

UCRL-JRNL-210209



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

The effect of problem perturbations on nonlinear dynamical systems and their reduced order models

Radu Serban, Chris Homescu, Linda Petzold

March 4, 2005

SIAM Journal on Scientific Computing

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

THE EFFECT OF PROBLEM PERTURBATIONS ON NONLINEAR DYNAMICAL SYSTEMS AND THEIR REDUCED ORDER MODELS *

RADU SERBAN [†], CHRIS HOMESCU [‡], AND LINDA R. PETZOLD [§]

Abstract. Reduced order models are used extensively in many areas of science and engineering for simulation, design, and control. Reduction techniques for nonlinear dynamical systems produce models that depend strongly on the nominal set of parameters for which the reduction is carried out. In this paper we address the following two questions: “What is the effect of perturbations in the problem parameters on the output functional of a nonlinear dynamical system?” and “To what extent does the reduced order model capture this effect?”

Key words. model reduction, small sample statistical condition estimation, adjoint method, valid parameter range, singular value analysis

AMS subject classifications.

1. Introduction. In this paper we consider nonlinear dynamical systems of the generic form

$$\mathcal{M}(t, y', y, p) = 0, \quad y(t_0) = y^0(p), \quad (1.1)$$

where y are the model states, y^0 are the initial conditions, and p are model parameters. The system may be an ordinary differential equation (ODE), a semi-discretized partial differential equation (PDE), or a differential algebraic equation (DAE). An important problem which arises in a wide variety of engineering and scientific applications is to find the amount of change which can be expected in a model response functional $h(t, y(t), p)$ at a later time t , as a function of changes in the parameters p . The direction in the parameter space of the perturbation that produces the maximum change in the output may also be needed. For example, in weather prediction one wishes to find those perturbations in the initial conditions that can produce the greatest change in the accuracy of the forecast, because they identify the locations where the acquisition of additional data can be most effective [6, 15].

For large-scale dynamical systems, such decisions making use of the sensitivity of the system to perturbations in the problem parameters may necessarily be based on the corresponding sensitivities of a reduced order model (ROM). A ROM corresponding to (1.1) is a system of the form

$$\widetilde{\mathcal{M}}(t, \tilde{y}', \tilde{y}, p) = 0, \quad \tilde{y}(t_0) = \tilde{y}^0(p), \quad (1.2)$$

such that its solution \tilde{y} is an approximation to the solution y of (1.1). The ROM must be computationally less expensive than the full model. Depending on the particular model reduction technique employed, this can be achieved either through a reduction in the dimension of the state space, or through a reduction in the complexity of

*This work was supported by DOE DE-FG03-00ER25430, NSF/NCSA ACI-9619019, NSF CCF 0428912, and NSF/ITR ACI-0086061. The work of the first author was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

[†]Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, California 94551 (radu@llnl.gov)

[‡]Department of Computer Science, University of California, Santa Barbara, California 93106 (homescu@cs.ucsb.edu)

[§]Department of Computer Science, University of California, Santa Barbara, California 93106 (petzold@engineering.ucsb.edu)

the forcing term, or both. Nonlinear ROMs have been obtained through Proper Orthogonal Decomposition (POD) [13, 19], reduction of complexity [18, 20], Balanced Truncation [8], Principal Component Analysis [11] and many other techniques [1]. An important property of nonlinear ROMs is that they are (necessarily) based on a nominal set of parameters or a limited range of operating conditions. Thus the question naturally arises, to what extent does a given nonlinear ROM capture the changes in the output functional of the full system that result from perturbations in the problem parameters?

We present in Section §2 an efficient method for the estimation of perturbation-induced errors in the solution of a dynamical system. The method is based on a combination of adjoint sensitivity analysis [4] and the small sample statistical condition estimate (SCE) method [12, 7]. We show how to apply these ideas to systems described by ODEs, semi-explicit index-1 DAEs, and Hessenberg index-2 DAEs [9, 3]. This algorithm is closely related to our earlier work on estimating errors for POD-based ROM applied to ODE systems [10], but extends those techniques to DAEs and to more general nonlinear reduced-order models.

Using this methodology, in Section §3 we show how the direction in parameter space of maximum error growth in the output functional can be computed with very few solutions of the adjoint model. This is related to singular vector (SV) analysis in the weather prediction literature, e.g., [5]. By interpreting the initial perturbation as the error (or uncertainty) at the initial time, it was found in [5] that the maximum possible error growth is the largest singular value of the operator obtained from the linearization of the differential system. In section §4 we tackle the issue of estimating the extent to which a given ROM preserves the value of the maximum perturbation-induced error and/or the direction of the perturbations that induce it. We introduce two types of similarity index which measure this property. The problem of estimating the error which combines the effects of both model reduction and perturbations in parameters is addressed in Section §5.

Numerical results are given in Section §6. Models described by both ODEs and DAEs, and corresponding ROMs generated by several different methods are presented. An example is given for which the ROM is quite accurate for the nominal values of the parameters, but fails to capture the behavior of the full model when the parameters are slightly perturbed. We demonstrate that the similarity index can be very effective in detecting this type of situation.

2. Estimation of Perturbation-Induced Errors. In this section we present a method for the estimation of errors induced by perturbations in the initial conditions of a dynamical system. This approach, based on adjoint sensitivity analysis and small-sample statistical condition estimation(SCE), will be used throughout this paper.

For dynamical systems of the form (1.1), we wish to find an estimate of the error $e(t) = Y(t) - y(t)$ at some time $t > t_0$, where $Y(t)$ is the solution of the perturbed system

$$\mathcal{M}(t, Y', Y, p + \delta p) = 0, \quad Y(t_0) = y^0(p + \delta p). \quad (2.1)$$

Subtracting (1.1) from (2.1) yields the following equation for the error $e(t)$:

$$\mathcal{M}_{y'}(t)e' + \mathcal{M}_y(t)e + \mathcal{M}_p(t)\delta p + \mathcal{O}(\|e\|^2) = 0,$$

where $\mathcal{M}_{y'}$, \mathcal{M}_y , and \mathcal{M}_p are the Jacobians of \mathcal{M} with respect to y' , y , and p , respectively, evaluated at $(t, y'(t), y(t), p)$. Thus, to a first order approximation, the

error satisfies

$$\mathcal{M}_{y'}(t)e' + \mathcal{M}_y(t)e + \mathcal{M}_p(t)\delta p = 0, \quad e(t_0) = y_p^0(p) \delta p. \quad (2.2)$$

Note that the problem of estimating the error for perturbations in the initial conditions is a particular case of the problem of estimating the error for perturbations in the parameters, for which $p = y^0$. For any given perturbation δp , one can solve the forward sensitivity equation (2.2) to obtain the time evolution of the error $e(t)$ at any time $t > t_0$. However, this forward approach to error estimation can become expensive if one wishes to estimate the error over a range of perturbations, because (2.2) must be solved for *each* perturbation of interest.

