

# A subspace approach to balanced truncation for model reduction of nonlinear control systems

Sanjay Lall<sup>1,\*</sup>,†, Jerrold E. Marsden<sup>2</sup> and Sonja Glavaški<sup>3</sup>

<sup>1</sup>*Department of Aeronautics and Astronautics, Stanford University, Stanford, CA 94305-4035, U.S.A.*

<sup>2</sup>*Control and Dynamical Systems 107-81, California Institute of Technology, Pasadena, CA 91125, U.S.A.*

<sup>3</sup>*Honeywell Technology Center, 3660 Technology Drive, Minneapolis, MN 55418, U.S.A.*

## SUMMARY

In this paper, we introduce a new method of model reduction for nonlinear control systems. Our approach is to construct an approximately balanced realization. The method requires only standard matrix computations, and we show that when it is applied to linear systems it results in the usual balanced truncation. For nonlinear systems, the method makes use of data from either simulation or experiment to identify the dynamics relevant to the input–output map of the system. An important feature of this approach is that the resulting reduced-order model is nonlinear, and has inputs and outputs suitable for control. We perform an example reduction for a nonlinear mechanical system. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: balanced truncation; model reduction; Karhunen–Loève expansion; total least-squares; Hankel norm

## 1. INTRODUCTION

One of the difficulties in designing controllers for complex physical systems is in the problem of *modelling*. Very often, we have mathematical models which are highly nonlinear partial differential equations, for which control design techniques are little understood.

An example is given by formation flight; accurate models are known for the fluid dynamics, and for the aircraft motion, which together can be accurately simulated. However, on paper, the equations of motion consist of partial and ordinary differential equations coupled via their boundary conditions, a formulation which offers little to the control designer. It is

---

\*Correspondence to: Sanjay Lall, Department of Astronautics and Aeronautics, Stanford University, Stanford, CA 94305-4035, U.S.A.

†E-mail: lall@stanford.edu

Contract/grant sponsor: AFOSR MURI; contract/grant number: F49620-96-1-0471

Contract/grant sponsor: EPRI/DOD CIN/SI; contract/grant number: W08333-06

Contract/grant sponsor: NSF/DARPA OPAAL; contract/grant number: DMS-9874082

*Received 7 September 2000*

*Revised 7 January 2001*

*Accepted 2 July 2001*

therefore a problem of considerable interest to construct explicit low-order models; once a control design has been performed using such a low-order model, it can be tested by simulating it in feedback with the full high-order simulation.

In this paper, we present a new method of model reduction for nonlinear control systems. The goal is to develop an intuitively motivated and systematic procedure for construction of low-order models for complex high-dimensional nonlinear systems. The focus is on preserving those features of the dynamics which are most relevant to the control design, in a similar way to that in which standard methods for model reduction of linear systems, such as balanced truncation, preserve the most controllable and observable states.

This is one of the main differences between modelling in the sense used by control systems engineers, and modelling as used for simulation or mechanical design. In the latter case, one is typically interested in the behaviour of an *autonomous* dynamical system, that is one with no external inputs driving the system. However, in the case of controlled systems, we are concerned with modelling the relationship between the system behaviour and the system inputs and outputs, or actuators and sensors. *Linear* input–output systems enjoy several such input–output model reduction techniques, including balanced truncation and optimal Hankel,  $H_\infty$  and  $H_2$  norm methods. In general, these methods are difficult to apply to complex high-dimensional nonlinear systems.

The approach presented here makes use of data from either simulation or experiment to identify those states of the system which are not affected by actuators, and which most affect the sensors. We construct an *empirical balanced realization* for nonlinear systems, which coincides with the usual balanced realization for linear systems. A Galerkin projection is then applied to the balanced realization to construct a low-dimensional nonlinear model. This reduced-order model may then be used for efficient controller synthesis, design optimization, and real-time control algorithms, where having a reduced-order model brings significant computational complexity benefits.

An example is given by linear-parameter-varying (LPV) control design [1], where the required optimization for control synthesis involves solving a linear matrix inequality feasibility problem where the number of variables grows as the square of the state-dimension of the system. A significant reduction in overall computational effort required is achieved by constructing a reduced-order model via simulation of the full-order model, and then using this reduced-order model for the control synthesis.

### 1.1. Prior work

Two of the most well-known methods of model reduction used in control are the Karhunen–Loève decomposition and the method of balanced truncation. The approach in this paper combines the ideas and methodology from both of these.

#### 1.1.1. The Karhunen–Loève expansion

The Karhunen–Loève expansion is a method of least-squares approximation which goes back to Karl Pearson [2]. The method is known in the literature by several names, including principal component analysis, proper orthogonal decomposition (POD), factor analysis, and total-least-squares estimation.

The use of the Karhunen–Loève expansion for model reduction was pioneered by Lumley [3], where it was used to model complex fluid flows. This has found wide applicability and led

to a method of understanding the important dynamical features seen in the fluid [4–6]. The KLE has also been used for model reduction of solids and structures [7, 8], for analysis of chaotic systems [9] and for analysis of systems with dynamical symmetries [10, 11]. For control systems, the KLE has also been used to design controllers for PDEs [12, 13].

### 1.1.2. *Balanced truncation*

The method of balanced truncation for model reduction of linear systems was proposed by Moore [14] in the context of realization theory, and is now a well-developed method of model reduction that appears in standard textbooks [15]. For linear systems, the approach requires only matrix computations, and has been very successfully used in control design. *A priori* error bounds in the induced 2-norm are known for the error between the original and the reduced system [16, 17].

The method of balanced truncation has been extensively developed for nonlinear systems by Scherpen [18–20], based on energy functions. A closed-loop approach is also presented in Reference [21], and balancing methods for bilinear systems also been developed [22–24]. Computational approaches for these methods have been investigated in Reference [25], and have been compared with the Karhunen–Loève method in Reference [26].

Another approach to balanced model reduction of nonlinear systems has been to make use of linear parameter-varying (LPV) systems theory, where the linear fractional transformation is used to construct uncertain multidimensional models capturing the nonlinear dynamics [27–30].

