_1 + E$$

Note that this reduced-order system remains internally balanced with gramian

$$\Sigma_1 = \mathrm{diag}\,\{1.5938 \times 10^{-2}, 2.7243 \times 10^{-3}\}$$

is controllable and observable, and is also asymptotically stable with matrix A_{11} having the eigenvalues $\lambda\{A_{11}\} = \{-1.1129, -2.4601\}$.

Case II (SPA): the slow singular perturbation approximation method. Using (3.2), we obtain the reduced-order system as

$$\bar{A} = \begin{bmatrix} -0.42491 & 1.2565 \\ -1.2565 & -3.7354 \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} -0.11638 \\ -0.14266 \end{bmatrix}$$

$$\bar{C} = [-0.1168 \quad 0.14266], \quad \bar{E} = 2.384 \times 10^{-4}$$

$$G_r(s) = \bar{C}(sI_2 - \bar{A})^{-1}\bar{B} + \bar{E}$$

Again, the reduced-order system is balanced with gramian Σ_1, controllable and observable, and also asymptotically stable with $\bar{A}$ having the eigenvalues $\lambda\{\bar{A}\} = \{-1.0026, -3.1578\}$.

Case III (DT/SPA): the mixture of Moore's direct-truncation method and the singular perturbation approximation. In this case we first use the direct truncation to obtain a third-order approximation; then we use the singular perturbation approximation to this third-order system to obtain the second-order approximation as

$$A_{m1} = \begin{bmatrix} -0.42405 & 1.2626 \\ -1.2626 & -3.7796 \end{bmatrix}, \quad B_{m1} = \begin{bmatrix} -0.11626 \\ -0.1435 \end{bmatrix}$$

$$C_{m1} = [-0.11626 \quad 0.1435], \quad E_{m1} = 2.5441 \times 10^{-4}$$

$$G_r(s) = C_{m1}(sI_2 - A_{m1})^{-1}B_{m1} + E_{m1}$$

This reduced-order system is balanced with gramian Σ_1, controllable and observable, and also asymptotically stable with A_{m1} having the eigenvalues $\lambda\{A_{m1}\} = \{-0.99696, -3.2067\}$.

Case IV (SPA/DT): the mixture of the singular perturbation approximation and Moore's direct-truncation reduction method. In this case we first employ the singular perturbation approximation method to find a third-order reduction; then we use the direct truncation to obtain the second-order approximation as

$$A_{m2} = \begin{bmatrix} -0.43869 & 1.1628 \\ -1.1628 & -3.0986 \end{bmatrix}, \quad B_{m2} = \begin{bmatrix} -0.11825 \\ -0.12993 \end{bmatrix}$$

$$C_{m2} = [-0.11825 \quad 0.12993], \quad E_{m2} = -1.6012 \times 10^{-5}$$

$$G_r(s) = C_{m2}(sI_2 - A_{m2})^{-1}B_{m2} + E_{m2}$$

Again this reduced-order system is internally balanced with gramian Σ_1, controllable and observable, and also asymptotically stable with A_{m2} having the eigenvalues $\lambda\{A_{m2}\} = \{-1.1231, -2.4142\}$.

We now consider the frequency errors for the above reductions. The Bode plots of the high-order system $G(s)$ and the reduced-order systems of the above four cases are shown in Figs 2 and 3. Figure 4 depicts the frequency error $\bar{\sigma}\{G(j\omega) - G_r(j\omega)\}$, where $G_r(s)$ is obtained from the above four cases. It is clear that the singular perturbation approximation method has very good reduction errors at low frequencies (when $\omega \leq 2 \text{ rad s}^{-1}$). In addition, the frequency error is less than the frequency error of the direct-truncation method until around $\omega = 15 \text{ rad s}^{-1}$. Now looking again at the

Bode plot of the system (Fig. 2), we see that this system is a typical low-pass filter. When $\omega = 15$ the magnitude of the system has decayed about -40 dB from its value in the pass band. Hence $0 \leqslant \omega \leqslant 15$ rad s^{-1} can be regarded as the working frequency range for this system, and within this range, the singular perturbation approximation method gives a better reduction than the direct-truncation method. Figure 5 shows the relative frequency errors of the reductions, $\bar{\sigma}\{G(j\omega) - G_r(j\omega)\}/\bar{\sigma}\{G(j\omega)\}$.

