# Balanced model reduction of linear systems with nonzero initial conditions: Singular perturbation approximation

Adnan Daraghmeh [a,*], Carsten Hartmann [b], Naji Qatanani [a]

[a] *Department of Mathematics, Faculty of Science, An-Najah National University, Nablus, Palestine*
[b] *Institute of Mathematics, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus 03046, Germany*

**A B S T R A C T**

In this article we study balanced model reduction of linear control systems using the singular perturbation approximation. Balanced model reduction techniques have been successfully applied to systems with *homogeneous* initial conditions, with one of their most important features being a priori $L^2$ and $H^\infty$ bounds for the approximation error. The main focus of this work is to derive an $L^2$ error bound for the singular perturbation approximation for system with *inhomogeneous* initial conditions, extending related work on balanced truncation. This $L^2$ error bound measures the difference between the input-output maps of the original and of the reduced initial value systems. The advantages and flexibility of this approach are demonstrated with a variety of numerical examples.

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

Model reduction is a major issue for control, optimization, and simulation of large-scale systems. In particular, for linear time-invariant systems, balanced truncation is a well-established tool for deriving reduced (i.e., low-dimensional) models that have an input-output behavior similar to the original model [2,36]; see also [6] and the references therein. The general idea of balanced truncation is to restrict the system onto the subspace of easily controllable and observable states which can be determined by computing Hankel singular values associated with the system. Moreover the method is known to preserve certain properties of the original system such as stability or passivity and gives an error bound that is easily computable [17]. Balanced model reduction of linear control systems has been under investigation for quite a long time due to the ubiquity of large-scale linear systems in a wide range of applications in science and engineering [2]. For example, the semi-discretization of partial differential equations describing physical, chemical or biological phenomena results in high-dimensional linear time-invariant (LTI) systems. We consider model reduction for linear time-invariant dynamical systems with inhomogeneous initial conditions and a state-space realization

$$\frac{dx}{dt} = Ax + Bu$$
$$y = Cx + Du$$
$$x(t_0) = x_0, \tag{1}$$

---

* Corresponding author.

*E-mail addresses:* adn.daraghmeh@najah.edu (A. Daraghmeh), hartmanc@b-tu.de (C. Hartmann), nqatanani@najah.edu (N. Qatanani).

where $x \in \mathbb{R}^n$ is an $n$-dimensional, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^l$ are, the states, inputs, and outputs of (1), and $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}, C \in \mathbb{R}^{l \times n}$, and $D \in \mathbb{R}^{l \times m}$ are matrices of appropriate size; $x_0 \in \mathbb{R}^n$ is the nonzero initial condition, prescribed at $t_0 = 0$. A common feature of (1) is that it is high-dimensional, with $n$ ranging from a few tens to several thousands as in control problems for large flexible space structures, and that it displays a variety of time scales. If the time scales in the system are well separated, it is possible to eliminate the fast degrees of freedom and to derive low-order reduced models, using averaging and homogenization techniques. Homogenization of feedback and open-loop control systems has been studied by various authors [1,7,12,22]. In many applications, however, an explicit smallness parameter that characterizes the time scales present in the dynamics is not available, so that perturbative methods such as averaging or homogenization cannot be applied.

### 1.1. Related work

Balanced model reduction going back to Moore [26] provides a rational basis for various approximation techniques such as Hankel norm approximations [35] or the singular perturbation approximation [27], and it includes easily computable error bounds [14]; see also [2,36] and the references given there. The general idea of balanced model reduction is to restrict the system to the subspace of easily controllable and observable states which can be determined by the Hankel singular values associated with the system. All these methods give a stable reduced system and a guaranteed upper bound of the approximation error, provided that the initial conditions are homogeneous, i.e. that $x_0 = 0$ in Eq. (1).

Balanced model reduction of linear systems with inhomogeneous initial conditions has received very little attention, with [4,11,18] being the only exceptions known to the authors. In these papers, the authors extend balanced truncation to the case of inhomogeneous initial conditions (in case of [18] using an $L^2$ regularization of the non-smooth input due to the initial data). In spirit, our approach here is similar to the paper [4] and the recent PhD thesis [11] by one of the authors of this article, but we consider a balanced version of the singular perturbation approximation (SPA). For homogeneous systems it is known that, although balanced truncation (BT) and SPA have the same $H^\infty$ error bound, the frequency characteristics of both methods are contrary to each other, in that balanced truncation yields a smaller error at high frequencies whereas SPA gives a better approximation at low frequencies [23]. Many of the properties of the SPA can be related to the properties of the corresponding truncated system through the balanced reciprocal system as is shown in [27], and our approach, as is detailed below, largely exploits the reciprocality relations between the corresponding transfer functions. Typical applications, in which nonzero initial conditions play a role, involve the optimal control of partial differential equations [8] and piecewise-linear model reduction that is used in circuit design [33], model predictive control [9], computational fluid dynamics [15], or uncertainty quantification [13], to mention just a few examples. The SPA comes into play whenever high model fidelity at low frequencies is sought, i.e., when the reduced-order model ought to approximate the slow modes accurately. Capturing the low frequency modes rather than the high frequency (i.e. fast) modes may be relevant when the model is used for predicting long-term effects (e.g. [31]). In contrast, balanced *truncation* that aims at approximating high frequency modes may be relevant for problems such as circuit design (e.g. [25]). We mention the closely related linear quadratic regulator problems, that are among the most actively investigated SPA of linear control problems [29,30]. Nevertheless the approaches therein are very different from ours, in that they are based on asymptotics and limit arguments, whereas our error bound is not based on a small parameter that goes to zero. Hence we believe that our approach can be relevant for obtaining high-fidelity reduced-order models when solving optimal feedback control problems in situations in which a small parameter is not available.

### 1.2. Contribution of this work and outline

In this article we present an extension balanced model reduction for linear control systems with nonzero initial conditions using the singular perturbation approximation. As one of the most important features of this method, it allows for a priori $L^2$ and $H^\infty$ bounds for the approximation error. This method has been successfully applied for systems with homogeneous initial conditions, however, the main focus in this work is to derive an $L^2$ error bound for singular perturbation approximation for system with inhomogeneous initial conditions, extending the work by Antoulas. The theoretical results are validated numerically.

This article is structured as follows: Section 2 reviews the standard $H^\infty$ error bound for balanced model reduction with zero initial conditions, introduces the main result and briefly discusses related work. In Section 3 we prove the main result, the SPA error bound in case of nonzero initial conditions, and discuss variations, along with computational aspects and related work. Numerical examples in Section 4 illustrate the theoretical findings. Finally, we conclude the article with a brief discussion in Section 5. The Appendix records various properties of reciprocal systems that are relevant for the proof of the error bound.

## 2. Preliminaries

Consider the linear time-invariant system of the form (1) with initial data $(t_0, x_0) = (0, 0)$. The associated transfer function is given by

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

We assume that the system characterized by the system matrices $A$, $B$, $C$, $D$ satisfies:

**Assumption 1.** The system (1) is asymptotically stable. Furthermore the matrix pair $(A, B)$ is controllable and the pair $(A, C)$ is observable.

By the above assumption, the controllability and observability Gramians $W_c$ and $W_o$ are the unique symmetric positive definite solutions to the pair of Lyapunov equations

$$AW_c + W_cA^T + BB^T = 0$$
$$W_oA + A^TW_o + C^TC = 0 \tag{2}$$

We call (1) balanced if the Gramians are equal and diagonal, $W_c = W_o = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ for some real numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$, called Hankel singular values (HSV). For a general system of the standard form (1), let $S \in \mathbb{R}^{n \times n}$ a regular matrix with the property

$$SW_cS^T = S^{-T}W_oS^{-1} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n). \tag{3}$$

Then the system with coefficient matrices

$$(SAS^{-1}, SB, CS^{-1}, D) \tag{4}$$

is balanced. In the following and unless further notice is given, we assume that our system is balanced:

**Assumption 2.** Let $1 \leq r \leq n$. The system (1) with coefficients $(A, B, C, D)$ is balanced, and the matrices $A$, $B$, $C$ are conformingly partitioned as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \ B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \ C = \begin{pmatrix} C_1 & C_2 \end{pmatrix}, \tag{5}$$

where $A_{11}, A_{12}, \ldots, B_1, \ldots$, etc. are matrices of sizes $r \times r$, $r \times (n - r)$, $\ldots$, $r \times m$, $\ldots$, etc.