Other approaches have also been used for nonlinear model reduction; for modular interconnected nonlinear systems, methods have been proposed in Reference [31], and for chemical plants, direct optimization techniques have been used [32]. Along trajectories, time-varying methods have been developed; see Reference [33] for details.

### 1.1.3. *Our approach*

Our approach in this paper has been to combine the features of the Karhunen–Loève expansion and balanced truncation, thereby developing a new method of model reduction for nonlinear systems. In particular, this approach is *data-based*, making use of data from simulation or experiment in the construction of the reduced-order model. In fact, Moore also proposed using principal components analysis for nonlinear systems [34], although this was not directly connected to the balanced truncation approach he had proposed in Reference [14].

By making use of data, we avoid some of the computational difficulties involved in computing energy functions for exact balanced truncation of nonlinear system [18]. The method provides an immediate computable procedure for model reduction of nonlinear input–output systems, requiring only standard matrix computations. When applied to linear input–output systems, it results in the usual balanced truncation of the system, with the accompanying *a priori* error bounds in the  $H_\infty$  and the Hankel norms.

Note that the method of balanced truncation is not optimal, in that in general, it does not achieve the minimal possible error in any known norm even for linear systems. However, it is a well-used method in control, which is both intuitively motivated by realization theory and known to perform well in engineering practice. The motivation and intended use of the methods in this paper is to construct deliberately approximate reduced-order models motivated by the ideas of balanced truncation.

The philosophy in this paper is also similar to that of identification, where data are used for construction of models. In particular, the method of subspace identification [35] is relevant here.

In this sense, our approach for model reduction can alternatively be viewed as a new approach to *identification*, and a way of incorporating nonlinear *a priori* mathematical models for the physics into the identification process.

This paper is a longer version of the conference paper [36]. This approach and a variation of it have also been used for modelling of a continuously stirred tank reactor (CSTR) in Reference [37]. We conclude the paper with an example to illustrate the application of the method to a nonlinear mechanical system.

## 2. MODEL REDUCTION FOR AUTONOMOUS SYSTEMS

In this section, we give an outline of existing techniques using the Karhunen–Loève decomposition for model reduction of nonlinear autonomous systems. The dynamics of such a physical system are often modelled by a high-dimensional nonlinear differential equation, of the form

$$\dot{x}(t) = f(x(t)) \quad (1)$$

This equation describes the system behaviour as the evolution of the state  $x(t)$  in a high-dimensional state space  $\mathbb{R}^n$ . The central idea used when applying the Karhunen–Loève expansion for model reduction is to search for a low-dimensional affine subspace of the state space, in which the dynamics of interest of the original system are contained. In general, a subspace containing the interesting state trajectories may not have low (or finite) dimension, and the decomposition provides a way of finding the best approximating low-dimensional subspace.

Once such a subspace is found, a Galerkin projection can be applied to project the dynamics onto it, so that the high-dimensional system is approximated by a small number of nonlinear ordinary differential equations.

Once an appropriate measure is assigned to the space of trajectories of the system, the problem of finding the subspace can be formulated as a 2-norm optimization problem. This has the advantage of requiring only linear matrix computations, despite its usefulness in application to nonlinear systems. The optimization need not be performed analytically; instead data from experiment or simulation can be used directly.

The method makes essential use of *empirical data*, taken either from experiments or from numerical simulation, consisting of sampled measurements  $\{x^{(1)}, \dots, x^{(N)}\}$  of  $x(t)$ . The next step is to perform a principal component analysis of this data, to find how well it may be approximated by projection onto a  $k$ -dimensional subspace of the original  $n$ -dimensional state space.

### 2.1. The Karhunen–Loève decomposition

The Karhunen–Loève decomposition provides a method for finding this best approximating subspace. We can characterize a subspace  $S \subset \mathbb{R}^n$  by the projection operator  $Q$  mapping  $\mathbb{R}^n$  onto  $S$ . The order of approximation is fixed at  $k$ , and we would like to find  $Q$  to minimize

$$H(Q) = \sum_{i=1}^N \|x^{(i)} - Qx^{(i)}\|_2^2 \quad (2)$$

the total squared perpendicular distance of the points from the  $k$ -plane. The following result is standard.

*Theorem 1* (Total least squares)

Let  $R$  be the correlation matrix of the data, defined by

$$R := \sum_{i=1}^N x^{(i)} x^{(i)*} \tag{3}$$

and let  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$  be the ordered eigenvalues of  $R$ . Then

$$\min_Q H(Q) = \sum_{i=n-k+1}^n \lambda_i \tag{4}$$

where the minimum is over all rank  $k$  projections  $Q$ .

In general,  $R$  may not have rank  $n$ , if the given data lie within a strict subspace of  $\mathbb{R}^n$ . Let  $s = \text{rank } R$ , and let  $\phi_1, \phi_2, \dots, \phi_s$  be orthonormal eigenvectors of  $R$ , corresponding to the non-zero  $\lambda_i$ . Each  $x^{(i)}$  can be written as

$$x^{(i)} = \sum_{j=1}^s a_{ij} \phi_j \tag{5}$$

where  $a_{ij} = \langle x^{(i)}, \phi_j \rangle$ , and  $\langle \phi_i, \phi_j \rangle = \delta_{ij}$ . The optimal  $k$ -dimensional subspace approximant is given by

$$\hat{x}^{(i)} = \sum_{j=1}^k a_{ij} \phi_j \tag{6}$$

Denote by  $P$  the  $k \times n$  matrix whose rows are  $\phi_1, \dots, \phi_k$ , so that  $PP^* = I$ . The projected approximant to  $x$  is given by  $P^*Px \in S$ , and  $y = Px$  is a representation in terms of the new coordinates  $\phi_i$  on  $S$ . This subspace approximant is then optimal, in the sense that the total ‘energy’ (2-norm) in the subspace is given by

$$\sum_{i=1}^N \|Px^{(i)}\|_2^2 = \sum_{j=1}^k \lambda_j \tag{7}$$

and this is the maximum achievable by any  $k$ -plane.