For this system we can consider the peak frequency error of the reductions, $\|G(j\omega) - G_r(j\omega)\|_\infty$; this criterion still favours the singular perturbation approximation method, as shown in Table 1. Note that the theoretical error bound for all the

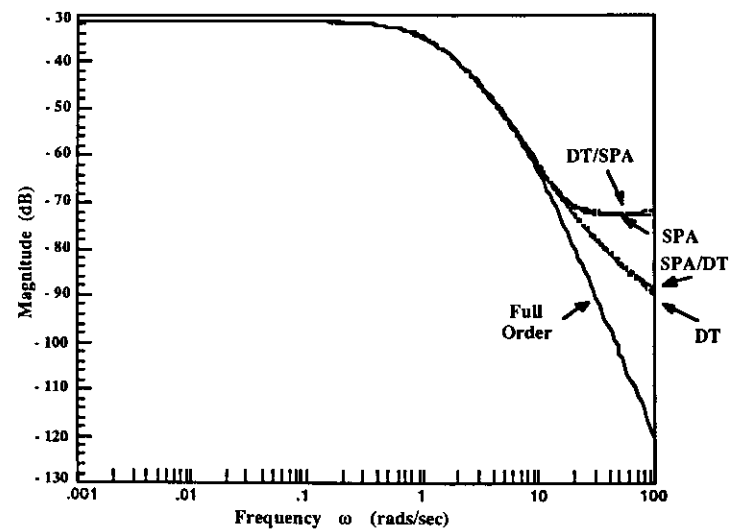

Figure 2. Bode plots (magnitude) of the systems.

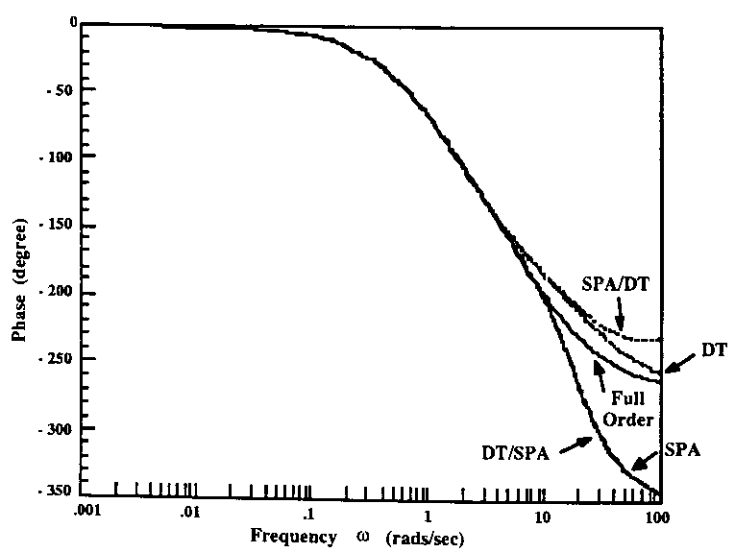

Figure 3. Bode plots (phase) of the systems.

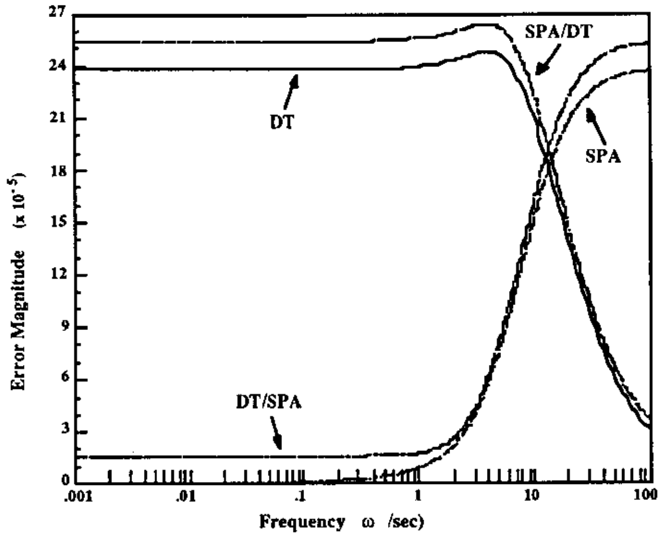

Figure 4. Frequency errors of the reduction (continuous-time).

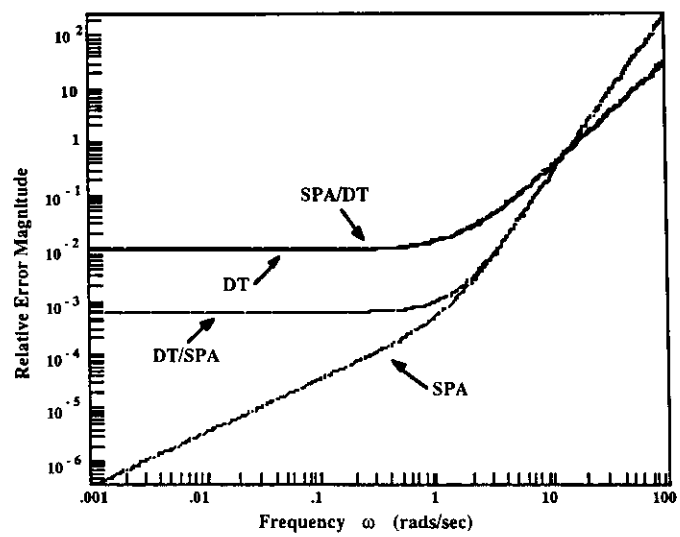

Figure 5. Relative frequency errors of the reduction (continuous-time).