Written in partitioned form, Eq. (1) reads

$$\dot{x}_1 = A_{11}x_1 + A_{12}x_2 + B_1u, \quad x_1(0) = 0$$
$$\dot{x}_2 = A_{21}x_1 + A_{22}x_2 + B_2u, \quad x_2(0) = 0$$
$$y = C_1x_2 + C_2x_2 + Du, \tag{6}$$

where the dot denotes time derivative. Applying balanced truncation to the balanced system (6) amounts to setting $x_2 = 0$, which yields a reduced $r$th order system of the form [2,36]

$$\dot{x}_r = A_{11}x_r + B_1u, \quad x_r(0) = 0$$
$$y_r = C_1x_r + Du, \tag{7}$$

with transfer function

$$G_r(s) = C_1(sI - A_{11})^{-1}B_1 + D \tag{8}$$

The next result is standard (see e.g. [2])

**Theorem 1.** Let $(A_r, B_r, C_r, D_r)$ be the rank-r approximation of (1) with transfer function $G_r$. If $\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ and $\Sigma_2 = \mathrm{diag}(\sigma_{r+1}, \ldots, \sigma_n)$ have no common element, then

$$\|G - G_r\|_{H^\infty} \leq 2\mathrm{tr}(\Sigma_2) \tag{9}$$

where $\|G\|_{H^\infty} = \max\{\|G(i\omega)\|_2 : \omega \in \mathbb{R}\|\}$ denotes the $H^\infty$ norm.

The aim of this paper is to study the SPA for systems with inhomogeneous initial data. The validity of the SPA, even for the homogeneous case, requires an extra assumption:

**Assumption 3.** All eigenvalues of the matrix $A_{22}$ have strictly negative real part.

Now, formally, the SPA amounts to setting $\dot{x}_2 = 0$ in (6) and solving the second equation for $x_2$, which yields a reduced system of the form (assuming that $u(0) = 0$)

$$\frac{d\tilde{x}_r}{dt} = (A_{11} - A_{12}A_{22}^{-1}A_{21})\tilde{x}_r + (B_1 - A_{12}A_{22}^{-1}B_2)u, \quad \tilde{x}_r(0) = 0$$
$$\tilde{y}_r = (C_1 - C_2A_{22}^{-1}A_{21})\tilde{x}_r + (D - C_2A_{22}^{-1}B_2)u, \tag{10}$$

with corresponding transfer function $\tilde{G}_r$. The error bound of Theorem 1 equally applies to $\tilde{G}_r$ as has been shown in [24]. As our result builds on the homogeneous SPA error bound, we give a short sketch of derivation in Appendix A for the reader's convenience.

**Remark 1.** Assumption 3 is standard for singularly perturbed systems of differential equations. It is not particularly restrictive as stability of $A_{22}$ follows from the stability of the balanced matrix $A$ under fairly mild conditions (cf. [2, Theorem 7.9, Part I] or [5]).

*2.1. Main result and related work*

Even though the error bound of Theorem 1 holds for both balanced truncation and the SPA, it is not valid if $x_0 \neq 0$, in fact the error can grow unboundedly, depending on the initial data. Balanced truncation of linear systems with nonzero initial conditions has been analysed only recently in [4,18], and this work is devoted to an extension of these results to the SPA. If there is a given initial condition $x_0$, a possible remedy (see [3]) is to do a translation $x \mapsto x - x_0$ of the state vector, but often the initial condition may either not be known *a priori* or change in the course of a simulation. Therefore we adopt a different approach that is in the spirit of the aforementioned paper [18] and assume that the initial conditions $x_0$ lie in the column space of some matrix $M \in \mathbb{R}^{n \times k}$.

An obvious idea could then be to consider $x_0 = Mz_0$ for some variable $z_0 \in \mathbb{R}^k$ and to treat $M$ as an extra input matrix, which would give rise to the augmented transfer function

$$G_a(s) = C(sI - A)^{-1}[B \, M] + D. \tag{11}$$

Treating the initial conditions on the same footing as the control $u$, however, is problematic, because such an input would be of the form $u_t = z_0 \delta(t - t_0)$, but clearly $\delta(\cdot) \notin L^2$, and hence the associated transfer function $G_a$ would not be well-defined. In the article [18], the authors have proposed to replace the Dirac delta function $\delta(\cdot)$ by a suitable $L^2$ regularization $\delta_\epsilon(\cdot)$ which then yields an error bound that depends on the particular regularization and the parameter $\epsilon > 0$ (and which necessarily degrades as $\epsilon \to 0$).

Here we follow a slightly different route and separately estimate the two sources of error in the $L^2$ norm coming from the smooth input $u$ and the initial conditions $x_0 = Mz_0$. To this end let $f : [t_0, \infty) \to \mathbb{R}^l$ be a measurable function and define the corresponding $L^2$-norm as

$$\|f\|_{L^2} = \sqrt{\int_{t_0}^{\infty} |f(t)|^2 \, dt}. \tag{12}$$

Our main result, Theorem 2 below, that will be proved in the next section is the following $L^2$ error bound for the singular perturbation approximation of (1):

$$\|y - \tilde{y}_r\|_{L^2} \leq c_r \|z_0\|_2 + 2\mathrm{tr}(\Sigma_2)\|u\|_{L^2} \tag{13}$$

with a constant $c_r = c_r(A, M, C)$ that is related to the $H^2$ norm of the auxiliary system $(A, M, C, 0)$ and that goes to zero as the rank $r$ of the reduced model goes to $n$.

A similar approach has been taken in the recent article [4] for balanced truncation, where the main difference is that therein—besides the fact that our result holds for the SPA—the control part $(A, B, C, D)$ and the auxiliary system $(A, M, C, 0)$ are balanced separately, whereas we suggest to balance only the control part and to reduce the auxiliary system correspondingly, which, depending on the rank of the matrix $M$, is computationally far less expensive. The advantages and disadvantages of both approaches will be discussed later on in Sections 3.2 and 4, but we stress that it possible to adapt our error bound to the framework of [4]. Related results are have been obtained in the PhD thesis [11], however the error bound for the auxiliary system is clearly suboptimal compared to the one derived in this paper.

## 3. Singular perturbation approximation of systems with nonzero initial conditions

We shall now derive an error bound for the SPA of a linear system with nonzero initial conditions. Our strategy is based the representation of the SPA as a reciprocal truncated system (see Appendix A). To this end recall that the reciprocal system (57) to (1) has the coefficients

$$\hat{A} = A^{-1}, \quad \hat{B} = A^{-1}B, \quad \hat{C} = -CA^{-1}, \quad \hat{D} = D - CA^{-1}B. \tag{14}$$

Let the initial condition of the reciprocal system be defined by

$$\hat{x}_0 = x_0 = (\eta, \xi)^T, \tag{15}$$

with $(\eta, \xi)$ denoting the partitioning of the initial conditions into non-negligible and negligible components. The partitioned reciprocal system with nonzero initial conditions now reads

$$\begin{aligned} \frac{d\hat{x}_1}{dt} &= \hat{A}_{11}\hat{x}_1 + \hat{A}_{12}\hat{x}_2 + \hat{B}_1 u, \quad \hat{x}_1(t_0) = \eta \\ \frac{d\hat{x}_2}{dt} &= \hat{A}_{21}\hat{x}_1 + \hat{A}_{22}\hat{x}_2 + \hat{B}_2 u, \quad \hat{x}_2(t_0) = \xi \\ \hat{y} &= \hat{C}_1\hat{x}_2 + \hat{C}_2\hat{x}_2 + \hat{D}u, \end{aligned} \tag{16}$$