### 2.2. Model reduction

The above procedure finds the optimal subspace; to allow affine variation we make use of the fact that the optimal *affine* subspace passes through the mean of the data. Hence, we construct the correlation matrix as

$$R_{ij} = \sum_{p=1}^N (x_i^{(p)} - \bar{x}_i)(x_j^{(p)} - \bar{x}_j) \tag{8}$$

where  $\bar{x} = 1/N \sum_{i=1}^N x^{(i)}$  is the mean of the data. The eigenvalues of  $R$  now provide us with information as to how close an approximation of the data is provided by a  $k$ -dimensional subspace; the goal is to choose  $k$  such that the fraction of the total ‘energy’ in the subspace

$$\sum_{i=1}^k \lambda_i / \sum_{i=1}^n \lambda_i \tag{9}$$

is close to one, yet  $k$  is sufficiently small. Clearly, this will not always be possible, with models which are better approximated by low-dimensional systems having relatively few large eigenvalues.

We now subtract the mean  $\bar{x}$  from the dynamic equation, defining  $e = x - \bar{x}$  to arrive at  $\dot{e}(t) = f(e(t) + \bar{x})$ . We would like to approximate  $e(t)$  in this equation by  $P^*Pe(t) \in S$ .

### 2.3. Galerkin projection

The Galerkin projection has been used extensively to construct numerical solutions to partial differential equations. It can also be used to construct lower-order *mathematical models* of a given dynamical system. The idea is to replace the given dynamics by an associated dynamics on a  $k$ -dimensional subspace  $S \subset \mathbb{R}^n$  of the original state space, by projecting the vector field onto the tangent space of  $S$ . Using the coordinates defined previously, the resulting reduced-order approximation is given by

$$\dot{y}(t) = Pf(P^*y(t) + \bar{x}) \quad (10)$$

The Karhunen–Loève method therefore projects the dynamics onto the sub-space containing most of the ‘statistical energy’ of the system. If we keep all of the eigenvectors corresponding to non-zero eigenvalues, then this subspace will contain all of the dynamics seen in the experiment. In general, we would expect that the more eigenvectors we keep, the better approximation we will obtain.

Clearly, the subspace obtained by the Karhunen–Loève method depends upon the inner product on the Euclidean space in which the state space is embedded; physical meaning is an important indicator of the appropriate choice of inner product. Since the 2-norm of the states is maximized, an inner product for which this norm corresponds to an approximation of the physical energy in the system may be an advantage.

Computational application of this method requires only standard matrix computations, despite its application to nonlinear systems. Also, the method has the feature that it separates the model reduction into two parts; that of finding a suitable subspace and that of performing the projection. This lends physical intuition to the procedure, and allows alternative projection schemes to be used.

We can expect such a procedure to work well for model reduction of the system within some given region of state space, and it is within such a specific region that data should be collected.

For control, a severe limitation is that the input–output behaviour of the system is so far not taken into account. In this paper, we combine the data-based approach of the Karhunen–Loève reduction procedure with the input–output focus of balanced truncation, and we now proceed with this.

## 3. MODEL REDUCTION OF CONTROLLED SYSTEMS

We now turn to the main problem addressed in this paper; that of model reducing nonlinear input–output systems of the form

$$\begin{aligned} \dot{x}(t) &= f(x(t), w(t)) \\ z(t) &= h(x(t)) \end{aligned} \quad (11)$$

Here  $x(t) \in \mathbb{R}^n$  is the state of the system,  $w(t) \in \mathbb{R}^p$ , and  $z(t) \in \mathbb{R}^q$ . The function  $w$  is regarded as an input signal to the system, and the function  $z$  as an output signal. The goal of *model reduction* is to construct another nonlinear system of differential equations

$$\begin{aligned} \dot{q}(t) &= \hat{f}(q(t), w(t)) \\ z(t) &= \hat{h}(q(t)) \end{aligned} \tag{12}$$

where  $q(t) \in \mathbb{R}^k$ , and  $k < n$ , such that the *input-output behaviour* of the two systems is similar, for states in some specific region of the state space. That is, for a class of inputs  $w$ , we would like the outputs  $z$  of the two systems to be close.

In this section, we develop an approach for model reduction of systems of form (11), taking explicit account of the input–output connection of the system, and generalizing standard methods from linear model reduction theory.

We construct a new method, based on the approach used for autonomous systems, but taking explicit account of the input–output connection of the system. The construction is designed to generalize standard methods from linear model reduction theory; before proceeding further, we need some preliminary definitions from the theory of linear systems.

### 3.1. Linear systems

For linear systems, Equations (11) become

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bw(t) \\ z(t) &= Cx(t) \end{aligned} \tag{13}$$

where  $x(t) \in \mathbb{R}^n$ , and  $A$ ,  $B$ , and  $C$  are matrices of appropriate dimension. The linear system is called *stable* if the eigenvalues of  $A$  all have real part strictly negative.

Suppose the system in Equation (13) is stable. Then, for  $u \in L_2(-\infty, 0]$  and denoting  $u_-(t) = u(-t)$ , the state at time zero  $x(0) = x_0$  is given by

$$x_0 = \int_0^\infty e^{As} B u_-(s) ds \tag{14}$$

This defines the *controllability operator*,  $\mathcal{C} : L_2[0, \infty) \rightarrow \mathbb{R}^n$  by  $x_0 = \mathcal{C}u_-$ . The following standard results can be found in Reference [38].

#### Lemma 2

Write  $Y := \mathcal{C}\mathcal{C}^*$ . Then  $Y$  is the smallest semipositive solution to the Lyapunov equation

$$AY + YA^* + BB^* = 0 \tag{15}$$

The system is called *controllable* if  $\text{Im } \mathcal{C} = \mathbb{R}^n$ , in which case  $Y > 0$  and Equation (15) has a unique solution. The matrix  $Y$  is called the *controllability gramian*.

As is standard, there is a similar notion for the output. Denote the future output by  $y_+ \in L_2[0, \infty)$ . Then, define the *observability operator*,  $\mathcal{O} : \mathbb{R}^n \rightarrow L_2[0, \infty)$  by  $y_+ = \mathcal{O}x_0$ , and hence  $\mathcal{O}x_0 = Ce^{At}x_0$ .