We can, with the SCE technique described below, obtain an estimate of the norm of the error at some arbitrary but fixed time $t > t_0$, i.e., $\|e(t)\|$, as long as scalar products of the form $z^T e(t)$ are available, for any given unit vector z . This estimate takes the form

$$\|e(t)\| \approx \frac{W_q}{W_n} \left(\sum_{j=1}^q |z_j^T e(t)|^2 \right)^{1/2},$$

where $\{z_j\}_{j=1\dots q}$ are q orthogonal vectors, uniformly and randomly selected from the unit sphere \mathbb{S}_{n-1} , where n is the dimension of y .

In the remainder of this section, after a short introduction to the SCE estimation technique, we present an adjoint-based method for efficiently computing the scalar products $z^T e(t)$. First, we consider dynamical systems that can be written as ODEs and then we present the derivation for DAE systems. We base our SCE estimates on the solution of adjoint systems that are *independent* of δp and can thus be used for the estimation of errors induced by *any* perturbation.

2.1. Small sample statistical method for condition estimation. The small sample statistical condition estimation (SCE) method, originally proposed in [12], offers an efficient means for condition estimation for general nonlinear functions, at the cost of allowing moderate relative errors in the estimate. The basic idea is described below (for complete details, see [12, 7]).

For any vector $v \in \mathbb{R}^n$, if u is selected uniformly and randomly from the unit sphere \mathbb{S}_{n-1} , the expected value of $u^T v$ is proportional to the norm of v :

$$E(|u^T v|) = W_n \|v\|.$$

The proportionality factor, called the *Wallis factor*, depends only on n . It is defined as $W_1 = 1$ and

$$W_n = \begin{cases} \frac{1 \cdot 3 \cdots (n-2)}{2 \cdot 4 \cdots (n-1)} & \text{for } n \text{ odd} \\ \frac{2 \cdot 2 \cdot 4 \cdots (n-2)}{\pi \cdot 1 \cdot 3 \cdots (n-1)} & \text{for } n \text{ even} \end{cases}$$

and can be approximated by $W_n \approx \sqrt{\frac{2}{\pi(n - \frac{1}{2})}}$.

Thus the expression $\xi = \frac{|u^T v|}{W_n}$ is used to estimate the norm $\|v\|$. This estimate is first order in the sense that the probability of a relative error in the estimate is

inversely proportional to the size of the error. That is, for $\gamma > 1$, we have

$$\Pr\left(\frac{\|v\|}{\gamma} \leq \xi \leq \gamma\|v\|\right) \geq 1 - \frac{2}{\pi\gamma} + O\left(\frac{1}{\gamma^2}\right).$$

Additional function evaluations can improve the accuracy of the estimate.

Suppose that we obtain estimates $\xi_1, \xi_2, \dots, \xi_q$ corresponding to orthogonal vectors u_1, u_2, \dots, u_q selected uniformly and randomly from the unit sphere \mathbb{S}_{n-1} . The expected value of the norm of the projection of v onto the span \mathcal{U} generated by u_1, u_2, \dots, u_q is

$$E\left(\sqrt{|u_1^T v|^2 + |u_2^T v|^2 + \dots + |u_q^T v|^2}\right) = \frac{W_n}{W_q} \|v\|.$$

The estimate $\nu(q) = \frac{W_q}{W_n} \sqrt{|u_1^T v|^2 + |u_2^T v|^2 + \dots + |u_q^T v|^2}$ is q -th order accuracy, i.e., a relative error of size γ in the estimate occurs with probability proportional to γ^{-q} [12]:

$$\begin{aligned} \Pr\left(\frac{\|v\|}{\gamma} \leq \nu(2) \leq \gamma\|v\|\right) &\approx 1 - \frac{\pi}{4\gamma^2}, \\ \Pr\left(\frac{\|v\|}{\gamma} \leq \nu(3) \leq \gamma\|v\|\right) &\approx 1 - \frac{32}{3\pi^2\gamma^3}, \\ \Pr\left(\frac{\|v\|}{\gamma} \leq \nu(4) \leq \gamma\|v\|\right) &\approx 1 - \frac{81\pi^2}{512\gamma^4}. \end{aligned}$$

Usually at most four random vectors are required in practice.

2.2. Ordinary differential equation systems. Consider the following system:

$$f(t, y', y, p) = 0, \quad y(t_0) = y^0(p), \quad (2.3)$$

for $t \in [t_0, t_f]$, $y, y^0 \in \mathbb{R}^n$ and $f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$. The partial derivative matrix $f_{y'}$ is nonsingular. The linearized error equation in this case is

$$M(t)e' + A(t)e = -f_p(t)\delta p, \quad e(t_0) = y_p^0(p)\delta p, \quad (2.4)$$

where the matrices of partial derivatives $M(t) = f_{y'}$, $A(t) = f_y$, and f_p are evaluated at $(t, y'(t), y(t), p)$. For the remainder of this section, for the sake of readability, we will drop the time argument for these matrices, except where absolutely necessary.

We begin by noting that the solution of (2.4) can be written as

$$e(t) = \Phi(t) \left(y_p^0 - \int_{t_0}^t \Phi^{-1}(\tau) M^{-1} f_p d\tau \right) \delta p,$$

where $\Phi(t) \in \mathbb{R}^{n \times n}$ is the fundamental matrix corresponding to (2.4), i.e.,

$$M\Phi' + A\Phi = 0, \quad \Phi(t_0) = I. \quad (2.5)$$

It is straightforward to verify that the solution λ of the adjoint system

$$(M^T \lambda)' - A^T \lambda = 0, \quad (2.6)$$

with final condition $\lambda(t) = z$, satisfies

$$\begin{aligned} \lambda^T(\tau) &= z^T \Phi(t) \Phi^{-1}(\tau) M^{-1}(\tau), \quad \forall \tau \in [t_0, t] \\ \lambda^T(t_0) M(t_0) &= z^T \Phi(t) \end{aligned} \quad (2.7)$$

and therefore,

$$z^T e(t) = \left(\lambda^T(t_0)M(t_0)y_p^0 - \int_{t_0}^t \lambda^T(\tau)f_p d\tau \right) \delta p. \quad (2.8)$$

The case of perturbations in initial conditions is obtained by setting $f_p(t) = 0$, $\forall t$ and $y_p^0 = I$, while the case of perturbations in the right-hand side only corresponds to $y_p^0 = 0$.

If we consider q orthogonal vectors z_j , $1 \leq j \leq q$, uniformly and randomly selected from the unit sphere \mathbb{S}_{n-1} , we obtain the SCE estimate

$$\begin{aligned} \|e(t)\| &\approx \frac{W_q}{W_n} \left(\sum_{j=1}^q |z_j^T e(t)|^2 \right)^{1/2} \\ &= \frac{W_q}{W_n} \left(\sum_{j=1}^q \left| \left(\lambda_j^T(t_0)M(t_0)y_p^0 - \int_{t_0}^t \lambda_j^T(\tau)f_p(\tau) d\tau \right) \delta p \right|^2 \right)^{1/2}, \end{aligned} \quad (2.9)$$

where λ_j are solutions of the adjoint systems (2.6) with final conditions based on z_j .

An important property of this error estimate is that the adjoint system (2.6) on which it is based is *independent* of the perturbation δp . Thus the same adjoint solutions λ_j can be used to estimate the error induced by any perturbation. Finally, we note that very few (less than four) unit vectors z_j (and hence adjoint system solutions) are usually sufficient for acceptable SCE estimates.