Formally, balanced truncation amounts to setting the second component $\hat{x}_2$—and thus $\xi$—to zero, which then yields the reduced system

$$\frac{d\hat{x}_r}{dt} = \hat{A}_{11}\hat{x}_r + \hat{B}_1 u, \quad \hat{x}_r(t_0) = \eta$$
$$\hat{y}_r = \hat{C}_1\hat{x}_r + \hat{D}_1 u, \tag{17}$$

with the coefficients as given by (62). A natural candidate for the SPA thus is[1]

$$\frac{d\tilde{x}_r}{dt} = (A_{11} - A_{12}A_{22}^{-1}A_{21})\tilde{x}_r + (B_1 - A_{12}A_{22}^{-1}B_2)u, \quad \tilde{x}_r(t_0) = \eta$$
$$\tilde{y}_r = (C_1 - C_2A_{22}^{-1}A_{21})\tilde{x}_r + (D - C_2A_{22}^{-1}B_2)u. \tag{18}$$

*Loss of reciprocality*

Before we estimate of the approximation error between (1) and (18), let us first study the relation between the corresponding transfer functions. To this end, let

$$\hat{y}(t) = \hat{C}e^{\hat{A}(t-t_0)}\hat{x}_0 + \hat{C}\int_{t_0}^{t} e^{\hat{A}(t-s)}\hat{B}u(s)ds + \hat{D}u(t) \tag{19}$$

be the solution of (16) with initial condition $\hat{x}_0 = Mz_0 = (\eta, \xi)$. Identifying the transfer function $\hat{H}$ with the Laplace transform of the linear map that maps the augmented input $(z_0, u) \in \mathbb{R}^k \times L^2([t_0, \infty), \mathbb{R}^m)$ to the function $\hat{y} \in L^2([t_0, \infty), \mathbb{R}^l)$, we obtain

$$\hat{H}(s) = \hat{g}(s) + \hat{G}(s) \tag{20}$$

for the transfer function of the inhomogeneous system, with

$$\hat{g}(s) = \hat{C}(sI - \hat{A})^{-1}M, \quad \hat{G}(s) = \hat{C}(sI - \hat{A})^{-1}\hat{B} + \hat{D} \tag{21}$$

representing the transfer functions of the auxiliary and the control system. By Lemma 5 in the appendix it holds that

$$G(s) = \hat{G}(1/s), \tag{22}$$

where $G$ is the transfer function associated with the linear map

$$(\Gamma u)(t) = C\int_{t_0}^{t} e^{A(t-s)}Bu(s)ds + Du(t), \tag{23}$$

i.e., the solution of the original system (1) with initial condition $x_0 = 0$. Nevertheless,

$$H(s) \neq \hat{H}(1/s) \tag{24}$$

since the transfer functions $\hat{g}$ and $g(s) = C(sI - A)^{-1}M$ associated with the auxiliary systems are not reciprocal to each other. As a consequence, the initial condition changes the reciprocality relation between the transfer functions which is due to the fact that we do not treat the matrix $M$ as an extra input matrix, like the matrix $B$.

**Remark 2.** It is possible to restore reciprocality by replacing the matrix $M$ in (21) by $\hat{M} = A^{-1}M = \hat{A}M$ and adding the constant feedthrough term

$$\hat{E} = -\hat{C}\hat{A}^{-1}\hat{M}z_0 \tag{25}$$

to the second equation in (16). However, unless $z_0 = 0$, this would mean that $y \notin L^2$, contradicting the overall balancing procedure (see also the paragraph "Discussion" below).

### 3.1. Approximation error

We will now estimate the approximation error between (1) and (18) by separately estimating the contributions coming from the auxiliary system and the control part. The following Theorem provides an error bound in $L^2([t_0, \infty), \mathbb{R}^l)$.

**Theorem 2.** Let (1) satisfy Assumptions 1–3, with nonzero initial conditions $x_0 = Mz_0$ for some matrix $M = (M_1^T M_2^T)^T \in \mathbb{R}^{n \times k}$ and an arbitrary vector $z_0 \in \mathbb{R}^k$. Further let $\tilde{y}_r$ be the output of (18) with initial conditions $\tilde{x}_r(0) = M_1 z_0$. Then

$$\|y - \tilde{y}_r\|_{L^2} \leq c_r \|z_0\|_2 + 2\text{tr}(\Sigma_2)\|u\|_{L^2}, \tag{26}$$

with

$$c_r = \sqrt{\text{tr}\left((M^T \ M_1^T)\begin{pmatrix} \Sigma & -U \\ -U^T & \Sigma_1 \end{pmatrix}\begin{pmatrix} M \\ M_1 \end{pmatrix}\right)} \tag{27}$$

---

[1] Throughout this article, we use the notation $\hat{y}, \hat{y}_r, \hat{G}, \hat{G}_r$, etc. to denote reciprocal systems and their approximants, and $\tilde{y}_r, \tilde{G}_r$ for the corresponding SPA.

where $\Sigma = \mathrm{diag}(\Sigma_1, \Sigma_2) \in \mathbb{R}^{n \times n}$ is the matrix of Hankel singular values of the homogeneous control system, and $U \in \mathbb{R}^{n \times r}$ denotes the solutions to the Sylvester equation

$$A^T U + U \tilde{A}_1 = -C^T \tilde{C}_1. \tag{28}$$

**Proof.** Without loss of generality, we set $t_0 = 0$ and introduce the shorthands

$$\gamma(t) = C e^{At} M, \quad \tilde{\gamma}_r = \tilde{C}_1 e^{\tilde{A}_1 t} \tilde{M}_1 \tag{29}$$

for the auxiliary flow part, with the coefficients

$$\tilde{A}_1 = A_{11} - A_{12} A_{22}^{-1} A_{21}, \quad \tilde{M}_1 = M_1, \quad \tilde{C}_1 = C_1 - C_2 A_{22}^{-1} A_{21}, \tag{30}$$

Further let $\Gamma, \tilde{\Gamma}_r$ denote the corresponding flow parts with $\Gamma$ as given by (23) for $t_0 = 0$ and

$$(\tilde{\Gamma}_r u)(t) = \tilde{C}_1 \int_{t_0}^{t} e^{\tilde{A}_1(t-s)} \tilde{B}_1 u(s) ds + \tilde{D}_1 u(t) \qquad (t_0 = 0) \tag{31}$$

where

$$\tilde{B}_1 = B_1 - A_{12} A_{22}^{-1} B_2, \quad \tilde{D}_1 = D - C_2 A_{22}^{-1} B_2. \tag{32}$$

The triangle inequality and the submultiplicativity of the $L^2$ norm then entail

$$\|y - \tilde{y}_r\|_{L^2} \leq \|\gamma - \tilde{\gamma}_r\|_{L^2} \|z_0\|_2 + \|\Gamma - \tilde{\Gamma}_r\|_{L^2} \|u\|_{L^2}. \tag{33}$$

Now, by Corollary 6 and the isometry between the time and frequency domain spaces $L^2$ and $H^2$, the second term can be bounded by the standard $H^\infty$ error bound (see e.g. [2, Proposition 5.2]):

$$\|\Gamma - \tilde{\Gamma}_r\|_{L^2} = \sup_{u \in L^2} \frac{\|(\Gamma - \tilde{\Gamma}_r) u\|_{L^2}}{\|u\|_{L^2}} = \sup_{\omega \in \mathbb{R}} \|\Gamma(i\omega) - \tilde{\Gamma}_r(i\omega)\|_2 \leq 2\mathrm{tr}(\Sigma_2) \tag{34}$$

The first term in (33) is recognized as the difference of the impulse responses of the full and the truncated reciprocal system, which equals (see e.g. [2, Section 5.5.1])

$$\|\gamma - \tilde{\gamma}_r\|_{L^2} = \sqrt{\int_0^\infty \mathrm{tr}\big((\gamma(t) - \tilde{\gamma}_r(t))^T (\gamma(t) - \tilde{\gamma}_r(t))\big) dt}. \tag{35}$$