#### Lemma 3

Write  $X := \mathcal{O}^*\mathcal{O}$ . Then  $X$  is the smallest semipositive solution to the Lyapunov equation

$$A^*X + XA + C^*C = 0 \tag{16}$$

The system is said to be *observable* if  $\ker \mathcal{O} = \{0\}$ , in which case  $X > 0$  and Equation (16) has a unique solution. The matrix  $X$  is called the *observability gramian*.

Both  $X$  and  $Y$  are  $n \times n$  matrices, and they are given by the following integral formulae:

$$Y = \int_0^{\infty} e^{At} B B^* e^{A^*t} dt$$

$$X = \int_0^{\infty} e^{A^*t} C^* C e^{At} dt$$
(17)

### 3.2. Gramians and principal component analysis

The method of principal component analysis relies on the use of data to construct the correlation matrix, with the underlying assumption that these data are collected from ‘typical’ system trajectories. Implicit in this analysis is that the trajectories from which data are sampled are parametrized by  $x_0$ , the initial state of the system.

Since we are now considering a controlled system with inputs, we can instead make the assumption that the initial state of the system is zero, and parametrize the trajectories for principal component analysis with respect to the system input  $u$ .

For theoretical purposes, we are not restricted to samples, and can simply construct the correlation matrix using the integral

$$R = \int_0^{\infty} (x(t) - \bar{x})(x(t) - \bar{x})^* dt$$
(18)

where  $x(t)$  is the state of the system at time  $t$ , and  $\bar{x}$  is the mean state.

Let  $\mathcal{T}^n$  be a set of  $r$  orthogonal  $n \times n$  matrices,  $\{T_1, \dots, T_r\}$ , and let  $\mathcal{M}$  be a set of  $s$  positive constants,  $\{c_1, \dots, c_s\}$ . Further define

$$\mathcal{E}^n = \{e_1, \dots, e_n; \text{standard unit vectors in } \mathbb{R}^n\}$$
(19)

Given a function  $u \in L_1$ , define the mean  $\bar{u}$  by

$$\bar{u} := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(t) dt$$
(20)

We make the standing assumptions that  $x, z \in L_1$  and  $z \in L_2$ . For given initial conditions, these assumptions are satisfied for exponentially stable systems.

#### Definition 4

Let  $\mathcal{T}^p, \mathcal{E}^p$  and  $\mathcal{M}$  be given sets as described above. For system (11), define the *empirical controllability gramian*  $\hat{Y}$  by

$$\hat{Y} = \sum_{l=1}^r \sum_{m=1}^s \sum_{i=1}^p \frac{1}{rsc_m^2} \int_0^{\infty} \Phi^{ilm}(t) dt$$
(21)

where  $\Phi^{ilm}(t) \in \mathbb{R}$  is given by

$$\Phi^{ilm}(t) := (x^{ilm}(t) - \bar{x}^{ilm})(x^{ilm}(t) - \bar{x}^{ilm})^*$$
(22)

and  $x^{ilm}(t)$  is the state of system (11) corresponding to the impulsive input  $w(t) = c_m T_1 e_i \delta(t)$ .



*Lemma 5*

For any sets  $\mathcal{T}^p$  and  $\mathcal{M}$ , the empirical controllability gramian  $\hat{Y}$  of the stable linear system  $\dot{x}(t) = Ax(t) + Bw(t)$  is equal to the usual controllability gramian  $Y$ .

*Proof.* For the linear system,

$$\begin{aligned} \Phi^{ilm}(t) &= c_m^2(e^{At}BT_l e_i)(e^{At}BT_l e_i)^* \\ &= c_m^2 e^{At}BT_l e_i e_i^* T_l^* B^* e^{A^*t} \end{aligned} \tag{23}$$

hence

$$\begin{aligned} \hat{Y} &= \int_0^\infty \sum_{l=1}^r \sum_{m=1}^s \sum_{i=1}^p \frac{1}{rs} e^{At}BT_l e_i e_i^* T_l^* B^* e^{A^*t} dt \\ &= \int_0^\infty \sum_{m=1}^s \frac{1}{s} e^{At}BB^* e^{A^*t} dt = Y \end{aligned} \tag{24}$$

which is the desired result. □

A variant of this construction was shown by Moore [14]; for impulsive inputs distributed in the above sense on the unit ball in  $\mathbb{R}^p$ , the Karhunen–Loève decomposition of the states leads to a construction of the controllability gramian for linear systems.

The empirical controllability gramian is a computable generalization of this to nonlinear systems; it has the property that the eigenvectors of  $\hat{Y}$  corresponding to non-zero eigenvalues span a subspace  $O \subset \mathbb{R}^n$  which contains the set of states reachable using the chosen initial impulsive inputs.

We are therefore led to a methodology for model reduction motivated by a standard idea from realization theory; truncate those states corresponding to small eigenvalues of  $\hat{Y}$ . In the context of nonlinear systems, one might perform a Galerkin projection onto the subspace spanned by the eigenvectors corresponding to the largest eigenvalues.

However, for the controlled systems, simply studying the input-state behaviour is not enough. The next definition is the analogue of the previous one for the output behaviour.

*Definition 6*

Let  $\mathcal{T}^n$ ,  $\mathcal{E}^n$ , and  $\mathcal{M}$  be given sets as described above. For system (11), define the *empirical observability gramian*  $\hat{X}$  by

$$\hat{X} = \sum_{l=1}^r \sum_{m=1}^s \frac{1}{rsc_m^2} \int_0^\infty T_l \Psi^{lm}(t) T_l^* dt \tag{25}$$

where  $\Psi^{lm}(t) \in \mathbb{R}^{n \times n}$  is given by

$$\Psi_{ij}^{lm}(t) := (z^{ilm}(t) - \bar{z}^{ilm})^* (z^{jlm}(t) - \bar{z}^{jlm}) \tag{26}$$

and  $z^{ilm}(t)$  is the output of system (11) corresponding to the initial condition  $x_0 = c_m T_l e_i$  with input  $w = 0$ .