In preparation for using the above estimate in comparing two different models, let us consider the perturbation-induced error in a model response functional $h(t, y(t), p)$. We seek the norm of this error, namely $\epsilon(t) := h(t, Y(t), p + \delta p) - h(t, y(t), p)$ (see Fig. 5.1). A linearization around $\delta p = 0$ yields

$$\epsilon(t) = h_y e(t) + h_p \delta p.$$

Thus

$$\begin{aligned} \|\epsilon(t)\| &\approx \frac{W_q}{W_{n_h}} \left(\sum_{j=1}^q |z_j^T h_y e(t) + z_j^T h_p \delta p|^2 \right)^{1/2} \\ &= \frac{W_q}{W_{n_h}} \left(\sum_{j=1}^q \left| \left(\lambda_j^T(t_0)M(t_0)y_p^0 - \int_{t_0}^t \lambda_j^T(\tau)f_p(\tau) d\tau + z_j^T h_p \right) \delta p \right|^2 \right)^{1/2}, \end{aligned}$$

where λ_j is the solution of the adjoint system (2.6) with the final condition $\lambda_j(t) = h_y^T(t)z_j$. We select the q vectors z_j uniformly and randomly from the unit sphere \mathbb{S}_{n_h-1} , where n_h is the dimension of h .

2.3. Semi-explicit index-1 differential-algebraic equation systems. A semi-explicit index-1 DAE takes the form

$$\begin{aligned} f(t, y', y, x, p) &= 0 \\ g(t, y, x, p) &= 0 \\ y(t_0) &= y^0(p), \quad x(t_0) = x^0(p), \end{aligned} \quad (2.10)$$

where y are the differential variables, x are the algebraic variables, and p are model parameters. The Jacobian matrices $f_{y'}$ and g_x are nonsingular. Note that the initial conditions y^0, x^0 are assumed to be consistent with the algebraic constraints, that is, $g(t_0, y^0, x^0, p) = 0$.

We are interested in estimating the norm of the error $e(t)$ between the solution $(Y(t), X(t))$ of the system with perturbed parameters and the solution $(y(t), x(t))$ of the unperturbed system. We decompose $e(t)$ into the differential and algebraic components: $e_d(t) = Y(t) - y(t)$ and $e_a(t) = X(t) - x(t)$. As in the ODE case, a linearization along the solution of the unperturbed DAE leads to

$$\begin{aligned} M(t)e'_y + A(t)e_d + B(t)e_a &= -f_p(t)\delta p \\ C(t)e_d + D(t)e_a &= -g_p(t)\delta p \\ e_d(0) &= y_p^0\delta p, \quad e_a(t_0) = x_p^0\delta p, \end{aligned} \tag{2.11}$$

where the matrices of partial derivatives, $M(t) = f_{y'}$, $A(t) = f_y$, $B(t) = f_x$, $C(t) = g_y$, $D(t) = g_x$, f_p and g_p are all evaluated at $(t, y(t), x(t), p)$. Consistency of the initial conditions for the above DAE is equivalent to the requirement that the perturbation δp satisfies

$$(C(t_0)y_p^0 + D(t_0)x_p^0 + g_p(t_0))\delta p = 0. \tag{2.12}$$

We reduce the analysis to the ODE case by using the essential underlying ODE (EUODE) [2]:

$$Me'_y + (A - BD^{-1}C)e_d = -(f_p - BD^{-1}g_p)\delta p, \quad e_d(t_0) = y_p^0\delta p. \tag{2.13}$$

The corresponding fundamental matrix Φ_1 satisfies

$$M\Phi'_1 + (A - BD^{-1}C)\Phi_1 = 0, \quad \Phi_1(t_0) = I, \tag{2.14}$$

and therefore

$$\begin{aligned} e_d(t) &= \Phi_1(t) \left(y_p^0 - \int_{t_0}^t \Phi_1^{-1}(\tau) M^{-1}(f_p - BD^{-1}g_p) d\tau \right) \delta p \\ e_a(t) &= -D^{-1}(Ce_d(t) + g_p\delta p). \end{aligned}$$

The SCE norm estimate of the error vector $[e_d^T(t); e_a^T(t)]^T$ is based on scalar products of the form

$$z_d^T e_d(t) + z_a^T e_a(t) = (z_d^T - z_a^T D^{-1}C) e_d(t) - z_a^T D^{-1}g_p\delta p,$$

for some vector $z = [z_d^T; z_a^T]^T \in \mathbb{S}_{n-1}$. It can be easily verified that the differential component λ of the solution of the adjoint system

$$\begin{aligned} (M^T\lambda)' - A^T\lambda - C^T\mu &= 0 \\ -B^T\lambda - D^T\mu &= 0 \\ \lambda(t) &= z_d - C^T D^{-T} z_a, \end{aligned} \tag{2.15}$$

satisfies $\lambda(\tau) = M^{-T}(\tau)\Phi_1^{-T}(\tau)\Phi_1^T(t)\lambda(t)$. As a consequence,

$$\begin{aligned} \left(\lambda^T(t_0)M(t_0)y_p^0 - \int_{t_0}^t \lambda^T(f_p - BD^{-1}g_p) d\tau \right) \delta p &= (z_d - C^T D^{-T} z_a)^T \\ &\times \left(\Phi_1(t)\Phi_1^{-1}(t_0)y_p^0 - \int_{t_0}^t \Phi_1(t)\Phi_1^{-1}(\tau)M^{-1}(f_p - BD^{-1}g_p) d\tau \right) \delta p. \end{aligned}$$

Thus

$$z_d^T e_d(t) + z_a^T e_a(t) = \left(\lambda^T(t_0)M(t_0)y_p^0 - \int_{t_0}^t \lambda^T(f_p - BD^{-1}g_p) d\tau - z_a^T D^{-1}g_p \right) \delta p.$$

or, using $\mu^T = -\lambda^T BD^{-1}$,

$$z_d^T e_d(t) + z_a^T e_a(t) = \left(\lambda^T(t_0)M(t_0)y_p^0 - \int_{t_0}^t (\lambda^T f_p + \mu^T g_p) d\tau - z_a^T D^{-1}g_p \right) \delta p. \quad (2.16)$$

Following the same argument as in §2.2, an SCE norm estimate for the perturbation-induced error $\epsilon(t)$ in a model response $h(t, y(t), x(t), p)$ can be obtained by replacing the final condition on λ with

$$\lambda(t) = (h_y - h_x D^{-1}C)^T z, \quad z \in \mathbb{S}_{n_h-1}$$

and using

$$z^T \epsilon(t) = \left(\lambda^T(t_0)M(t_0)y_p^0 - \int_{t_0}^t (\lambda^T f_p + \mu^T g_p) d\tau - z^T h_x D^{-1}g_p + z^T h_p \right) \delta p.$$

If we consider only perturbations in the initial conditions ($f_p = g_p = y_p^0 = 0$, implying $x_p^0 = I$) and if we are interested in the norm of the entire error vector $e(t)$, the following refinement of the above procedure, which takes advantage of the geometry of the problem, results in more accurate SCE estimates. We begin by observing that, since $e_d(t)$ and $e_a(t)$ must be consistent with the algebraic constraints, we need only consider unit vectors z that satisfy the linearized constraint $C(t)z_d + D(t)z_a = 0$. Let $n_y = \dim(y)$ be the number of degrees of freedom of the system (2.10). We start by selecting q vectors z_d uniformly and randomly from \mathbb{S}_{n_y-1} and, for each one of them, construct the vector $z_a = -D^{-1}(t)C(t)z_d$. We then orthonormalize the set of q vectors $z = [z_d^T; z_a^T]^T$ and thus obtain q orthogonal unit vectors in \mathbb{R}^n which are uniformly and randomly distributed over the sphere \mathbb{S}_{n_y-1} embedded on the linear subspace of \mathbb{R}^n of codimension $n - n_y$.