The difference $\gamma_e = \gamma - \tilde{\gamma}_r$ can be formally represented by an LTI system $\Sigma_e = (A_e, M_e, C_e, 0)$ with the coefficients

$$\Sigma_e = \left( \begin{array}{ccc|c} A_{11} & A_{12} & 0 & M_1 \\ A_{21} & A_{22} & 0 & M_2 \\ 0 & 0 & \tilde{A}_1 & \tilde{M}_1 \\ \hline C_1 & C_2 & -\tilde{C}_1 & 0 \end{array} \right). \tag{36}$$

By interchanging the integral with the trace in (35) and using the invariance of the trace under cyclic permutations, it can be seen that the $L^2$ norm of $\hat{\gamma}_e$ can be expressed in terms of the observability Gramian $P_e$ associated with $\Sigma_e$:

$$\|\gamma_e\|_{L^2}^2 = \int_0^\infty \mathrm{tr}\big(C_e e^{A_e t} M_e M_e^T e^{A_e^T t} C_e^T\big) dt = \mathrm{tr}(M_e^T P_e M_e), \tag{37}$$

with the Gramian $P_e$ being the solution of the Lyapunov equation

$$A_e^T P_e + P A_e = -C_e^T C_e. \tag{38}$$

The specific form of the Gramian,

$$P_e = \begin{pmatrix} \Sigma & -U \\ -U^T & \Sigma_1 \end{pmatrix}, \tag{39}$$

follows from the Sylvester equation (28) and by noting that both the original system and its the SPA are output balanced with Gramians $\Sigma$ and $\Sigma_1$, respectively, i.e.,

$$A^T \Sigma + \Sigma A = -C^T C, \quad \tilde{A}_1^T \Sigma_1 + \Sigma_1 \tilde{A}_1 = -\tilde{C}_1^T \tilde{C}_1, \tag{40}$$

The second Lyapunov equation is balanced, as follows from Lemma 4 and the fact that the coefficients $\tilde{A}_1$ and $\tilde{C}_1$ are reciprocal to the truncated balanced coefficients (62). □

### 3.2. Discussion

We will now discuss the error bound and its relation to existing work. In particular, we shall discuss the role of the initial conditions in the construction of the reduced model, which is *not* included in the balancing procedure and, as a consequence, in the construction of the low-dimensional subspace, on which the reduced dynamics live. First of all, observe that an equivalent representation of (27) is given by

$$c_r = \sqrt{\mathrm{tr}\left(M_2^T \Sigma_2 M_2 + 2M_1^T \Sigma_1 M_1 - 2M_1^T U_1 M_1 - 2M_2^T U_2 M_1\right)} \tag{41}$$

where $U_1 \in \mathbb{R}^{r \times r}$ and $U_2 \in \mathbb{R}^{(n-r) \times r}$ denote the partitioning of $U = (U_1^T, U_2^T)^T \in \mathbb{R}^{n \times r}$. It is plausible that, depending on the specific application, one has control over the initial condition $\eta \in \mathrm{range}(M_1)$ by virtue of $M_1$, but not over $\xi \in \mathrm{range}(M_2)$. Therefore, we define

$$\theta(M_2) = c_r \tag{42}$$

and ask for which choices of $M_2$—with $M_1$ given—the error due to the initial conditions will be extremal, with the aim of identifying typical scenarios and best-case scenarios for the approximation error. Clearly, by the submultiplicativity of the Frobenius norm,

$$\sup_{\|M_2\| \leq 1} \theta(M_2) < \infty \tag{43}$$

where $\|\cdot\|$ is any matrix norm, in other words, the approximation error is bounded from above. Now recall that the SPA via the reciprocal system (17) amounts to setting $M_2 = 0$. However, since $\Sigma_2 > 0$, the function $\theta$ has a unique minimum that is attained at

$$M_2^* = \Sigma_2^{-1} U_2 M_1, \tag{44}$$

which is our best-case scenario for the approximation error, where

$$\theta(M_2^*) - \theta(0) = -M_1^T U_2^T \Sigma_2^{-1} U_2 M_1 \leq 0. \tag{45}$$

Therefore truncating the initial condition at $M_2 = 0$, which was the basis of the derivation of the SPA (18) from the reciprocal system (17), will not be optimal in general in terms of the error bound, even though the difference may be negligible in practice (see Section 4).

*Time scale separation limit*

Another scenario is obtained by choosing $M_2$, so that the initial conditions remain consistent with the "standard" SPA of uncontrolled linear system [20,21]. To see what this means, let $0 < \epsilon \ll 1$ and consider the dynamics

$$
\begin{aligned}
\dot{x}_1^\epsilon &= A_{11} x_1^\epsilon + A_{12} x_2^\epsilon, \quad x_1^\epsilon(0) = M_1 z_0 \\
\epsilon \dot{x}_2^\epsilon &= A_{21} x_1^\epsilon + A_{22} x_2^\epsilon, \quad x_2^\epsilon(0) = M_2 z_0 \\
y^\epsilon &= C_1 x_2^\epsilon + C_2 x_2^\epsilon,
\end{aligned}
\tag{46}
$$

which is a variant of (6) with inhomogeneous initial conditions, $x^\epsilon(0) \neq 0$, but with zero control $u = 0$.

For $0 < \epsilon \ll 1$ the system is stiff and under the Assumptions 1–3, the limiting dynamics as $\epsilon \downarrow 0$ is governed by the differential algebraic equation[2]

$$
\begin{aligned}
\dot{x}_1 &= A_{11} x_1 + A_{12} x_2, \quad x_1(0) = M_1 z_0 \\
0 &= A_{21} x_1 + A_{22} x_2 \\
y &= C_1 x_2 + C_2 x_2
\end{aligned}
\tag{47}
$$

It can be shown (e.g. [19]) that a uniform $\mathcal{O}(\epsilon)$-approximation on the time interval $[0, T]$ requires that the initial conditions in (46) satisfy the algebraic equation

$$A_{21} x_1^\epsilon(0) + A_{22} x_2^\epsilon(0) = 0; \tag{48}$$

otherwise the approximation error will be $\mathcal{O}(1)$ in a transient initial layer of length $\sqrt{\epsilon}$, and it will be $\mathcal{O}(\epsilon)$ for all times $t$ beyond the initial layer. A uniform $\mathcal{O}(\epsilon)$-approximation for all $t \in [0, T]$ for any $T > 0$ therefore requires that

$$M_2 = -A_{22}^{-1} A_{21} M_1, \tag{49}$$

in case of which the limiting Eq. (47) can be reduced to a reduced-order system for $x_1$ only, which is exactly the uncontrolled form of (18):

---

[2] Here $\epsilon$ plays either the role of the largest entry $\sigma_r$ of $\Sigma_2$, assuming that the entries of $\Sigma_1$ are $\mathcal{O}(1)$, or the ratio $\sigma_{r+1}/\sigma_r$ of the largest entry of $\Sigma_2$ and the smallest entry of $\Sigma_1$; cf. [16].

$$\dot{x}_1 = \left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)x_1, \quad x_1(0) = M_1 z_0$$
$$y = \left(C_1 - C_2 A_{22}^{-1} A_{21}\right)x_1. \tag{50}$$

The fact that (50) is the unique best-approximation of the singularly perturbed system of Eqs. (46) as $\epsilon \downarrow 0$ suggests that (44) reduces to (49) in some limit; this what we shall call the *time scale separation scenario*. The following can be shown by a formal asymptotic expansion of the solution to the Sylvester equation (28) and the Lyapunov equation (40) for the balanced observability Gramian in powers of $\epsilon$:

**Conjecture 3.** *Let the matrix $A = A^{\epsilon}$ in (1) be of the form*

$$A^{\epsilon} = \begin{pmatrix} A_{11} & A_{12} \\ \frac{1}{\epsilon}A_{21} & \frac{1}{\epsilon}A_{22} \end{pmatrix}. \tag{51}$$

*Further let $M_2^* = M_2^{*,\epsilon}$ be given by (44). Then, under the conditions of Theorem 2,*

$$M_2^{*,\epsilon} \to -A_{22}A_{21}M_1 \quad \text{as } \epsilon \downarrow 0. \tag{52}$$

We stress that the last statement—that we call a "Conjecture" rather than a "Theorem" as we do not provide a rigorous proof—is an asymptotic result, whereas Theorem 2 is not. As a consequence, setting $M_2 = M_2^*$ or even $M_2 = 0$ in (1) may lead to a smaller overall approximation error than $M_2 = -A_{22}A_{21}M_1$ for finite $\Sigma_2$. We will come back to this point in the discussion of the numerical example in Section 4.