*Lemma 7*

For any non-empty sets  $\mathcal{T}^n$  and  $\mathcal{M}$ , the empirical observability gramian  $\hat{X}$  of the stable linear system  $\dot{x}(t) = Ax(t) + Bw(t)$ ,  $z(t) = Cx(t)$  is equal to the usual observability gramian  $X$ .

*Proof.* For the linear system,

$$\begin{aligned}\Psi_{ij}^{lm}(t) &= c_m^2 (C e^{At} T_l e_i)^* (C e^{At} T_l e_j) \\ &= c_m^2 e_i^* T_l^* e^{A^* t} C^* C e^{A t} T_l e_j\end{aligned}\quad (27)$$

hence

$$\Psi^{lm}(t) = c_m^2 T_l^* e^{A^* t} C^* C e^{A t} T_l \quad (28)$$

and

$$\hat{X} = \sum_{m=1}^s \frac{1}{s} \int_0^\infty e^{A^* t} C^* C e^{A t} dt = X \quad (29)$$

which is the desired result.  $\square$

We now have the tools we need for empirical analysis of the input–output behaviour of the nonlinear system. Rather than searching for exact controllability and observability submanifolds within the state space, our approach is to search for subspaces which approximate these manifolds. The advantage of this approach lies in the computation; all that is required is the solution of standard linear matrix eigenvalue problems.

Unlike in the autonomous case, we now have *two* important subspaces of the state space, and their corresponding eigenvalues. We can proceed in the same manner as for linear systems, and make use of the ideas of balanced realization theory to decide on which subspace to project. We therefore return to a description of linear systems theory.

Corresponding to a linear system there is an associated Hankel operator  $\Gamma$  which maps past inputs to future outputs. The singular values of the Hankel operator are the eigenvalues of  $\mathcal{O}\mathcal{C}\mathcal{C}^*\mathcal{O}^*$ , which are the same as those of the  $n \times n$  matrix  $\mathcal{O}^*\mathcal{O}\mathcal{C}\mathcal{C}^* = XY$ . The *balanced realization* gives a way to find which particular states correspond to which Hankel singular value. We may change state co-ordinates via any non-singular linear transformation  $T$  without affecting the input–output behaviour, in which case the above Lyapunov Equations (15) and (16) imply that the gramians transform correspondingly according to  $Y \rightarrow T Y T^*$  and  $X \rightarrow T^{-1} X T^{-1}$ . For linear systems, the choice of  $T$  does not affect the Hankel singular values. A realization  $(A, B, C)$  is called *balanced* if the controllability gramian  $Y$  and the observability gramian  $X$  are equal and diagonal. The Hankel singular values correspond to states through which the input is transmitted to the output. The *balanced realization* gives a way to find which particular states correspond to which Hankel singular value.

The Hankel singular values  $\sigma_i$  then indicate the importance of the corresponding state in the balanced realization in the transfer of energy from past inputs to future outputs. This leads to a method of model reduction known as *balanced truncation*, introduced by Moore [14] in the context of realization theory. The procedure is to truncate those states from the balanced realization corresponding to small Hankel singular values  $\sigma_i$ . If the states are ordered according to decreasing singular value, this is equivalent to applying a Galerkin projection to the balanced realization, where  $P = [I \ 0]$ .

### 3.3. Empirical balanced truncation

The empirical gramians give a quantitative method for deciding upon the importance of particular subspaces of the state space, with respect to the inputs and outputs of the system. We

propose to use these for model reduction of nonlinear systems in the same way as for linear systems; find a linear change of coordinates such that the empirical gramians are balanced, and perform a Galerkin projection onto the states corresponding to the largest eigenvalues of  $\hat{X}\hat{Y}$ .

Since for linear systems, the empirical gramians are exactly the usual gramians, this method is exactly balanced truncation when it is applied to a linear system. When applied to a nonlinear system, it requires only matrix computations, and results in a new nonlinear model.

Let  $T$  be the change of co-ordinates such that the system is balanced; that is  $TYT^* = T^{-1}XT^{-1} = \Sigma$ , and let  $P = [I \ 0]$  be the  $k \times n$  projection matrix. Applying the previous analysis leads to a reduced-order model given by

$$\begin{aligned}\dot{y}(t) &= PTf(T^{-1}P^*y(t), w(t)) \\ z(t) &= h(T^{-1}P^*y(t))\end{aligned}\tag{30}$$

This is the final reduced-order model. This *empirical balanced truncation* gives a reduced-order model which takes into account the input–output behaviour and is directly computable from data.

For linear systems, the Hankel singular values are unaffected by co-ordinate changes, even though the Gramian matrices themselves are not coordinate invariant. For nonlinear systems, this property no longer holds.

#### 4. COMPUTATION

We can directly apply Definitions 4 and 6 to nonlinear systems for construction of the empirical gramians from data. The data may be taken either from simulation or from experiment. All that is needed is numerical approximation of the above integrals, using sampled-data from both  $x(t)$  and  $z(t)$ . It is also possible to take advantage of knowledge of  $f$  to calculate  $\dot{x}(t)$  in this computation to achieve higher-order accuracy, and increase the allowable sample time.

To perform the required experiments, it is necessary to be able to set the initial state of the system, to measure the state at all times, and to apply approximately impulsive inputs. These conditions may be satisfied for certain mechanical systems, although in the case of fluid dynamics, for example, it may be difficult to obtain the required state measurements and we may be forced to rely on data from simulation.

We also need to choose the sets  $\mathcal{T}$  and  $\mathcal{M}$ . A reasonable simple choice is  $\mathcal{T} = \{I, -I\}$ , since this corresponds to using both positive and negative inputs (or initial states) on each input separately. For distributed actuators, larger sets may be justified. The sets  $\mathcal{M}$  specify the size of inputs and states we are interested in. The choice is motivated by the magnitudes of the inputs and states actually seen in experiments, with the goal that during computation of the empirical Gramians, the dynamics should evolve in a region of state-space close to that in which the closed-loop system will be operating. It is not necessary to use the same  $\mathcal{M}$  for the controllability and observability experiments.