2.4. Hessenberg index-2 differential-algebraic equation systems. Consider next a dynamical system expressed in Hessenberg index-2 DAE form

$$\begin{aligned} f(t, y', y, x, p) &= 0 \\ g(t, y, p) &= 0 \\ y(t_0) &= y^0(p), \quad x(t_0) = x^0(p), \end{aligned} \quad (2.17)$$

where y are the differential variables, x are the algebraic variables, and p are model parameters. The matrices $f_{y'}$ and $g_y f_{y'}^{-1} f_x$ are nonsingular and the initial conditions are assumed to be consistent, i.e., $g(t_0, y^0, p) = 0$ and $\frac{d}{dt}g(t_0, y^0, p) = 0$.

Let $A(t) = f_y$, $B(t) = f_x$, $C(t) = g_y$, and $M(t) = f_{y'}$. To keep the derivation as simple as possible, we start by assuming that $M(t) = I$. Results for a general mass matrix will be given at the end of this section. With $M(t) = I$, the linearized error equation becomes

$$\begin{aligned} e_y' + A(t)e_d + B(t)e_a &= -f_p(t)\delta p \\ C(t)e_d &= -g_p(t)\delta p \\ e_d(t_0) &= y_p^0 \delta p, \quad e_a(t_0) = x_p^0 \delta p \end{aligned} \quad (2.18)$$

for the errors $e_d(t) = Y(t) - y(t)$ and $e_a(t) = X(t) - x(t)$, where $Y(t)$, $X(t)$ is the solution of (2.17) with perturbed parameters. Consistency of the initial conditions in (2.18) requires

$$\begin{aligned} Cy_p^0 \delta p &= -g_p \delta p \\ -(C' - CA)y_p^0 \delta p + CBx_p^0 \delta p &= (g'_p - Cf_p) \delta p \end{aligned} \quad (2.19)$$

at $t = t_0$.

For the Hessenberg index-2 DAE (2.18), the EUODE is derived as follows [2]. If B is sufficiently smooth, there exists a smooth bounded matrix function $R(t) \in \mathbb{R}^{(n_y - n_x) \times n_y}$ whose linearly independent, normalized rows form a basis for the nullspace of B^T . Thus, $RB = 0$ and $\begin{bmatrix} R \\ C \end{bmatrix}$ is invertible.

With the change of variables $u(t) = R(t)e_d(t)$, e_d is obtained as

$$e_d = \begin{bmatrix} R \\ C \end{bmatrix}^{-1} \begin{bmatrix} u \\ -g_p \delta p \end{bmatrix} = Su - Fg_p \delta p,$$

where $S(t) \in \mathbb{R}^{n_y \times (n_y - n_x)}$ satisfies $R(t)S(t) = I$ and $C(t)S(t) = 0$ and $F = B(CB)^{-1}$. The EUODE is obtained as $u' = (R' - RA)Su - (R' - RA)Fg_p \delta p - Rf_p \delta p$. Before proceeding further, we note that

$$\begin{aligned} F &= B(CB)^{-1} \\ SR &= I - B(CB)^{-1}C, \end{aligned}$$

and thus the EUODE can be written as

$$u' = (R' - RA)Su + R((B' + AB)(CB)^{-1}g_p - f_p) \delta p, \quad u(t_0) = u_0, \quad (2.20)$$

where we have used $R'F = R'B(CB)^{-1} = -RB'(CB)^{-1}$. Note that to satisfy the first condition in (2.19), a perturbation δp must be such that

$$y_p^0 \delta p = S(t_0)u_0 - B(t_0)(C(t_0)B(t_0))^{-1}g_p(t_0)\delta p, \quad (2.21)$$

for an arbitrary u_0 .

If Φ_2 is the fundamental matrix corresponding to (2.20), that is,

$$\Phi_2' = (R' - RA)S\Phi_2, \quad \Phi_2(t_0) = I,$$

then

$$\begin{aligned} u(t) &= \Phi_2(t) \left(u_0 + \int_{t_0}^t \Phi_2^{-1}(\tau) R((B' + AB)(CB)^{-1}g_p - f_p) \delta p d\tau \right) \\ e_d(t) &= Su(t) - Fg_p \delta p = Su(t) - B(CB)^{-1}g_p \delta p \\ e_a(t) &= (CB)^{-1} \left((C' - CA)e_d(t) + (g'_p - Cf_p) \delta p \right). \end{aligned}$$

Consider next the adjoint system

$$\begin{aligned} \lambda' - A^T \lambda - C^T \mu &= 0 \\ -B^T \lambda &= 0, \end{aligned} \quad (2.22)$$

whose EUODE can be obtained with the change of variables $v = S^T \lambda$ (implying $\lambda = R^T v$) as

$$v' = \left(S'^T + S^T A^T \right) R^T v = -S^T (R' - RA)^T v, \quad v(t) = \nu.$$

Since the EUODE of the adjoint is the adjoint of the EUODE (2.20) we get

$$\lambda(\tau) = R^T(\tau)v(\tau) = R^T(\tau)\Phi_2^{-T}(\tau)\Phi_2^T(t)\nu. \quad (2.23)$$

The scalar products required by the SCE procedure in this case are of the form

$$\begin{aligned} z_d^T e_d(t) + z_a^T e_a(t) &= z_d^T e_d(t) + z_a^T (CB)^{-1} (C' - CA) e_d(t) + z_a^T (CB)^{-1} (g'_p - C f_p) \delta p \\ &= (z_d^T + z_a^T (CB)^{-1} (C' - CA)) (S u(t) - B (CB)^{-1} g_p) \delta p + z_a^T (CB)^{-1} (g'_p - C f_p) \delta p. \end{aligned}$$

Using (2.23) and (2.21) we get

$$\begin{aligned} &\left(\lambda^T(t_0) y_p^0 + \int_{t_0}^t \lambda^T ((B' + AB)(CB)^{-1} g_p - f_p) d\tau \right) \delta p \\ &= \nu \Phi_2(t) \left(u_0 + \int_{t_0}^t \Phi_2^{-1}(\tau) R ((B' + AB)(CB)^{-1} g_p - f_p) \delta p d\tau \right), \end{aligned}$$

where we have used $R(t_0)S(t_0) = I$ and $R(t_0)B(t_0) = 0$. Taking into account $\mu^T = -\lambda^T (B' + AB)(CB)^{-1}$ and selecting the final condition for (2.22) to be

$$\lambda(t) = (I - B(CB)^{-1}C)^T (z_d + (C' - CA)^T (CB)^{-T} z_a),$$

we get

$$\begin{aligned} z_d^T e_d(t) + z_a^T e_a(t) &= \left(\lambda^T(t_0) y_p^0 - \int_{t_0}^t (\lambda^T f_p + \mu^T g_p) d\tau \right. \\ &\quad \left. - z_d^T B(CB)^{-1} g_p + z_a^T (CB)^{-1} (g'_p - (C' - CA)B(CB)^{-1} g_p - C f_p) \right) \delta p. \end{aligned}$$