	DT	SPA	DT/SPA	SPA/DT	Theoretical bound
$\|G(j\omega) - G_r(j\omega)\|_\infty$	2·4802	2·3693	2·5284	2·6402	2·7042
Exact error at DC	2·384	0·0	0·16012	2·5441	—

Table 1. Frequency errors of the reduction (Example 1) ($\times 10^{-4}$).

above reductions is $2\,\mathrm{tr}(\Sigma_2) = 2 \times (1{\cdot}272 \times 10^{4} + 8{\cdot}006 \times 10^{-6}) = 2{\cdot}7042 \times 10^{-4}$. It is also interesting to compare the exact errors at DC of the four reduction methods, as also shown in Table 1. In fact, for Cases II, III and IV one can write down directly the exact errors at DC, since the singular perturbation approximation gives no error at DC, and the direct-truncation reduction of order 1 gives the exact error at DC as $2\sigma_4$ and $2\sigma_3$ respectively.

5.2. *Example 2*

We now consider a discrete-time system. For comparison purposes, we simply use the bilinear mapping (4.1) to discretize the system in Example 1, (5.1). Using the relation (4.2) together with (5.2) and (5.3), we have

$$G(z) = H(zI_4 - \Phi)^{-1}\Gamma + E_d$$

where

$$\Phi = \begin{bmatrix} -0{\cdot}1372 & -0{\cdot}30259 & 0{\cdot}02607 & -0{\cdot}01093 \\ 0{\cdot}30259 & 0{\cdot}65545 & 0{\cdot}07482 & -0{\cdot}02894 \\ 0{\cdot}02607 & -0{\cdot}07482 & 0{\cdot}89126 & 0{\cdot}09597 \\ 0{\cdot}01093 & -0{\cdot}02894 & -0{\cdot}09597 & 0{\cdot}57533 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} -0{\cdot}12405 \\ -9{\cdot}6875 \times 10^{-3} \\ 6{\cdot}1354 \times 10^{-5} \\ -3{\cdot}2595 \times 10^{-6} \end{bmatrix}$$

$$H = [-0{\cdot}12405 \quad -9{\cdot}6875 \times 10^{-3} \quad 6{\cdot}1354 \times 10^{-5} \quad -3{\cdot}2595 \times 10^{-6}]$$

$$E_d = 9{\cdot}4697 \times 10^{-3}$$

Now, from the result of Lemma 4.1, we know that this realization is also (discrete-time) internally balanced with the same gramian Σ as shown in (5.4).

As in the continuous-time case, we now use the direct-truncation method, the singular perturbation approximation method and their mixture to obtain a second-order approximation of the above system. We have the following cases.

Case I (DT): direct truncation of the balanced system. We obtain

$$\Phi_{11} = \begin{bmatrix} -0{\cdot}1372 & -0{\cdot}30259 \\ 0{\cdot}30259 & 0{\cdot}65545 \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} -0{\cdot}12405 \\ -9{\cdot}6875 \times 10^{-3} \end{bmatrix}$$

$$H_1 = [-0{\cdot}12405 \quad 9{\cdot}6875 \times 10^{-3}], \quad E_d = 9{\cdot}4697 \times 10^{-3}$$

$$G_r(z) = H_1(zI - \Phi_{11})^{-1}\Gamma_1 + E_d$$

Note now that this reduced-order system is *no longer balanced*. However, it is still controllable and observable and asymptotically stable with Φ_{11} having the eigenvalues $\lambda\{\Phi_{11}\} = \{3{\cdot}163 \times 10^{-3}, 0{\cdot}51509\}$.

Case II (SPA): slow singular perturbation approximation method. Using the reduction method defined in (4.4), we obtain

$$\bar{\Phi} = \begin{bmatrix} -0{\cdot}13124 & -0{\cdot}31964 \\ 0{\cdot}31964 & 0{\cdot}60667 \end{bmatrix}, \quad \bar{\Gamma} = \begin{bmatrix} -0{\cdot}12404 \\ -9{\cdot}6494 \times 10^{-3} \end{bmatrix}$$

$$\bar{H} = [-0{\cdot}12404 \quad 9{\cdot}6494 \times 10^{-3}], \quad \bar{E}_d = 9{\cdot}4697 \times 10^{-3}$$

$$G_r(z) = \bar{H}(zI - \bar{\Phi})^{-1}\bar{\Gamma} + \bar{E}_d$$

This reduced-order system is internally balanced with the gramian $\Sigma_1 = \text{diag}\{1{\cdot}5938 \times 10^{-2}, 2{\cdot}7243 \times 10^{-3}\}$, controllable and observable, and asymptotically stable with Φ having the eigenvalues $\lambda\{\Phi\} = \{5{\cdot}3446 \times 10^{-2}, 0{\cdot}42199\}$.