As for the Karhunen–Loève reduction of autonomous systems, a crucial feature of this procedure is that, in general, we should expect it to work well only in *limited regions of state-space*. We do not expect to be able to approximate global behaviour of the system well using a linear projection; however, it is perhaps more reasonable to expect to be able to construct low-order models for systems within some given operating region. For some systems, one may prefer not to

use impulsive inputs, and in particular it is known that the effects of persistent excitation are often important for nonlinear systems. In this case, the procedure for computation of the empirical controllability gramian may need modification, and similar results may be derived for other classes of input signals.

A simple numerical technique for balancing the empirical gramians  $\hat{X}$  and  $\hat{Y}$  is as follows. First, apply the Cholesky factorization [39] to  $\hat{Y}$  so that  $\hat{Y} = ZZ^*$ , with  $Z$  lower triangular with non-negative diagonal entries. Let  $U\Sigma^2U^*$  be a singular value decomposition of  $Z^*\hat{X}Z$ , and let  $T = \Sigma^{1/2}U^*Z^{-1}$ . Then  $T\hat{Y}T^* = \Sigma$ , and  $T^{-1}\hat{X}T^{-1} = \Sigma$ , as desired.

We can now change state co-ordinates of the nonlinear system, and truncate using the Galerkin projection. The rows of  $T$  may be thought of as giving the ‘modes’ of the system associated with the Hankel singular values.

## 5. MECHANICAL LINKS EXAMPLE

Not only are high-order systems hard to control, but it is difficult to develop intuition as to their behaviour. One of the advantages of the procedure developed here is that it can be viewed as a selection of appropriate mode shapes on which to project the nonlinear dynamics. These mode shapes often are physically meaningful, and in this section we give examples for a simple, mildly nonlinear mechanical system.

The system is shown in Figure 1. It consists of five uniform rigid rods in two-dimensions, connected via torsional springs and dampers. The lowest rod is pinned to ground with a torsional spring, so that the system has a stable equilibrium in the upright vertical position. There is no gravity. The system has a single input, a torque about the lowest pin joint, and a single output, the horizontal displacement of the end of the last rod from the vertical symmetry axis. The potential

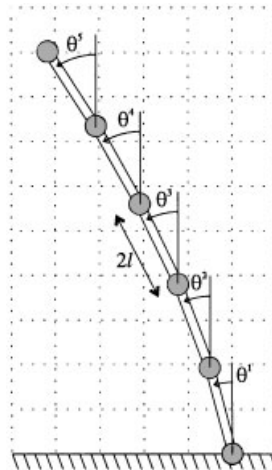


Figure 1. Example mechanical system.

energy  $V$  of this system is

$$V = \frac{1}{2}k(\theta^1)^2 + \frac{1}{2}k \sum_{i=2}^n (\theta^i - \theta^{i-1})^2 \tag{31}$$

and the kinetic energy  $T$  is

$$T = \frac{1}{2} \sum_{i=1}^n \frac{ml^2}{3} (\dot{\theta}^i)^2 + \frac{1}{2} \sum_{i=1}^n m((x^i)^2 + (y^i)^2) \tag{32}$$

where  $x^i, y^i$  are the Cartesian coordinates of the centre of mass of the  $i$ th rod, given by

$$x_i = \begin{cases} -l \sin \theta^i & \text{if } i = 1, \\ -l \sin \theta^i - 2l \sum_{p=1}^{i-1} \sin \theta^p & \text{if } i = 2, \dots, n \end{cases} \tag{33}$$

$$y_i = \begin{cases} l \cos \theta^i & \text{if } i = 1, \\ l \cos \theta^i + 2l \sum_{p=1}^{i-1} \cos \theta^p & \text{if } i = 2, \dots, n \end{cases} \tag{34}$$

The Lagrangian  $L = T - V$ , and the equations of motion are then

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}^i} - \frac{\partial L}{\partial \theta^i} = F_i \tag{35}$$

where  $F$  is the forcing term containing dissipative forces and the external force term  $w$ ,

$$F_i = \begin{cases} -b\dot{\theta}^i + w & \text{if } i = 1, \\ -b(\dot{\theta}^i - \dot{\theta}^{i-1}) & \text{if } i = 2, \dots, n \end{cases} \tag{36}$$

The measurement equation is

$$z = h(\theta, \dot{\theta}) := -2l \sum_{i=1}^n \sin \theta^i \tag{37}$$

The constants are given by  $b = 0.5, k = 3, m = 1, l = 1$ , and  $n = 5$ .

5.1. *Linearized model*

We first analyse the linearization of the system about its stable equilibrium; Figure 2 shows the configuration part of the first four of the ten mode shapes and the corresponding singular values of the balanced realization. The singular values have been normalized so that they sum to one.

5.2. *Empirical balancing*

Application of the above empirical model reduction procedure to the nonlinear Lagrangian model of this system leads to a set of corresponding modes. We choose  $\mathcal{M} = \{0.4\}$  and  $\mathcal{F} = \{I\}$  for both controllability and observability gramians. In this case, Figure 3 shows these modes; we can see that the first two mode shapes have split into three.

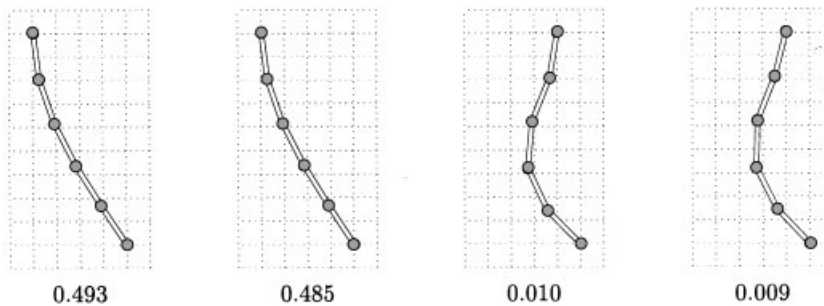


Figure 2. The balanced modes of the linearization, with the corresponding singular value of each mode.