If the final condition for λ is replaced with

$$\lambda(t) = (I - B(CB)^{-1}C)^T (h_y^T + (C' - CA)^T (CB)^{-T} h_x^T) z,$$

for some $z \in \mathbb{R}^{n_h}$, the resulting adjoint variable can be used in evaluating the scalar product $z^T \epsilon(t)$, which is necessary in an SCE estimate of the norm of the perturbation-induced error in the model response $h(t, y(t), x(t), p)$, as

$$\begin{aligned} z^T \epsilon(t) &= \left(\lambda^T(t_0) y_p^0 - \int_{t_0}^t (\lambda^T f_p + \mu^T g_p) d\tau \right) \delta p - z^T h_y B(CB)^{-1} g_p \\ &\quad + z^T h_x (CB)^{-1} (g'_p - (C' - CA)B(CB)^{-1} g_p - C f_p) + z^T h_p \delta p. \end{aligned}$$

Finally, for a general mass matrix $M(t)$, we note that the matrix $CM^{-1}B$ must always be nonsingular. Then the scalar product $z_d^T e_d(t) + z_a^T e_a(t)$ can be obtained, using the solution $\{\lambda, \nu\}$ of the following adjoint system

$$\begin{aligned} (M^T \lambda)' - A^T \lambda - C^T \mu &= 0 \\ -B^T \lambda &= 0 \end{aligned} \quad (2.24)$$

$$M^T(t)\lambda(t) = (I - M^{-1}B(CM^{-1}B)^{-1}C)^T (z_d + (C' - CM^{-1}A)^T (CM^{-1}B)^{-T} z_a),$$

as

$$\begin{aligned}
z_d^T e_d(t) + z_a^T e_a(t) &= \left(\lambda^T(t_0) M(t_0) y_p^0 - \int_{t_0}^t (\lambda^T f_p + \mu^T g_p) d\tau \right. \\
&\quad \left. - z_d^T M^{-1} B (C M^{-1} B)^{-1} g_p \right. \\
&\quad \left. + z_a^T (C M^{-1} B)^{-1} (g'_p - (C' - C M^{-1} A) M^{-1} B (C M^{-1} B)^{-1} g_p - C M^{-1} f_p) \right) \delta p.
\end{aligned} \tag{2.25}$$

3. Singular Vector (SV) Analysis. SV analysis, introduced by Lorenz in the 1960's [17], is employed to compute the largest error growth and to help estimate model predictability. Finding the direction of maximal error growth implies solving a generalized eigenvalue problem for a matrix involving the product $\Phi^T \Phi$, where Φ is the fundamental matrix of the differential system¹. Solving the resulting eigenvalue problem is quite computationally expensive. It is typically accomplished using Lanczos-type algorithms which require only the evaluation of matrix-vector products. However, every such product involves a forward integration of a linearized model followed by a backward in time solution of the corresponding adjoint model. Based on the combination of SCE estimate and adjoint models, we describe below a method for finding an approximation to the leading SV (representing the direction of maximal error growth) that requires only very few solutions of the adjoint model (equal to the number of unit vectors considered in the SCE estimate).

Consider again the error $\epsilon(t) := h(t, Y(t), p + \delta p) - h(t, y(t), p)$ in some model response functional $h(t, y, p)$. We wish to find the direction of perturbations in the parameters p that gives rise to the maximum error $\|\epsilon(t)\|$ at some time $t \geq t_0$. In other words, we seek

$$\mathcal{E} := \sup_{\substack{\|u\| = 1 \\ u \in \mathbb{R}^{n_u}}} \|\epsilon(t, u)\|, \tag{3.1}$$

where ϵ is the response error induced by a perturbation $\delta p = \delta p(u)$ in the parameters p . The directions u over which we search for the maximum error norm growth are constrained to belong to a search space \mathbb{R}^{N_u} such that the resulting perturbations $\delta p(u)$ are admissible. The dimension N_u is model and problem dependent and will be defined below for each of the dynamical systems discussed in §2.

Using an SCE estimate for the error norm in (3.1) we get

$$\mathcal{E}^2 = \left(\frac{W_q}{W_{n_h}} \right)^2 \sup_{\substack{\|u\| = 1 \\ u \in \mathbb{R}^{n_u}}} \sum_{j=1}^q |z_j^T \epsilon(t, u)|^2,$$

where, for the dynamical systems considered in this paper, the scalar products $z_j^T \epsilon(t, u)$ were derived in §2.2–2.4. The essence of these formulas is that, in all cases,

$$z_j^T \epsilon(t, u) = \alpha_j^T \delta p,$$

¹In the meteorological and weather prediction literature this matrix is often called the *propagator*.

with α_j defined in terms of the solution λ_j of an appropriate adjoint system. Therefore

$$\begin{aligned} \mathcal{E}^2 &= \left(\frac{W_q}{W_{n_h}} \right)^2 \sup_{\substack{\|u\|=1 \\ u \in \mathbb{R}^{n_u}}} \sum_{j=1}^q |\alpha_j^T \delta p(u)|^2 = \left(\frac{W_q}{W_{n_h}} \right)^2 \sup_{\substack{\|u\|=1 \\ u \in \mathbb{R}^{n_u}}} \delta p^T(u) \left(\sum_{j=1}^q \alpha_j \alpha_j^T \right) \delta p(u) \\ &= \left(\frac{W_q}{W_{n_h}} \right)^2 \sup_{\substack{\|u\|=1 \\ u \in \mathbb{R}^{n_u}}} u^T P^T \left(\sum_{j=1}^q \alpha_j \alpha_j^T \right) P u, \end{aligned}$$

where we have assumed that the admissible perturbations can be written as $\delta p(u) = Pu$, for some matrix P . Let $\Psi := P^T \left(\sum_{j=1}^q \alpha_j \alpha_j^T \right) P$. The matrix Ψ is symmetric and has $\text{rank}(\Psi) \leq q$. Thus its singular value decomposition (SVD) is $\Psi = U^T \Sigma U = \sum_{j=1}^q \sigma_j u_j u_j^T$, and $u^T \Psi u = \sum_{j=1}^q (\sqrt{\sigma_j} u^T u_j)^2$. If the singular values of Ψ are numbered in decreasing order, it is easy to see that

$$\sup_{\substack{\|u\|=1 \\ u \in \mathbb{R}^{n_u}}} u^T \Psi u = \sigma_1,$$

which is attained for $u^* = u_1$. Thus

$$\mathcal{E} = \frac{W_q}{W_{n_h}} \sqrt{\sigma_1}, \quad (3.2)$$

and the direction of perturbations along which the error is maximized is $\delta p^* = Pu_1$.

Next we establish the dimension n_u in (3.1) and the corresponding matrix P (defining the allowable parameter perturbations $\delta p(u) = Pu$) for the dynamical systems discussed in the previous section.