Case III (SPA/DT): the mixture of the singular perturbation approximation method and the direct truncation approximation (Theorem 4.4). In this case we first reduce the system to a third-order one by the singular perturbation approximation; then we directly truncate this third-order system to obtain the reduced-order system as

$$\Phi_m = \begin{bmatrix} -0{\cdot}13748 & -0{\cdot}30184 \\ 0{\cdot}30184 & 0{\cdot}65742 \end{bmatrix}, \quad \Gamma_m = \begin{bmatrix} -0{\cdot}12405 \\ -9{\cdot}6872 \times 10^{-3} \end{bmatrix}$$

$$H_m = [-0{\cdot}12405 \quad 9{\cdot}6872 \times 10^{-3}], \quad E_m = 9{\cdot}4697 \times 10^{-3}$$

$$G_r(z) = H_m(zI - \Phi_m)^{-1}\Gamma_m + E_{dm}$$

Again this reduced-order system is *not* internally balanced. But it is controllable and observable, and asymptotically stable with Φ_m having the eigenvalues $\lambda\{\Phi_m\} = \{1{\cdot}3957 \times 10^{-3}, 0{\cdot}51855\}$.

Consider the frequency errors for the above model reductions. The frequency errors $\bar{\sigma}\{G(e^{j\theta}) - G_r(e^{j\theta})\}$ are depicted in Fig. 6, where the $G_r(z)$ are obtained from the above three cases. This figure shows clearly that the singular perturbation reduction method has much smaller reduction errors at low frequencies. The relative frequency errors $\bar{\sigma}(G(e^{j\theta}) - G_r(e^{j\theta})\}/\bar{\sigma}\{G(e^{j\theta})\}$ for the above reduction are shown in Fig. 7.

The theoretical upper bound for the L_∞-norm frequency error of the above reductions is again $2\,\text{tr}(\Sigma_2) = 2{\cdot}7024 \times 10^{-4}$. For comparison, we list the L_∞-norm errors actually reached by the above reductions in Table 2. We also can compare the reduction errors at DC (i.e. $z = 1$) for the three reduction methods. However, we cannot write down the exact error at DC for Case III now as in the continuous-time

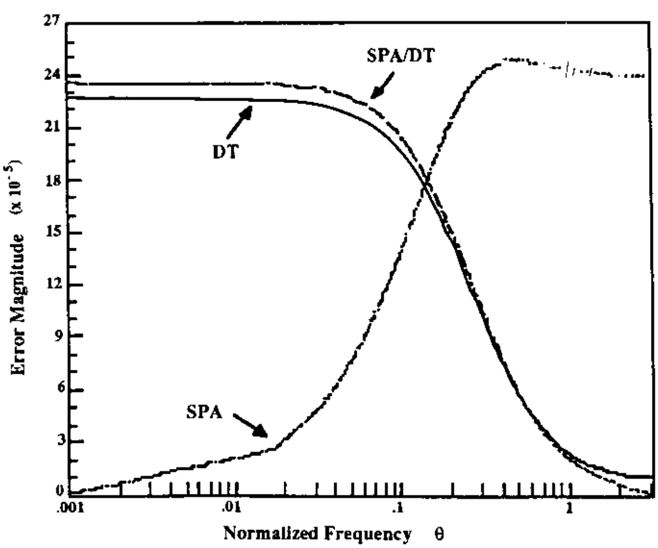

Figure 6. Frequency errors of the reduction (discrete-time).

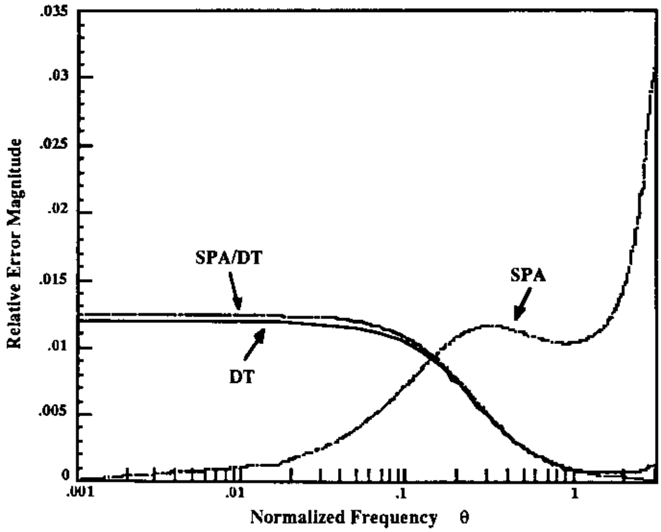

Figure 7. Relative frequency errors of the reduction (discrete-time).

case, since, for discrete-time systems, the reduction error bound for the direct truncation of a balanced system is no longer tight, even when the order reduction is 1 (see e.g. Al-Saggaf 1986).

	DT	SPA	SPA/DT	Theoretical error
$\|G(j\omega)-G_r(j\omega)\|_\infty$	2·2602	2·4803	2·3552	2·7042
Exact error at DC	2·2602	0·0	2·3553	—

Table 2. Frequency errors of the reduction (Example 2) ($\times 10^{-4}$).