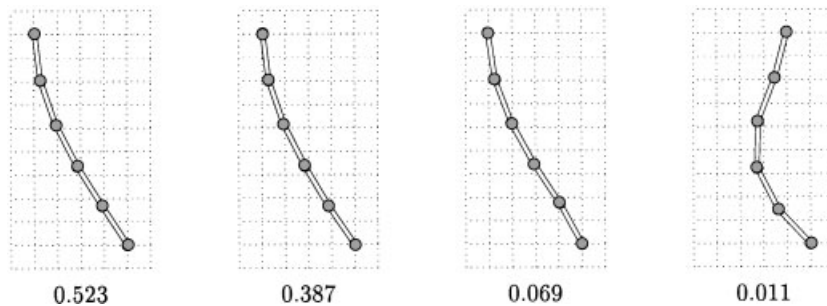


Figure 3. The empirical balanced modes of the nonlinear model, for  $c_i = 0.4$ .

Figure 4 shows the mode shapes using  $\mathcal{M} = \{1\}$  and  $\mathcal{T} = \{I\}$  for both controllability and observability gramians, so we are concentrating on the dynamics resulting from the larger force inputs. Here, the mode shapes become significantly different from those of the linearization.

Since this system is mechanical, the dynamics have a Lagrangian structure. In the absence of forcing and dissipation, Lagrangian systems conserve energy as well as quantities associated with the symmetries of the system. The dynamics of a mechanical system also satisfy a variational principle, and the evolution maps are symplectic transformations. All of these properties can be viewed as fundamental to a model of a mechanical system. Methods for model reduction, which take account of this underlying geometry and preserve it, are developed in References [7, 40]. The importance of this is evidenced in Figure 2, where each mode shape appears twice; this is a consequence of the underlying correspondence between configuration variables and their generalized momenta. These repeated structures are captured by the linearized method, and the method in this paper should be improved to take account of this repeated structure. One way to achieve this would be by combining the techniques in this paper with those of [7, 40] and constructing the empirical gramians on the configuration space rather than the phase space. In this way, the underlying geometric structure is preserved by the reduction, and the resulting reduced-order system is itself a Lagrangian system. Similarly, taking account of symmetry

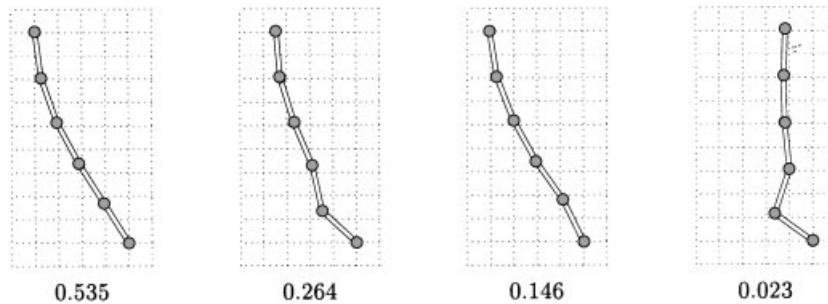


Figure 4. The empirical balanced modes of the nonlinear model, for  $c_i = 1$ .

[10, 11] and other special structure present in the original system can be of great use, leading to reduced computational requirements and producing more accurate reduced-order models.

## 6. CONCLUSIONS

In this paper, we have presented new methods for model reduction of nonlinear systems. These methods combine the physical intuition of Karhunen–Loève techniques, as they have been applied to analysis of complex flows, with the control engineering methodology of balanced truncation.

The resulting method of model reduction is based on data. It requires only simple matrix computations, and may be applied directly to nonlinear systems. When applied to linear systems, the reduced-order model constructed is exactly the usual balanced truncation of the system.

The choice of impulsive inputs, and initial states distributed uniformly on the unit sphere, are perhaps somewhat arbitrary methods of parametrizing trajectories for nonlinear systems. Many other methods might be suggested, and for control a very reasonable strategy would be to apply a Karhunen–Loève decomposition of the states seen during simulation of *closed-loop* behaviour. There are several other possibilities that might be suggested, such as subspace-identification procedures to compute both empirical gramians simultaneously; our focus here has been the construction of a computable extension of balanced truncation for nonlinear systems.

Qualitatively, the method preserves those states of the dynamics which are reachable using small input energy, and those states which result in large output energy, where by energy we mean the usual 2-norm. The empirical Gramian matrices are used numerically to measure these effects and it is the energy transfer, that is the product of these quantities, which is used to determine which states to keep and which to reduce.

The effectiveness of this procedure is determined by its usefulness in control applications, and to evaluate this is necessary to design a controller for the reduced-order nonlinear system and check its performance and stability, both in simulation with the full-order model and with the real system. One feature of the empirical procedure described in this paper is that often the simulation itself has been implemented using known heuristics; for example vortex methods are often used to simulate incompressible flow. Using data from such a simulation for model reduction, these heuristics can be incorporated into the control design process.

## ACKNOWLEDGEMENTS

We thank John Doyle, Petr Krysl, Peter Schröder, and Clancy Rowley for helpful comments and inspiration. This research was partially supported by AFOSR MURI grant F49620-96-1-0471, EPRI/DOD CIN/SI project WO8333-06 and NSF/DARPA OPAAL grant DMS-9874082.