1. For an ODE system, any perturbation δp is admissible. Thus $n_u = n_p \equiv \dim(p)$ and $P = I_{n_p}$.
2. For a semi-explicit index-1 DAE, requiring the initial condition for (2.12) to be consistent is equivalent to requiring that δp be in the null-space of the matrix $Q_1 = C(t_0)y_p^0 + D(t_0)x_p^0 + g_p(t_0)$. Then $P = \mathcal{N}(Q_1)$ and $n_u = \dim(\mathcal{N}(Q_1))$, where $\mathcal{N}(Q_1)$ is a basis for the null-space of Q_1 . In the particular case of perturbations in initial conditions $e_d(t_0) = \delta p$ and $e_a(t_0) = -D^{-1}(t_0)C(t_0)\delta p$. Therefore $\mathcal{U} \equiv \mathbb{R}^{n_y}$ and $P = I_{n_y}$.
3. For a Hessenberg index-2 DAE, similarly to the index-1 DAE case, the conditions (2.19), modified here to account for a general mass matrix, are equivalent to the requirement that δp be in the null-space of

$$Q_2 = \begin{bmatrix} C y_p^0 + g_p \\ -(C' - C M^{-1} A) y_p^0 + C M^{-1} B x_p^0 - g'_p + C M^{-1} f_p \end{bmatrix}$$

at $t = t_0$. Thus $P = \mathcal{N}(Q_2)$ and $n_u = \dim(\mathcal{N}(Q_2))$. In the case of perturbations in the initial conditions, $e_d(t_0) = \delta p$ and $e_a(t_0) = (C(t_0)B(t_0))^{-1}(C'(t_0) - C(t_0)A(t_0))\delta p$. Consistent perturbations must then satisfy $\delta p = S(t_0)u$, with $u \in \mathbb{R}^{n_y - n_x}$. Note that $n_y - n_x$ represents the number of degrees of freedom of the system (2.18).

Finally, if the parameter perturbations δp are to be further restricted to the range of some matrix Π , then $P = \Pi$ in the ODE case, and $P = \Pi \mathcal{N}(Q\Pi)$, where $Q = Q_1$ for index-1 DAEs or $Q = Q_2$ for index-2 DAEs.

4. Quality Measures for Reduced Order Models. The quality of a given ROM can be judged by the extent to which it preserves the value of the maximum possible error and/or the direction of perturbations that induce it. We begin by constructing a matrix whose entries

$$m_{ij} = \langle u_i, \tilde{u}_j \rangle^2$$

are the squared scalar products of the i -th SV of the full model and the j -th SV of the reduced model. These entries represent the amount of energy ² of the i -th full model SV that can be explained by the j -th reduced model SV. The SV analysis must be based on a model response that is common to both the full and reduced-order models. Based on the matrix m_{ij} , a similarity index can be defined ([16]) as

$$\mathcal{S}_1 = \frac{1}{q} \sum_{i,j=1}^q m_{ij}. \quad (4.1)$$

Although the proposed method (based on SCE norm estimation) provides only an approximation to \mathcal{S}_1 , since the evaluation of \mathcal{S}_1 requires several singular vectors, it is more computationally attractive than the alternative offered by Lanczos-type algorithms, not to mention the option of computing the entire fundamental matrix Φ .

A second measure of similarity can be defined in terms of the ratio between the norms of the errors induced in the full model and in the reduced order model by a perturbation of unit norm along the dominant SV of the full model

$$\mathcal{S}_2 = \min \left(\frac{\|\tilde{\epsilon}(t, u_1)\|}{\|\epsilon(t, u_1)\|}, \frac{\|\epsilon(t, u_1)\|}{\|\tilde{\epsilon}(t, u_1)\|} \right). \quad (4.2)$$

Since both similarity indexes are related to the assessment of ROM quality, one may ask when one index is more relevant than the other. Each similarity index evaluates different aspects of the ROM quality. The first index, \mathcal{S}_1 , quantifies the similarity between subspaces spanned by the first q SVs for the full model and, respectively, for the reduced model. Thus it provides an overall picture of ROM approximation of the phenomena described by the full model. The second index, \mathcal{S}_2 , gives a more localized result, by estimating the extent to which the ROM preserves the error due to perturbations along the direction of maximal error growth, given by the dominant SV of the full model.

Finally, we note that both indexes take values between 0 and 1 and that for identical models, they will equal unity. The index \mathcal{S}_2 is a more stringent ROM quality test, in the sense that, typically, $\mathcal{S}_2 < \mathcal{S}_1$.

5. ROM Error under Perturbations. Consider next the problem of estimating the error $\|\Theta(t)\|$, which combines the effects of both model reduction and perturbations in parameters (see Fig. 5.1). For an arbitrary, but fixed, time $t > t_0$, let ³

$$\begin{aligned} \Theta(t) &= h(Y) - \tilde{h}(\tilde{Y}) = (h(y) - \tilde{h}(\tilde{y})) + (h(Y) - h(y)) - (\tilde{h}(\tilde{Y}) - \tilde{h}(\tilde{y})) \\ &= \theta(t) + \epsilon(t) - \tilde{\epsilon}(t) \\ &\approx \theta(t) + h_y e(t) - \tilde{h}_{\tilde{y}} \tilde{e}(t) + (h_p - \tilde{h}_p) \delta p, \end{aligned} \quad (5.1)$$

²The energy depends on the particular inner product used. Here we use the usual Euclidean inner product.

³For the sake of clarity, only the relevant function arguments were kept.

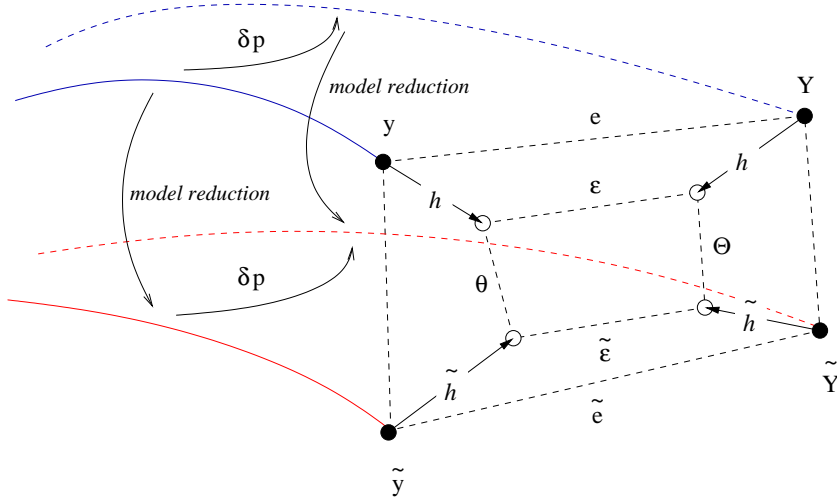


FIG. 5.1. Graphical representation of model reduction and perturbation-induced errors. Solid-line trajectories and quantities in lower case (y, \tilde{y}, θ) correspond to unperturbed systems, while dash-line trajectories and states in upper-case (Y, \tilde{Y}, Θ) correspond to the perturbed systems. Quantities with a tilde correspond to reduced order models.

where $\theta(t) = h(t, y(t), p) - \tilde{h}(t, \tilde{y}(t), p)$ is the error in the model response due solely to model reduction, $e(t) = Y(t) - y(t)$ is the error introduced by the perturbation in the original system, and $\tilde{e}(t) = \tilde{Y}(t) - \tilde{y}(t)$ is the error introduced by the perturbation in the reduced system. We distinguish two cases. In both, the projections $z_j^T h_y e(t)$ and $z_j^T \tilde{h}_{\tilde{y}} \tilde{e}(t)$ are based on the solution of adjoint problems corresponding to the full and reduced models, respectively, using the procedure of §2.