6. Conclusions

In this paper we have shown that direct-truncation reduction and the slow singular perturbation approximation of a stable internally balanced continuous-time system are two fully compatible model order-reduction techniques, in the sense that both methods yield a minimal, stable and balanced reduced-order system with the same L_∞-norm frequency error bound on the reduction. We have also shown that, although the upper bound for both methods is the same, the actual frequency errors of these two reduction methods are quite different. Direct-truncation reduction tends to have smaller errors at high frequencies and larger errors at low frequencies, whereas the singular perturbation approximation will have larger errors at high frequencies and smaller errors at low frequencies, while directly matching the DC gain of the reduced-order system with the DC gain of the original system.

We have also established that a certain bilinear mapping preserves not only the balanced structure between the continuous-time system and the discrete-time system, but also the singular perturbation approximation structure between the reduced-order continuous-time system and the reduced-order discrete-time system. Hence the results on the singular perturbation approximation of the continuous-time, stable and balanced system can be easily extended to the discrete-time case. However, it should be pointed out that in the discrete-time case the direct-truncation reduction method is *not* fully compatible with the singular perturbation reduction method in the sense that the former method gives a stable, minimal but not balanced reduced-order system.

Finally, we mention that the above two model order-reduction techniques can be used in a mixed way in the order-reduction procedure of a balanced system (certain restrictions must be imposed in the discrete-time case).

Appendix A

Proof of Theorem 3.4

We need to prove that for $\sigma_0 = j\xi$, with ξ real,

$$\|G(j\omega) - \bar{C}(\sigma_0)[j\omega I_r - \bar{A}(\sigma_0)]^{-1}\bar{B}(\sigma_0) - \bar{E}(\sigma_0)\|_\infty \leqslant 2\,\mathrm{tr}\,(\Sigma_2) \qquad \text{(A 1)}$$

where $G(s) = C(sI_n - A)^{-1}B + E$ is a stable and internally balanced system with the gramian Σ, where $\Sigma = \mathrm{diag}\,\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$ with $\sigma_i \geqslant \sigma_{i+1}$, $i = 1, 2, \ldots, n-1$, and is partitioned conformally as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \quad C_2], \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \qquad \text{(A 2)}$$

Also, $\bar{A}(\sigma)$, $\bar{B}(\sigma)$, $\bar{C}(\sigma)$ and $\bar{E}(\sigma)$ are defined as

$$\left.\begin{aligned} \bar{A}(\sigma) &= A_{11} + A_{12}(\sigma I_{n-r} - A_{22})^{-1}A_{21} \\ \bar{B}(\sigma) &= B_1 + A_{12}(\sigma I_{n-r} - A_{22})^{-1}B_2 \\ \bar{C}(\sigma) &= C_1 + C_2(\sigma I_{n-r} - A_{22})^{-1}A_{21} \\ \bar{E}(\sigma) &= E + C_2(\sigma I_{n-r} - A_{22})^{-1}B_2 \end{aligned}\right\} \qquad \text{(A 3)}$$

It can be shown easily by the block-matrix inversion lemma (Kailath 1980) that

$$\begin{aligned} \bar{C}(\sigma_0)[j\omega I_r - \bar{A}(\sigma_0)]^{-1}\bar{B}(\sigma_0) - \bar{E}(\sigma_0) &= C\left[\begin{pmatrix} j\omega I_r & 0 \\ 0 & \sigma I_{n-r} \end{pmatrix} - A\right]^{-1} B + E \\ &= C\left[\begin{pmatrix} j\omega I_r & 0 \\ 0 & j\xi I_{n-r} \end{pmatrix} - A\right]^{-1} B + E \end{aligned}$$

Since (A, B, C) is balanced, we have

$$\left.\begin{aligned} A\Sigma + \Sigma A^{\mathrm{T}} + BB^{\mathrm{T}} &= 0 \\ A^{\mathrm{T}}\Sigma + \Sigma A + C^{\mathrm{T}}C &= 0 \end{aligned}\right\} \qquad \text{(A 4)}$$

From these Lyapunov equations we have

$$\begin{aligned} &\left(\begin{bmatrix} j\omega I_r & 0 \\ 0 & j\xi I_{n-r} \end{bmatrix} - A\right)^{-1}\Sigma + \Sigma\left(\begin{bmatrix} -j\omega I_r & 0 \\ 0 & -j\xi I_{n-r} \end{bmatrix} - A^{\mathrm{T}}\right)^{-1} \\ &= \left(\begin{bmatrix} j\omega I_r & 0 \\ 0 & j\xi I_{n-r} \end{bmatrix} - A\right)^{-1} BB^{\mathrm{T}}\left(\begin{bmatrix} -j\omega I_r & 0 \\ 0 & -j\xi I_{n-r} \end{bmatrix} - A^{\mathrm{T}}\right)^{-1} \end{aligned} \qquad \text{(A 5 a)}$$

$$(-j\omega I_n - A^{\mathrm{T}})^{-1}\Sigma + \Sigma(j\omega I_n - A)^{-1} = (-j\omega I_n - A^{\mathrm{T}})^{-1}C^{\mathrm{T}}C(j\omega I_n - A)^{-1} \qquad \text{(A 5 b)}$$