*Balanced approximation of the auxiliary system*

Another idea to incorporate the initial conditions into the construction of the reduced order model is to balance the homogeneous control system (i.e. for $x_0 = 0$) and the auxiliary system (i.e. for $u = 0$) separately. This route has been taken in [4] for the case of balanced truncation. When the auxiliary system $(A, M, C, 0)$ is balanced, the error term $c_r$ in Theorem 2 simplifies according to

$$\tilde{c}_r = \sqrt{\text{tr}\left(\tilde{C}_2 \Sigma_2 \tilde{C}_2^T\right) + 2\text{tr}\left(\tilde{A}_{12}\tilde{\Sigma}_2 V_2\right)}, \tag{53}$$

where $V = (V_1^T, V_2^T) \in \mathbb{R}^{(r+(n-r)) \times r}$ is the solution of the Sylvester equation

$$\tilde{A}V + V\tilde{A}_1^T = -\tilde{M}\tilde{M}_1^T, \tag{54}$$

and $\tilde{A}$, $\tilde{A}_{12}$, $\tilde{C}_2$, etc. denote the corresponding matrix blocks of the balanced auxiliary system $(A, M, C, 0)$. Note that $\tilde{\Sigma}_2 \neq \Sigma_2$ in general, since the latter is computed by balancing of the original system $(A, B, C, D)$ rather than the auxiliary system $(A, M, C, 0)$.

The simplified error term (53) that follows from Lemma 7.13 and Remark 7.2.3 in [2] has a suggestive interpretation: The first term represents the $H^2$ norm of the truncated subsystem, whereas the second term is twice the inner product between the off-diagonal entries of $\tilde{A}$ and $V_2$ weighted with the small Hankel singular values $\Sigma_2$. It is negligible whenever the balanced auxiliary system has small off-diagonal $\tilde{A}$-matrix entries or when $\Sigma_2$ is negligible. In particular, when $\tilde{M}_2 \approx 0$ in the balanced representation, then $\Sigma_2 \approx 0$ and therefore $\tilde{c}_r \approx 0$.

**Remark 3.** To our knowledge, our work is the only one that consider model reduction by SPA for systems with nonzero initial conditions. In contrast to the closely related balanced truncation approach in [4] that is better suited to approximate the high-frequency response of the system, which also includes possible transients due to the non-zero initial conditions, the SPA aims at the slow, i.e. low-frequency modes. Although the idea of reducing the error due to the initial conditions by balancing the auxiliary system is appealing, this step comes at the additional numerical cost of computing another balancing transformation (including the inversion of $\tilde{A}_{22}$ to compute the Schur complement). This step may be costly, especially if the matrix $M$ has large rank. However, the slow modes targeted by the SPA are less sensitive to the short-time transients which explains why, even without the extra balancing step, the SPA already yields a good approximation, and this is in line with the numerical results shown below that demonstrate that the dynamics is essentially governed by the slow control for various different initial conditions. This may explain that for the numerical examples at hand, the extra balancing step of the auxiliary system typically does not lead to higher model fidelity, even when the matrix $M$ has larger rank than $B$. All our numerical examples do not show a huge impact of $M$ on the slow modes that are approximates by the SPA, and we believe that this is a feature of the SPA and therefore there is no need to include it in the construction of the SPA subspace. Clearly, there might be cases in which the fidelity of the reduced system is very sensitive to the particular $B$ and $M$ matrices, so that it may pay off to compute another balancing transformation for the auxiliary part in order to reduce the approximation error.

## 4. Numerical illustration

We illustrate our theoretical findings with a numerical example and make some comparison with the related approaches [18] and [4]. To this end we consider the CD player system from the SLICOT library [10] with $n = 120$ degrees of freedom. To begin with we compare standard balanced truncation and the singular perturbation approximation of the system with
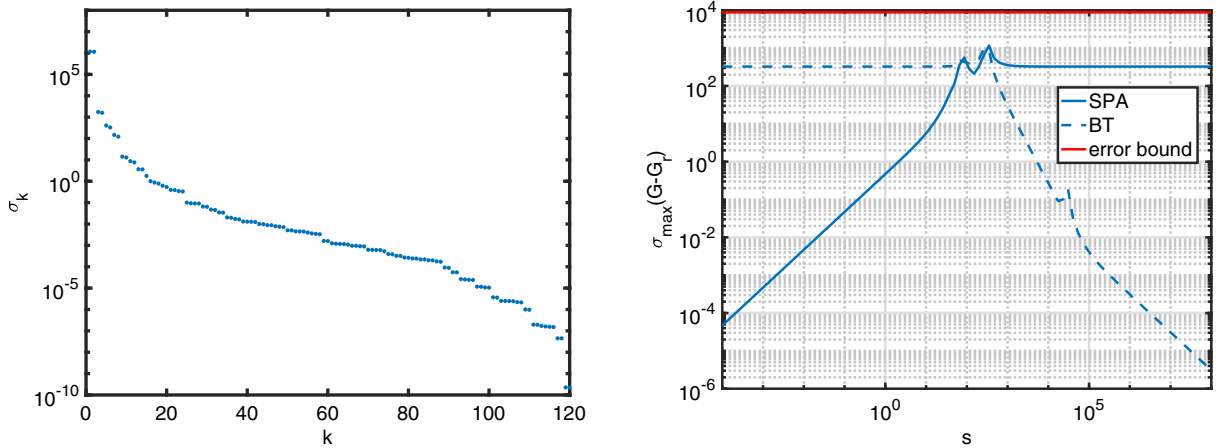
**Fig. 1.** CD player system with zero initial conditions: Hankel singular values (left panel) and frequency plots for singular perturbation approximation (SPA) and balanced truncation.
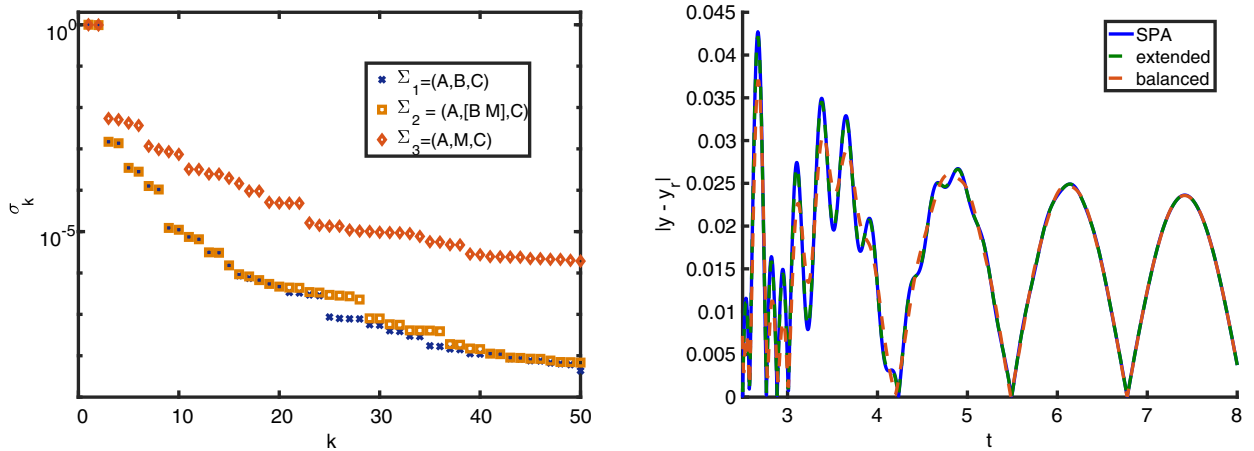