1. If the components of the error $\theta(t)$ are known, we can directly use an SCE estimate for the norm $\|\Theta(t)\|$:

$$\begin{aligned} \|\Theta(t)\| &= \|\theta(t) + h_y e(t) - \tilde{h}_{\tilde{y}} \tilde{e}(t)\| \\ &\approx \frac{W_q}{W_{n_h}} \left[\sum_{j=1}^q \left| z_j^T \theta(t) + z_j^T (h_p - \tilde{h}_p) \delta p + z_j^T h_y e(t) - z_j^T \tilde{h}_{\tilde{y}} \tilde{e}(t) \right|^2 \right]^{1/2}. \end{aligned}$$

2. If only the norm $\|\theta(t)\|$ is available, we use the bounds

$$\begin{aligned} \left| \|\theta(t) + (h_p - \tilde{h}_p) \delta p\| - \|h_y e(t) - \tilde{h}_{\tilde{y}} \tilde{e}(t)\| \right| &\leq \|E(t)\| \\ &\leq \|e(t) + (h_p - \tilde{h}_p) \delta p\| + \|h_y e(t) - \tilde{h}_{\tilde{y}} \tilde{e}(t)\| \end{aligned}$$

and the following SCE estimate:

$$\|h_y e(t) - \tilde{h}_{\tilde{y}} \tilde{e}(t)\| \approx \frac{W_q}{W_{n_h}} \left[\sum_{j=1}^q \left| z_j^T h_y e(t) - z_j^T \tilde{h}_{\tilde{y}} \tilde{e}(t) \right|^2 \right]^{1/2}.$$

6. Numerical Results. We have implemented our methodology to problems representative of the dynamical systems considered in this paper. We consider models

described by systems of ODEs: the 2-D Brusselator (modeling autocatalytic, oscillating chemical reactions) and the GRI combustion model (from chemical kinetics); and by DAEs (the heat shock model in biology).

6.1. 2D Brusselator. The following 2D time-dependent PDE models a chemically reacting system:

$$\begin{aligned} u_t &= \epsilon_1 \Delta u + A + u^2 v - (B + 1)u \\ v_t &= \epsilon_2 \Delta v + Bu - u^2 v, \end{aligned}$$

where the diffusion coefficients are $\epsilon_1 = \epsilon_2 = 2 \cdot 10^{-3}$ and the chemical parameters are given by $A = 1$ and $B = 3.4$. The PDE is defined over the square $\Omega = [0, 1] \times [0, 1]$, for $t \in [0, 1]$. Homogeneous Neumann conditions are imposed at the boundary of the domain, while the initial conditions are

$$u(x, y, 0) = 1 - 0.5 \cdot \cos(\pi y), \quad v(x, y, 0) = 3.5 - 2.5 \cdot \cos(\pi x). \quad (6.1)$$

Using the method of lines, we convert this PDE to an ODE initial value problem (IVP) using central finite differences to approximate the spatial derivatives on a grid of $N_x \times N_y = 16 \times 16$ interior points, with the boundary conditions explicitly used to define the necessary ghost boundary points. The resulting system, of the form (2.3), has dimension $n = 512$.

POD provides a method for finding the best approximating affine subspace to a given set of data. When using POD for model reduction of dynamical systems, the data are time snapshots of the solution obtained via numerical simulations or from experiments. Consider next the solutions of (1.1) at m time points, collected in the $n \times m$ matrix $\mathcal{Y} = [y(t_1) - \bar{y}, y(t_2) - \bar{y}, \dots, y(t_m) - \bar{y}]$, where \bar{y} is the mean of these observations. POD seeks a subspace $S \in \mathbb{R}^n$ and the corresponding projection matrix $P = \rho \rho^T$ so that the total square distance

$$\|\mathcal{Y} - P\mathcal{Y}\|^2 = \sum_{i=1}^m \|(y(t_i) - \bar{y}) - P(y(t_i) - \bar{y})\|^2$$

is minimized. The solution to the above minimization problem can be easily obtained by considering the SVD of the correlation matrix $R = \mathcal{Y}\mathcal{Y}^T$ (see [19, 10] for details). A POD-based reduced model that approximates the original problem can then be constructed by projecting onto S the model (1.1):

$$\rho^T \mathcal{M}(t, \rho \tilde{y} + \bar{y}, \rho \tilde{y}') = 0, \quad \tilde{y}(t_0) = \rho^T (y_0 - \bar{y}).$$

Returning to the Brusselator model, a POD-based ROM is constructed using 100 equally-spaced snapshots from the trajectory corresponding to the initial conditions (6.1) and retaining 12 modes. We note that the resulting ROM retains 99.52% of the “energy” of the full system. Moreover, comparing the solutions of the full and reduced models at the final time, we find that

$$\frac{\|y(t) - \tilde{y}(t)\|}{\|y(t)\|} \cdot 100.0 = 0.3661\%. \quad (6.2)$$

We are therefore led to believe that this ROM (of dimension only 12) can very accurately represent the solution of full model. However, using the analysis of the preceding

TABLE 6.1

Brusselator example: SV analysis results. For both the full and reduced models the following quantities are given: the estimated maximum perturbation-induced error norm given by (3.2) at the final time $t = 1.0$, the actual norm of the error in the estimated direction of maximum error growth and the norms of errors induced by several random perturbations.

Full model		Reduced-order model	
\mathcal{E}	6.70690e+01	$\tilde{\mathcal{E}}$	6.50318e+00
$\ e(t, u_1)\ $	5.18372e+01	$\ \tilde{e}(t, \tilde{u}_1)\ $	5.85795e+00
$\ e(t, u_{rnd})\ $	2.82639e+00	$\ \tilde{e}(t, \tilde{u}_{rnd})\ $	2.54461e+00
	5.18880e+00		2.58347e+00
	3.79540e+00		3.32946e+00
	5.81685e+00		1.59148e+00
	4.14449e+00		1.65752e+00

sections, it turns out that this ROM does not perform well under perturbations in the initial conditions.