Observe that

$$(j\omega I_n - A)^{-1} - \left(\begin{bmatrix} -j\omega I_r & 0 \\ 0 & -j\xi I_{n-r} \end{bmatrix} - A^{\mathrm{T}}\right)^{-1}$$

$$= (j\omega I_n - A)^{-1}(\xi - \omega) j \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix}\left(\begin{bmatrix} -j\omega I_r & 0 \\ 0 & -j\xi I_{n-r} \end{bmatrix} - A^{\mathrm{T}}\right)^{-1} \quad \text{(A 6)}$$

Assume now that $r = n - 1$; we have

$$\bar{\sigma}^2\{C(j\omega I_n - A)^{-1}B + E - \bar{C}(\sigma_0)[j\omega I_r - \bar{A}(\sigma_0)]^{-1}\bar{B}(\sigma_0) - \bar{E}(\sigma_0)\}$$

$$= \bar{\sigma}^2\left\{C(j\omega I_n - A)^{-1}B - C\left(\begin{bmatrix} j\omega I_r & 0 \\ 0 & j\xi I_{n-r} \end{bmatrix} - A\right)^{-1}B\right\}$$

$$= \bar{\sigma}^2\left\{C(j\omega I_n - A)^{-1}(\xi - \omega) j \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} -j\omega I_r & 0 \\ 0 & -j\xi I_{n-r} \end{bmatrix} - A^{\mathrm{T}}\right)^{-1}B\right\}$$

$$= \lambda_{\max}\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}(-j\omega I_n - A^{\mathrm{T}})^{-1}C^{\mathrm{T}}C(j\omega I_n - A)^{-1}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}(\xi - \omega)^2\right.$$

$$\times \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} j\omega I_r & 0 \\ 0 & j\xi I_{n-r} \end{bmatrix} - A\right)^{-1}$$

$$\left.\times BB^{\mathrm{T}}\left(\begin{bmatrix} -j\omega I_r & 0 \\ 0 & -j\xi I_{n-r} \end{bmatrix} - A^{\mathrm{T}}\right)^{-1}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$$

$$= \lambda_{\max}\{[\Delta^{-1} + (\Delta^*)^{-1}]\sigma_n(\xi - \omega)^2\sigma_n[\tilde{\Delta}^{-1} + (\tilde{\Delta}^*)^{-1}]\}$$

where Δ^{-1} and $\tilde{\Delta}^{-1}$ are defined as follows

$$\begin{bmatrix} 0 & 0 \\ 0 & \Delta^{-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}(j\omega I_n - A)^{-1}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & \tilde{\Delta}^{-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} j\omega I_{n-1} & 0 \\ 0 & j\xi \end{bmatrix} - A\right)^{-1}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

By the matrix-inversion lemma (Kailath 1980), it is easy to see that

$$\Delta = j\omega - A_{22} - A_{21}(j\omega I_{n-1} - A_{11})^{-1}A_{12}$$

$$\tilde{\Delta} = j\xi - A_{22} - A_{21}(j\omega I_{n-1} - A_{11})^{-1}A_{12}$$

Define

$$\bar{\Delta} = -A_{22} - A_{21}(j\omega I_{n-1} - A_{11})^{-1}A_{12} \triangleq \alpha + j\beta$$

where α and β are real functions of ω; then

$$\Delta = j\omega + \alpha + j\beta = \alpha + j(\omega + \beta)$$

$$\tilde{\Delta} = j\xi + \alpha + j\beta = \alpha + j(\xi + \beta)$$

So we have

$$\bar{\sigma}^2\{C(j\omega I_n - A)^{-1}B + E - \bar{C}(\sigma_0)[j\omega I_{n-1} - \bar{A}(\sigma_0)]^{-1}\bar{B}(\sigma_0) - \bar{E}(\sigma_0)\}$$
$$= \sigma_n^2 \frac{2\alpha}{\alpha^2 + (\omega+\beta)^2}(\xi-\omega)^2 \frac{2\alpha}{\alpha^3 + (\xi+\beta)^2}$$
$$= 4\sigma_n^2 \frac{\alpha^2(\xi-\omega)^2}{\alpha^2 + (\omega+\beta)^2][\alpha^2 + (\xi+\beta)^2]}$$

Define

$$f \triangleq [\alpha^2 + (\omega+\beta)^2][\alpha^2 + (\xi+\beta)^2] - \alpha^2(\xi-\omega)^2$$

We have

$$f = (\alpha^2 + \omega^2 + 2\omega\beta + \beta^2)(\alpha^2 + \xi^2 + 2\xi\beta + \beta^2) - \alpha^2(\xi^2 - 2\xi\omega + \omega^2)$$
$$= (\alpha^2+\beta^2)^2 + (\alpha^2+\beta^2)(2\xi\beta + 2\omega\beta) + (\xi+\omega\beta)^2 + \omega^2\xi^2$$
$$+ 2\omega^2\xi\beta + 2\omega\beta\xi^2 + 2\xi\omega(\alpha^2+\beta^2)$$
$$= (\alpha^2 + \beta^2 + \xi\beta + \omega\beta + \xi\omega)^2 \geqslant 0$$

for all real α, β, ξ, ω. This means that

$$\bar{\sigma}\{C(j\omega I_n - A)^{-1}B + E - \bar{C}(\sigma_0)[j\omega I_{n-1} - \bar{A}(\sigma_0)]^{-1}\bar{B}(\sigma_0) - \bar{E}(\sigma_0)\} \leqslant 2\sigma_n \quad \text{(A 7)}$$

for all real ω and ξ.