**Fig. 2.** CD player system with nonzero initial conditions: the left panel shows the normalized Hankel singular values for the original system, the extended system and the auxiliary system, the right panel shows the corresponding approximation errors for the dynamics with random Gaussian initial conditions $x_0 = Mz_0$ and control $u = 0$ (see text below for details).

homogeneous initial conditions $x = 0$. The left panel of Fig. 1 shows the Hankel singular values (HSV) of the CD player system. We observe that the low-lying spectrum of HSV displays a significant gap between $\sigma_2$ and $\sigma_3$ where the HSV drop by about three orders of magnitude, so $r = 2$ will be our candidate for the rank of the reduced order system. The right panel of the figure shows the differences between the full and reduced transfer functions for balanced truncation (BT) and the singular perturbation approximation (SPA), and clearly illustrates the "reciprocal" approximation characteristics of the two methods: while the approximation error of the truncated transfer function

$$G_r(s) = C_1 \big( sI - A_{11} \big)^{-1} B_1 \qquad (55)$$

is lowest for high frequencies, the corresponding singular perturbation approximation $\hat{G}_r(s) = G_r(1/s)$ with

$$\hat{G}_r(s) = \hat{C}_1 \big( sI - \hat{A}_{11} \big)^{-1} \hat{B}_1 \qquad (56)$$

is good at approximating the low frequencies, i.e. the slow modes. The higher fidelity of the SPA at lower frequencies is in accordance with the interpretation of the SPA either as an averaging method or as a Padé approximation of the transfer function about the expansion point $s = 0$. By reciprocality, this means that balanced truncation is usually better in capturing the higher frequencies, which can be observed numerically and which corresponds to a Padé approximation at $s = \infty$. Hence we expect the SPA to be the method of choice whenever the slow modes or low frequencies are the relevant ones, whereas model reduction by balanced truncation should be used otherwise, because it does neither require to compute Schur complements nor to compute the full-blown balancing transformation, both of which are numerically tedious when $r$ is small so that $n - r \approx n \gg 1$.

**Table 1**
Absolute and relative $L^2$ norm of $y - y_r$ and error bound for the CD player.

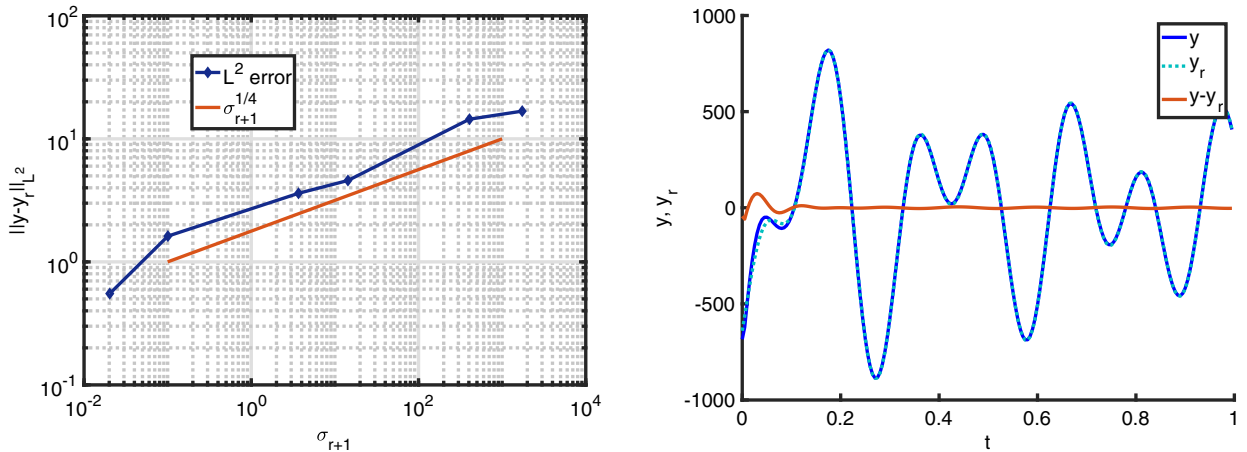| $r$ | Rel. $L^2$ error | Abs. $L^2$ error | Error bound |
|---|---|---|---|
| 2 | 0.0147 | 16.7954 | $9.9553 \cdot 10^3$ |
| 4 | 0.0126 | 14.4464 | $2.438 \cdot 10^3$ |
| 8 | 0.0040 | 4.5866 | 141.7349 |
| 12 | 0.0032 | 3.6086 | 39.2650 |
| 24 | 0.0014 | 1.6194 | 4.0903 |
| 34 | 0.0005 | 0.5523 | 1.1649 |



**Fig. 3.** $L^2$ approximation error as a function of the first negligible HSV $\sigma_{r+1}$ (left panel) and typical output of the CD player system (right panel). Note the mismatch between the original system and its approximant in the initial transient phase until $t \approx 0.1$ due to inconsistent initial conditions (see the text for further explanation).

*Singular perturbation approximation with nonzero initial conditions*

We will now look at the approximation for inhomogeneous initial conditions $x_0 = Mz_0 \neq 0$. To begin with, we set $u = 0$ and study approximations of $\gamma(t) = C \exp(At)M$ by its rank-$r$ approximants $\gamma_r(t)$ for initial conditions that are $\mathcal{O}(1)$ with respect to the characteristic length scale of the system. Specifically, we consider a matrix $M \in \mathbb{R}^{n \times k}$ with $k$ between 1 and $n = 120$ and $z_0 \in \mathbb{R}^k$ being a Gaussian random variable with mean zero and unit covariance. The right panel of Fig. 2 shows the error $|y(t) - y_r(t)|$ for a rank-2 approximation with a typical matrix $M$ of column rank $k = 3$ for three different cases:

1. Case "SPA" shows the approximation for $\tilde{\gamma}_r(t) = \tilde{C}_1 \exp(\tilde{A}_1 t) \tilde{M}_1$ where the coefficients $(\tilde{A}_{11}, \tilde{M}_1, \tilde{C}_1)$ are computed via balancing $(A, B, C)$ according to Theorem 2.
2. Case "extended" refers to $\gamma_r^e(t) = C_1^e \exp(A_{11}^e t) M_1^e$ that is computed from the extended system $(A, (B \,|\, M), C)$ by balanced truncation, following the procedure of [18].
3. Case "balanced" refers to the balanced truncation of the auxiliary system $(A, M, C)$.

In all three cases the relative $L^2$ norm over a time interval [0,10] is about 4%, and different choices of $M$ or $z_0$ yield consistent results, namely: that the three approximations of the semigroup $\gamma(t)$ are very similar, which suggests that in our case the approximation is dominated by the matrix $A$ and the gap in the Hankel singular value spectrum.

For the system at hand, this behavior is remarkably robust, and the corresponding first 10 Hankel singular values that are shown in the left panel of Fig. 2 do not change at all when $k$, the effective dimension of $x_0$, varies between $k = 1$ and $k = n = 120$. The same goes for the relative $L^2$ error of the auxiliary system that shows no significant differences between the different approximants. This behavior can be explained by the strongly dissipative character of the system $(A, \cdot, C)$ that is induced by the eigenvalue of the matrix $A$ with the largest real part and that completely dominates the exponential convergence towards the fixed point 0; it is a well-known fact from the theory of singularly perturbed differential equations (e.g. [28]) that an $\mathcal{O}(1)$ mismatch in the initial conditions between the original and the reduced system will result in a transient initial layer of length $\sqrt{\epsilon}$, in which the approximation error remains $\mathcal{O}(1)$ before it drops to $\mathcal{O}(\epsilon)$, where $\epsilon \approx \sigma_{r+1}/\sigma_r$. Since $\sigma_{r+1}/\sigma_r \approx 10^{-3}$ in our case, it is neither advisable to include an extra balancing step in order to reduce the auxiliary part of the system nor to math the initial conditions. See also the discussion on page 10.