## REFERENCES

1. Wu F, Yang XH, Packard A, Becker G. Induced  $L_2$ -norm control for LPV systems with bounded parameter variation rates. *International Journal of Robust and Nonlinear Control* 1996; **6**(9–10):983–998.
2. Pearson K. On lines and planes of closest fit to points in space. *Philosophical magazine* 1901; **2**:609–629.
3. Lumley JL. The structure of inhomogeneous turbulence. In *Atmospheric turbulence and wave propagation*, Yaglom AM, Tatarski VI (eds). Nauka: Moscow, 1967; 166–178.
4. Berkooz G, Holmes P, Lumley JL. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics* 1993; **25**:539–575.
5. Holmes P, Lumley JL, Berkooz G. *Turbulence, Coherent Structures, Dynamical Systems, and Symmetry*. Cambridge University Press: Cambridge, 1996.
6. Sirovich L. Turbulence and dynamics of coherent structures—parts I–III. *Quarterly of Applied Mathematics* 1987; **45**(3):561–590.
7. Krysl P, Lall S, Marsden JE. Dimensional model reduction in nonlinear finite-element dynamics of solids and structures. *International Journal for Numerical Methods in Engineering*, to appear.
8. Azeez MFA, Vakakis AF. Using Karhunen–Loève decomposition to analyze the vibro-impact response of a rotor. *Seventh Conference on Nonlinear Vibrations, Stability, and Dynamics of Structures*, Blacksburg, VA, July 1998.
9. Sirovich L, Rodriguez JD. Coherent structures and chaos: a model problem. *Physics Letters A* 1987; **120**(5):211–214.
10. Glavaški S, Marsden JE, Murray RM. Model reduction, centering, and the Karhunen–Loève expansion. *Proceedings of the IEEE Conference on Decision and Control*, vol. 37, 1998; 2071–2076.
11. Rowley CW, Marsden JE. Reconstruction equations and the Karhunen–Loève expansion for systems with symmetry. *Physica D* 2000; **142**:1–19.
12. Atwell JA, King BB. Computational aspects of reduced order feed-back controllers for spatially distributed systems. *Proceedings of the IEEE Conference on Decision and Control*, December 1999, 4301–4306.
13. Bloch AM, Marsden JE. Controlling homoclinic orbits. *Theoretical and Computational Fluid Mechanics* 1989; **1**:179–190.
14. Moore BC. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Transactions on Automatic Control* 1981; **26**(1):17–32.
15. Dullerud GE, Paganini F. *A Course in Robust Control Theory: a Convex Approach*. Springer: Berlin, 2000.
16. Glover K. All optimal Hankel-norm approximations of linear multivariable systems and their L-infinity error bounds. *International Journal of Control* 1984; **39**:1115–1193.
17. Enns DF. *Model reduction for control system design*. PhD thesis, Stanford University, 1984.
18. Scherpen JMA. Balancing for nonlinear systems. *Systems and Control Letters* 1993; **21**(2):143–153.
19. Scherpen JMA, van der Schaft AJ. Normalized coprime factorizations and balancing for unstable nonlinear-systems. *International Journal of Control* 1994; **60**(6):1193–1222.
20. Scherpen JMA. H-infinity balancing for nonlinear systems. *International Journal of Robust and Nonlinear Control* 1996; **6**:645–668.
21. Pavel L, Fairman FW. Controller reduction for nonlinear plants — an  $L_2$  approach. *International Journal of Robust and Nonlinear Control* 1997; **7**:475–505.
22. Al-Baiyat SA, Bettayeb M. A new model reduction scheme for k-power bilinear systems. *Proceedings of the IEEE Conference on Decision and Control*, 1993.
23. Al-Baiyat SA, Bettayeb M, Al-Saggaf UM. New model reduction scheme for bilinear systems. *International Journal of Systems Science* 1994; **25**:1631–1642.
24. Gray WS, Mesko J. Energy functions and algebraic gramians for bilinear systems. *Proceedings of the 4th IFAC Nonlinear Control Systems Design Symposium*, 1998.
25. Newman AJ, Krishnaprasad PS. Computation for nonlinear balancing. *Proceedings of the IEEE Conference on Decision and Control*, 1998.
26. Newman AJ, Krishnaprasad PS. Nonlinear model reduction for RTCVD. *Proceedings of the 32nd Conference on Information Sciences and Systems*, Princeton, NJ, March 1998.
27. Wood GD, Goddard PJ, Glover K. Approximation of linear parameter-varying systems. *Proceedings of the IEEE Conference on Decision and Control*, 1996; 406–411.



28. Beck CL, Doyle JC, Glover K. Model reduction of multidimensional and uncertain systems. *IEEE Transactions on Automatic Control* 1996; **41**:1466–1477.
29. Beck CL, Doyle JC. A necessary and sufficient minimality condition for uncertain systems. *IEEE Transactions on Automatic Control* 1999; **44**:1802–1813.
30. Lall S, Beck C. Model reduction of complex systems in the linear-fractional framework. *IEEE Conference on Computer-Aided Control System Design*, 1999.
31. Rathinam M, Petzold L. An iterative method for simulation of large-scale modular systems using reduced order models. *Proceedings of the IEEE Conference on Decision and Control*, 2001.
32. Petzold L, Zhu W. Model reduction for chemical kinetics: an optimization approach. *AiChE Journal* 1999; **45**(4):869–886.
33. Lall S, Beck C. Error bounds for model reduction via balanced truncation of linear time-varying systems. *IEEE Transactions on Automatic Control*, submitted for publication.
34. Moore BC. Principal component analysis in nonlinear systems: preliminary results. *Proceedings of the IEEE Conference on Decision and Control* 1979; 1057–1060.
35. van der Veen A-J, Deprettere EF, Swindlehurst AL. Subspace-based signal analysis using singular value decomposition. *Proceedings of the IEEE* 1993; **81**(9):1277–1308.
36. Lall S, Marsden JE, Glavaški S. Empirical model reduction of controlled nonlinear systems. *Proceedings of the IFAC World Congress*, vol. F, July 1999; 473–478.
37. Hahn J, Edgar TF. Reduction of nonlinear models using balancing of empirical gramians and galerkin projections. *Proceedings of the American Control Conference*, December 2000, 2864–2868.
38. Kailath T. *Linear Systems*. Prentice-Hall: Englewood Cliffs, NJ, 1980.
39. Golub GH, Van Loan CF. *Matrix Computations*. Johns Hopkins University Press: Washington, D.C., 1996.
40. Lall S, Krysl P, Marsden JE. Structure-preserving model reduction of mechanical systems. In preparation.
41. Bloch AM. Estimation, principal components, and Hamiltonian systems. *Systems and Control Letters* 1985; **6**:103–108.
42. Glavaški S. *Robust system analysis and nonlinear system model reduction*. PhD thesis, California Institute of Technology, May 1998.
43. Gray WS, Mesko J. General input balancing and model reduction for linear and nonlinear systems. *Proceedings of the European Control Conference*, 1997.