If $r < n-1$, we define

$$P_n(s) \triangleq C(sI_n - A)^{-1}B$$
$$P_r(s,\sigma) \triangleq C\left(\begin{bmatrix} sI_r & 0 \\ 0 & \sigma I_{n-r} \end{bmatrix} - A\right)^{-1} B$$

We obtain

$$\|C(j\omega I_n - A)^{-1}B + E - \bar{C}(\sigma_0)[j\omega I_r - \bar{A}(\sigma_0)]^{-1}\bar{B}(\sigma_0) - \bar{E}(\sigma_0)\|_\infty$$
$$= \|P_n(j\omega) - P_r(j\omega,\sigma_0)\|_\infty$$

We now have

$$\|P_n(j\omega) - P_r(j\omega,\sigma_0)\|_\infty \leqslant \|P_n(j\omega) - P_{n-1}(j\omega,\sigma_0)\|_\infty$$
$$+ \|P_{n-1}(j\omega,\sigma_0) - P_{n-2}(j\omega,\sigma_0)\|_\infty$$
$$+ \ldots + \|P_{r+1}(j\omega,\sigma_0) - P_r(j\omega,\sigma_0)\|_\infty$$

We have shown in (A 7) that $\|P_n(j\omega) - P_r(j\omega,\sigma_0)\|_\infty \leqslant 2\sigma_n$. To prove that $\|P_k(j\omega) - P_{k-1}(j\omega,\sigma_0)\|_\infty \leqslant 2\sigma_k$ $(k = n-1, n-2, \ldots, r+1)$, we can use a similar procedure to the above. We now only show that $\|P_{n-1}(j\omega) - P_{n-2}(j\omega,\sigma_0)\|_\infty \leqslant 2\sigma_{n-1}$. We have

$$P_{n-1}(s,\sigma_0) = C\left(\begin{bmatrix} sI_{n-2} & & \\ & s & \\ & & \sigma_0 \end{bmatrix} - A\right)^{-1} B$$

$$P_{n-2}(s,\sigma_0)=C\left(\begin{bmatrix} sI_{n-2} & & \\ & \sigma_0 & \\ & & \sigma_0\end{bmatrix}-A\right)^{-1}B$$

Hence

$$P_{n-1}(s,\sigma_0)-P_{n-2}(s,\sigma_0)=C\left\{\left(\begin{bmatrix} sI_{n-2} & & \\ & s & \\ & & \sigma_0\end{bmatrix}-A\right)^{-1}\right.$$

$$\left.-\left(\begin{bmatrix} sI_{n-2} & & \\ & \sigma_0 & \\ & & \sigma_0\end{bmatrix}-A\right)^{-1}\right\}B$$

$$=C\left(\begin{bmatrix} sI_{n-2} & & \\ & s & \\ & & \sigma_0\end{bmatrix}-A\right)^{-1}\begin{bmatrix} 0 & & \\ & \sigma_0-s & \\ & & 0\end{bmatrix}$$

$$\times\left(\begin{bmatrix} sI_{n-2} & & \\ & \sigma_0 & \\ & & \sigma_0\end{bmatrix}-A\right)^{-1}B$$

Then, using a similar technique to that used before, we obtain

$$\bar{\sigma}^2[P_{n-1}(j\omega,\sigma_0)-P_{n-2}(j\omega,\sigma_0)]=\lambda_{\max}\left\{\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}[\Delta^{-1}+(\Delta^*)^{-1}]\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}(\xi-\omega)^2\right.$$

$$\left.\times\,\sigma_{n-1}^2\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\times[\tilde{\Delta}^{-1}+(\tilde{\Delta}^*)^{-1}]\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\right\}$$

where

$$\Delta=\begin{bmatrix} j\omega & 0\\ 0 & j\xi\end{bmatrix}-A_{22}-A_{21}(j\omega I_{n-2}-A_{11})^{-1}A_{12}$$

$$\tilde{\Delta}=\begin{bmatrix} j\xi & 0\\ 0 & j\xi\end{bmatrix}-A_{22}-A_{21}(j\omega I_{n-2}-A_{11})^{-1}A_{12}$$

Define

$$\bar{\Delta}=-A_{22}-A_{21}(j\omega I_{n-2}-A_{11})^{-1}A_{12}\triangleq\begin{bmatrix}\bar{\Delta}_{11} & \bar{\Delta}_{12}\\ \bar{\Delta}_{21} & \bar{\Delta}_{22}\end{bmatrix}$$