Even though we believe that this is observation is typical, we stress that there may be cases in which the fidelity of the reduced system is very sensitive to the particular $B$ and $M$ matrices, so that it may pay off to compute another balancing transformation for the auxiliary part in order to reduce the overall approximation error; adapting the error bound to this case is straightforward as is described on page 11; cf. also [4].

As the final step, we compute the $L^2$ bound of the difference between the output $y$ of the original system and the output $y_r$ of the SPA (50) with nonzero initial condition and $u \neq 0$ in the time-domain [0,50], and we compare it with the SPA error bound (26) and (27) derived in Theorem 2. We again choose random initial conditions $x_0 = Mz_0$ with $M \in \mathbb{R}^{n \times k}$ having rank $k = 3$. The results are shown in Table 1 for various orders $r$ of the SPA and clearly demonstrate convergence with increasing rank $r$. Note that even though the error bound seems to be rather conservative, given the true $L^2$ error, but it becomes asymptotically tighter as $r$ approaches the system dimension.

For further illustration, the left panel of Fig. 3 shows a typical realization of the full system and the reduced 2-dimensional SPA for a control

$$u(t) = 10^{-4} \exp(0.2t)(\sin(-40t),\ \cos(-40t))^T.$$

The plot gives a visual confirmation of the good fidelity of the SPA and again illustrates the the short initial boundary layer for $t$ between 0 and approximately 0.1 that is caused by the mismatch in the initial conditions.

## 5. Conclusions

In this article we have described a new approach for constructing a reduced order model for linear time-invariant dynamical systems with inhomogeneous initial conditions. We have derived an infinite time-horizon $L^2$ error for the singular perturbation approximation of linear control systems with nonzero initial conditions.

The error bound augments the usual balanced truncation error bound for the SPA of linear systems with zero initial conditions. The extra error contribution comes in form of the $H^2$ norm difference between the corresponding homogeneous (i.e., control-free) flow maps and can be computed by solving a matrix Sylvester equation. That is, computing the error bound comes at little additional numerical cost, compared to computing the balancing transformation, in particular the balancing transformation is computed only for the state space system, without the taking into account the initial condition. Generally, the augmented error bound is relatively conservative, but so is the usual balanced truncation error bound for systems with zero initial conditions, and indeed the error term coming from the initial condition is roughly of the same order of magnitude as the original balanced truncation bound.

## Appendix A. The reciprocal system

In this section we introduce the reciprocal system and discuss some properties of its transfer function. To this end let $x_0 = 0$ in (1). The reciprocal system to (1) is defined as

$$\begin{aligned} \frac{d\hat{x}}{dt} &= \hat{A}\hat{x} + \hat{B}u, \ \hat{x}(t_0) = \hat{x}_0 \\ \hat{y} &= \hat{C}\hat{x} + \hat{D}u \end{aligned} \tag{57}$$

with the coefficients

$$\hat{A} = A^{-1}, \quad \hat{B} = A^{-1}B, \quad \hat{C} = -CA^{-1}, \quad \hat{D} = D - CA^{-1}B \tag{58}$$

and initial condition

$$\hat{x}_0 = 0. \tag{59}$$

The next two auxiliary lemmas (see e.g. [24,34]) explain the idea behind reciprocality. We sketch the proofs for the readers' convenience.

**Lemma 4.** *The system* $(A, B, C, D)$ *is balanced and stable with Hankel singular values* $\Sigma$ *if and only if the reciprocal system* $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ *is balanced and stable with the same Hankel singular values* $\Sigma$.

**Proof.** If (1) is balanced, its coefficients satisfy the pair of Lyapunov equations (2) with $W_c = \Sigma = W_o$ where $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$. Multiplying the first equation in (2) from the left by $A^{-1}$ and from the right by $A^{-T}$, we see that the pair $(\hat{A}, \hat{B})$ solves the same Lyapunov equation as the pair $(A, B)$ for $W_c = \Sigma$. By a similar argument, it follows that the matrix pairs $(\hat{A}, \hat{C})$ and $(A, C)$ solve the same Lyapunov equation for $W_o = \Sigma$. $\square$

**Lemma 5.** *Let* $\hat{G}$ *be the transfer function of the reciprocal system* (57) *with zero initial conditions* $\hat{x}_0 = 0$. *Then*

$$\hat{G}(s) = G(1/s). \tag{60}$$

**Proof.** Since $A$ is invertible, a simple calculation shows that

$$
\begin{aligned}
G(s) &= C(sI - A)^{-1}AA^{-1}B + D \\
&= Cs^{-1}(A^{-1} - s^{-1}I)^{-1}A^{-1}B + D \\
&= -C(s^{-1}I - A^{-1} + A^{-1})(s^{-1}I - A^{-1})^{-1}A^{-1}B + D \\
&= -CA^{-1}(s^{-1}I - A^{-1})^{-1}A^{-1}B + D - CA^{-1}B \\
&= \hat{C}(s^{-1}I - \hat{A})^{-1}\hat{B} + \hat{D}
\end{aligned}
$$

□

As a consequence of the previous lemmas, the reduced system

$$
\begin{aligned}
\frac{d\hat{x}_r}{dt} &= \hat{A}_{11}\hat{x}_r + \hat{B}_1 u, \quad \hat{x}_r(0) = 0 \\
\hat{y}_r &= \hat{C}_1\hat{x}_r + \hat{D}_1 u
\end{aligned}
\tag{61}
$$

with coefficients

$$
\begin{aligned}
\hat{A}_{11} &= \left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)^{-1} \\
\hat{B}_1 &= \left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)^{-1}\left(B_1 - A_{12}A_{22}^{-1}B_2\right) \\
\hat{C}_1 &= -\left(C_1 - C_2A_{22}^{-1}A_{21}\right)\left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)^{-1} \\
\hat{D}_1 &= D - CA^{-1}B,
\end{aligned}
\tag{62}
$$

reciprocal to the SPA (10). Note that the coefficients $(\hat{A}_{11}, \hat{B}_1, \hat{C}_1, \hat{D}_1)$ are simply the truncated version of the system $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ where $\hat{D}_1 = \hat{D}$ and the Schur complement-like expressions for the truncated coefficients follow from the properties of inverses of partitioned matrices (see e.g. [32, Section 9.1.3]). The particular form of $\hat{D}_1$ entails reciprocality between (10) and (61) since

$$
\hat{D}_1 - \hat{C}_1\hat{A}_{11}^{-1}\hat{B}_1 = \hat{D}_1 + CA^{-1}B - C_2A_{22}^{-1}B_2 = D - C_2A_{22}^{-1}B_2,
\tag{63}
$$

as a consequence of which we recover

$$
\tilde{A}_1 = \hat{A}_{11}^{-1}, \quad \tilde{B}_1 = \hat{A}_{11}^{-1}\hat{B}_1, \quad \tilde{C}_1 = \hat{C}_1\hat{A}_{11}^{-1}, \quad \tilde{D}_1 = \hat{D}_1 - \hat{C}_1\hat{A}_{11}^{-1}\hat{B}_1
$$

as the coefficients of the SPA (10).