Then the (1, 1) block of Δ^{-1} becomes

$$(\Delta^{-1})_{11}=[j\omega+\bar{\Delta}_{11}-\bar{\Delta}_{12}(j\xi+\bar{\Delta}_{22})^{-1}\bar{\Delta}_{21}]^{-1}$$

and we obtain

$$(\tilde{\Delta}^{-1})_{11} = [j\xi + \bar{\Delta}_{11} - \bar{\Delta}_{12}(j\xi + \bar{\Delta}_{22})^{-1}\bar{\Delta}_{21}]^{-1}$$

Let

$$\alpha + j\beta \triangleq \bar{\Delta}_{11} - \bar{\Delta}_{12}(j\xi + \bar{\Delta}_{22})^{-1}\bar{\Delta}_{21}$$

where α and β are real functions of ω and ξ; then

$$\begin{aligned}\bar{\sigma}^2[P_{n-1}(j\omega, \sigma_0) - P_{n-2}(j\omega, \sigma_0)] &= \lambda_{\max}\{[(j\omega + \alpha + j\beta)^{-1} + (-j\omega + \alpha - j\beta)^{-1}](\xi - \omega)^2\sigma_{n-1}^2 \\ &\qquad \times [(j\xi + \alpha + j\beta)^{-1} + (-j\xi + \alpha - j\beta)^{-1}]\} \\ &= 4\sigma_{n-1}^2 \frac{\alpha^2(\xi - \omega)^2}{[\alpha^2 + (\omega + \beta)^2][\alpha^2 + (\xi + \beta)^2]} \leqslant 4\sigma_{n-1}^2\end{aligned}$$

for all real α, β, ξ, ω. Hence

$$\|P_{n-1}(j\omega, \sigma_0) - P_{n-2}(j\omega, \sigma_0)\|_\infty \leqslant 2\sigma_{n-1}$$

This concludes the proof. □

References

AL-SAGGAF, U. M., 1986, On model reduction and control of discrete time systems. Ph.D. thesis, Department of Electrical Engineering, Stanford University.

AL-SAGGAF, U. M., and FRANKLIN, G. F., 1986, On model reduction. *Proceedings of the 25th Conference on Decision and Control*, Athens, Greece, pp. 1064–1069.

ENNS, D. F., 1984 a, Model reduction for control systems design. Ph.D. thesis, Department of Aeronautics and Astronautics, Stanford University; 1984 b, Model reduction with balanced realizations: an error bound and a frequency weighted generalization. *Proceedings of the 23rd Conference on Decision and Control*, Las Vegas, U.S.A., pp. 127–132.

FERNANDO, K. V., and NICHOLSON, H., 1982 a, Singular perturbation model reduction of balanced system. *I.E.E.E. Transactions on Automatic Control*, **27**, 466–468; 1982 b, Singular perturbational model reduction in the frequency domain. *Ibid.*, **27**, 969–970; 1983 a, On the structure of balanced and other principal representations of SISO systems. *Ibid.*, **28**, 228–231; 1983 b, Singular perturbational approximations for discrete-time balanced systems. *Ibid.*, **28**, 240–242; 1983 c, Reciprocal transformations in balanced model-order reduction. *Proceedings of the Institution of Electrical Engineers*, Pt D, **130**, 359–362.

GLOVER, K., 1984, All optimal Hankel-norm approximations of linear multivariable systems and their L_∞-error bounds. *International Journal of Control*, **39**, 1115–1193.

JAMSHIDI, M., 1981, An overview on the aggregation of large-scale systems. Preprints of 8th IFAC World Congress, Paper No. 44.1, Kyoto, Japan.

KAILATH, T., 1980, *Linear Systems* (Englewood Cliffs, New Jersey: Prentice-Hall).

KOKOTOVIC, P. V., O'MALLEY, R. E., and SANNUTI, P., 1976, Singular perturbations and order reduction in control theory—an overview. *Automatica*, **12**, 123–132.

MOORE, B. C., 1982, Principal component analysis in linear systems: controllability, observability, and model reduction. *I.E.E.E. Transactions on Automatic Control*, **26**, 17–32.

PERNEBO, L., and SILVERMAN, L. M., 1982, Model reduction via balanced state space representations. *I.E.E.E. Transactions on Automatic Control*, **27**, 382–387.

SAKSENA, V. R., O'REILLY, J., and KOKOTOVIC, P. V., 1984, Singular perturbations and time-scale methods in control theory: survey 1976–1983. *Automatica*, **20**, 273–293.

SANTIAGO, J. M., and JAMSHIDI, M., 1986, On the extensions of the balanced approach for model reduction. *Control Theory and Advanced Technology*, **2**, 207–226.

SREERAM, V., and AGATHOKLIS, P., 1989, Model reduction using balanced realizations with improved low order frequency behaviour. *Systems and Control Letters*, **12**, 33–38.