Another (nontrivial) consequence is that the reciprocality relation entails the known error bound of Theorem 1 for the singular perturbation approximation (10), specifically, for the difference $G - \tilde{G}_r$ (see [24, Theorem 3.2] for further details):

**Corollary 6.** *Under the previous assumptions, it holds that*

$$
\|G - \tilde{G}_r\|_{H^\infty} \le 2\mathrm{tr}(\Sigma_2).
\tag{64}
$$

**Proof.** Using the reciprocality relations with $s^{-1} = -i\omega$ for $s = i\omega$, it follows that

$$
\begin{aligned}
\|G(i\omega) - \tilde{G}_r(i\omega)\|_2 &= \|G(i\omega) - \hat{G}(-i\omega) + \hat{G}(-i\omega) - \hat{G}_r(-i\omega) + \hat{G}_r(-i\omega) - \tilde{G}_r(i\omega)\|_2 \\
&\le \|G(i\omega) - \hat{G}(-i\omega)\|_2 + \|\hat{G}(-i\omega) - \hat{G}_r(-i\omega)\|_2 + \|\hat{G}_r(-i\omega) - \tilde{G}_r(i\omega)\|_2 \\
&= \|\hat{G}(-i\omega) - \hat{G}_r(-i\omega)\|_2.
\end{aligned}
$$

Theorem 1 thus implies that

$$
\|G - \tilde{G}_r\|_{H^\infty} \le \|\hat{G} - \hat{G}_r\|_{H^\infty} \le 2\mathrm{tr}(\Sigma_2).
\tag{65}
$$

□

## References

[1] O. Alvarez, M. Bardi, Viscosity solutions methods for singular perturbations in deterministic and stochastic control, SIAM J. Control Optim. 40 (4) (2002) 1159–1188.

[2] A.C. Antoulas, Approximation of Large-scale Dynamical Systems, Advances in design and control, Society for Industrial and Applied Mathematics, 2005.

[3] U. Baur, P. Benner, L. Feng, Model order reduction for linear and nonlinear systems: a system-theoretic perspective, Arch. Comput. Meth. Eng. 21 (4) (2014) 331–358.

[4] C. Beattie, S. Gugercin, V. Mehrmann, Model reduction for systems with inhomogeneous initial conditions, Syst. Control Lett. 99 (2017) 99–106.

[5] P. Benner, T. Breiten, C. Hartmann, B. Schmidt, Model reduction of controlled Fokker–Planck and Liouville–von Neumann equations, SIAM J. Appl. Dyn Syst. (2017). available from https://arxiv.org/abs/1706.09882.

[6] P. Benner, M. Embree, R.B. Lehoucq, C. Kelley, A mathematical biography of Danny C. Sorensen, Linear Algebra Appl. 415 (2006) 231–234.

[7] A. Bensoussan, G. Blankenship, Singular perturbations in stochastic control, in: Singular Perturbations and Asymptotic Analysis in Control Systems, Springer, 1987, pp. 171–260.

[8] R. Buchholz, H. Engel, E. Kammann, F. Tröltzsch, On the optimal control of the Schlögl-model, Comput. Optim. Appl. 56 (1) (2013) 153–185.

[9] A.L. Cervantes, O.E. Agamennoni, J.L. Figueroa, A nonlinear model predictive control system based on wiener piecewise linear models, J. Process Control 13 (7) (2003) 655–666.

[10] Y. Chahlaoui, P. Van Dooren, Benchmark examples for model reduction of linear time invariant dynamical systems, in: P. Benner, D.C. Sorensen, V. Mehrmann (Eds.), Dimension Reduction of Large-Scale Systems, Lecture Notes in Computer Science and Engineering, 45, Springer, Berlin, 2005, pp. 379–392.

[11] A. Daraghmeh, Model order reduction of linear control systems: comparison of balanced truncation and singular perturbation approximation with application to optimal control, Ph.D. Thesis, Berlin Freie Universität, Germany, 2016.

[12] L.C. Evans, The perturbed test function method for viscosity solutions of nonlinear PDE, Proc. R. Soc. Edinb. Sect. A Math. 111 (3-4) (1989) 359–375.

[13] D. Galbally, K. Fidkowski, K. Willcox, O. Ghattas, Non-linear model reduction for uncertainty quantification in large-scale inverse problems, Int. J. Numer. Meth. Eng. 81 (12) (2010) 1581–1608.

[14] K. Glover, All optimal hankel-norm approximations of linear multivariable systems and their $l_\infty$-error bounds, Int. J. Control 39 (6) (1984) 1115–1193.

[15] D. Gratton, K. Willcox, Reduced-order, trajectory piecewise-linear models for nonlinear computational fluid dynamics, in: Proceedings of the 34th AIAA Fluid Dynamics Conference and Exhibit, 2004-2329, 2004, pp. 1–12.

[16] C. Hartmann, B. Schäfer-Bung, A. Thöns-Zueva, Balanced averaging of bilinear systems with applications to stochastic control, SIAM J. Control Optim. 51 (2013) 2356–2378, doi:10.1137/100796844.

[17] C. Hartmann, V.-M. Vulcanov, C. Schütte, Balanced truncation of linear second-order systems: a hamiltonian approach, Multiscale Model. Simul. 8 (2010) 1348–1367.

[18] M. Heinkenschloss, T. Reis, A. Antoulas, Balanced truncation model reduction for systems with inhomogeneous initial conditions, Automatica 47 (2011) 559–564, doi:10.1016/j.automatica.2010.12.002.

[19] P.V. Kokotovic, Applications of singular perturbation techniques to control problems, SIAM Rev. 26 (4) (1984) 501–550.

[20] P.V. Kokotovic, R.E. O'Malley, P. Sannuti, Singular perturbations and order reduction in control theory–an overview, Automatica 12 (2) (1976) 123–132.

[21] P.V. Kokotovic, R.A. Yackel, Singular perturbation of linear regulators: basic theorems, IEEE Trans. Autom. Control 17 (1) (1972) 29–37.

[22] P.-L. Lions, P.E. Souganidis, Correctors for the homogenization of Hamilton–Jacobi equations in the stationary ergodic setting, Commun. Pure Appl. Math. 56 (10) (2003) 1501–1524.

[23] Y. Liu, B.D. Anderson, Singular perturbation approximation of balanced systems, Int. J. Control 50 (4) (1989) 1379–1405.

[24] Y. Liu, B.D.O. Anderson, Singular perturbation approximation of balanced systems, Int. J. Control 50 (1989) 1379–1405, doi:10.1080/00207178908953437.

[25] K. Mohaghegh, M. Striebel, E. Ter Maten, R. Pulch, Nonlinear model order reduction based on trajectory piecewise linear approach: comparing different linear cores, in: J. Roos, L. Costa (Eds.), Scientific Computing in Electrical Engineering SCEE 2008, Mathematics in Industry, 14, Springer, Berlin, Heidelberg, 2010, pp. 563–570.

[26] B. Moore, Principal component analysis in linear systems: controllability, observability, and model reduction, IEEE Trans. Autom. Control 26 (1) (1981) 17–32.

[27] G. Muscato, G. Nunnari, L. Fortuna, Singular perturbation approximation of bounded real balanced and stochastically balanced transfer matrices, Int. J. Control 66 (2) (1997) 253–270.

[28] R. O'Malley, Singular Perturbation Methods for Ordinary Differential Equations, Springer, New York, 1991.

[29] R. O'Malley Jr, The singularly perturbed linear state regulator problem, SIAM J. Control 10 (3) (1972) 399–413.

[30] R.E. O'Malley Jr, On two methods of solution for a singularly perturbed linear state regulator problem, SIAM Rev. 17 (1) (1975) 16–37.

[31] L. Özkan, M.V. Kothare, C. Georgakis, Model predictive control of nonlinear systems using piecewise linear models, Comput. Chem. Eng. 24 (2) (2000) 793–799.

[32] K.B. Petersen, M.S. Pedersen, The Matrix Cookbook, Technical University of Denmark, 2012.

[33] M. Rewienski, J. White, A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices, IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. 22 (2) (2003) 155–170.

[34] R. Saragih, Fatmawati, Singular perturbation approximation of balanced infinite-dimentional system, Int. J. Control Autom. 6 (2013) 409–420.

[35] S. Skogestad, I. Postlethwaite, Multivariable Feedback Control: Analysis and Design, 2, Wiley, New York, 2007.

[36] K. Zhou, J.C. Doyle, K. Glover, Robust and Optimal Control, Prentice Hall, 1998.