We have applied the approach presented in §3 to estimate the leading singular vectors for both the full model and the ROM, using $q = 12$ and $q = 4$ unit vectors for the SCE, respectively. The results are summarized in Table 6.1. Next, applying the approach presented in §4 and using the first four leading singular vectors, Eq. (4.1) gives a similarity index of only $\mathcal{S}_1 = 0.03287$, while Eq. (4.2) gives $\mathcal{S}_2 = 0.02328$. The reason for this poor agreement is that the direction of maximum error growth for the full system is almost orthogonal to the POD subspace in which the ROM is bound to evolve. Indeed, if u_1 is the leading singular vector for the full system, we have that $\|(I - \rho * \rho^T)u_1\| = 0.95266$, while $\|\rho^T u_1\| = 0.30404$, where ρ is the POD projection matrix. Moreover, the leading singular vector \tilde{u}_1 of the ROM fails to approximate well even the projection of u_1 on the POD subspace: the angle $\angle(\tilde{u}_1, \rho^T u_1) = 150.87^\circ$ (an angle of 0° or 180° indicates perfect agreement). The discrepancies between the leading singular vector of the full model and that of the ROM can also be seen in Fig. 6.1 in which the singular vectors are color coded and superimposed over the unperturbed initial conditions.

As a final verification, introducing a perturbation along u_1 of magnitude of only $1/100 \cdot \|y_0\|$, the quantity (6.2) becomes

$$\frac{\|Y(t) - \tilde{Y}(t)\|}{\|Y(t)\|} \cdot 100.0 = 16.268\%,$$

which is in perfect agreement with the similarity index \mathcal{S}_2 (indeed, $0.02328 = \mathcal{S}_2 \approx 0.3661/16.268 = 0.0225$).

6.2. GRI combustion model. Given n_s chemical species with mass fractions y_i and m reactions, the constant-volume batch reactor is modeled by the ODE system

$$y' = S(y)F(y), \quad y(t_0) = y_0, \quad (6.3)$$

where $S \in \mathbb{R}^{n_s \times m}$ includes the stoichiometric matrix and the thermodynamic equations. $F \in \mathbb{R}^m$ is the vector of nonlinear responses corresponding to elementary reactions, with $F_i(y)$ denoting the reaction rate of the i -th physical elementary reaction.

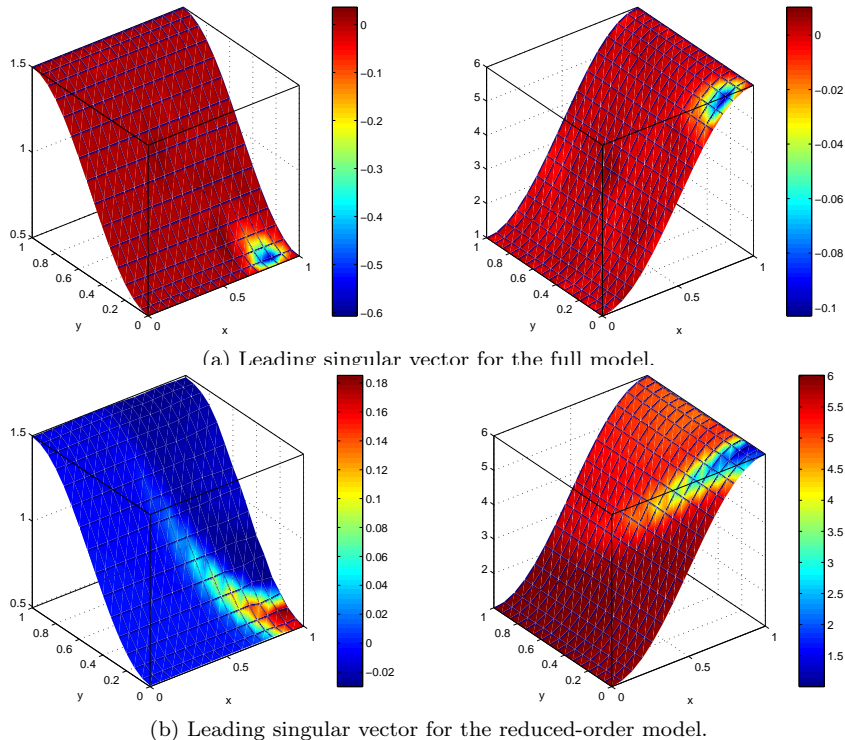


FIG. 6.1. *Brusselator example: Singular vectors color coded and superimposed over the initial conditions (u component on the left, v component on the right).*

A ROM can be constructed by choosing some (small) subset of the reactions to use in the mechanism so that the behavior of the reduced system is as close as possible to that of the original system, given a range of operating conditions. Often this is done by a chemist based on physical knowledge and intuition. A numerical method for the construction of such reduced models is given by [18]. In that approach, one seeks a ROM in the form

$$\tilde{y}' = S(\tilde{y})DF(\tilde{y}), \quad \tilde{y}(t_0) = y_0, \quad (6.4)$$

such that the norm of the error $e = y - \tilde{y}$ is minimized over a given interval $[t_0, t]$. In (6.4), $D \in \mathbb{R}^{m \times m}$ is a diagonal matrix whose diagonal elements d_i are either 1 or 0 (depending on whether or not the reaction i is selected for the reduced mechanism). The problem of finding the reduced mechanism is written as a discrete constrained optimization where the minimum is over d_1, d_2, \dots, d_m , with $d_i \in \{0, 1\}$ and $\sum d_i = k \ll m$ is prescribed.

The GRI combustion model is a set of $m = 177$ reactions describing the combustion of natural gas and involves $n_s = 32$ chemical species. For more details on this model, see [18, 20]. The resulting ODE system has dimension $n = n_s + 1 = 33$, the last differential equation describing the evolution of the temperature; i.e., $y = [c, T/T_0]$, where c are the species concentrations and T is the temperature, scaled by its initial value T_0 .

We consider two different ROMs, corresponding to $k = 32$ and $k = 78$ reactions selected in the reduced mechanism and denoted by ROM₃₂ and ROM₇₈, respectively.

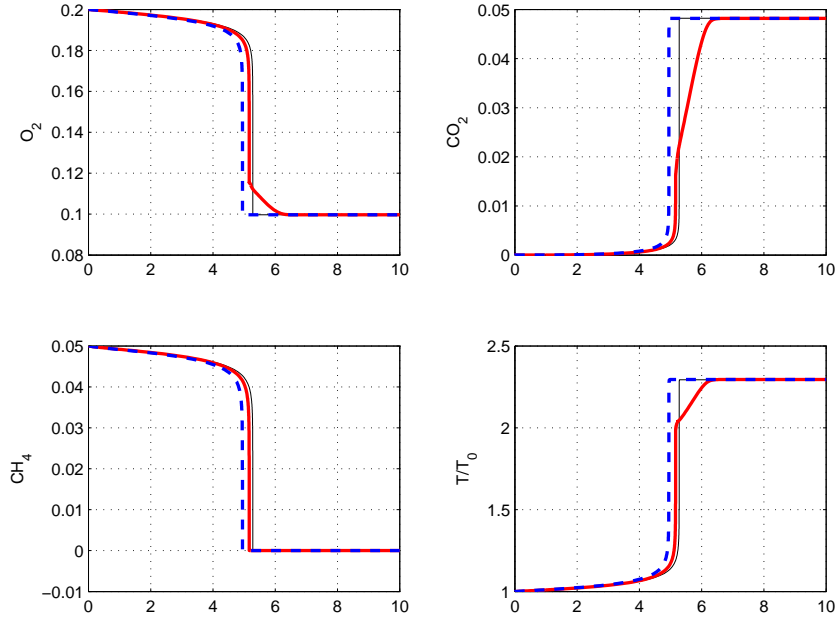


FIG. 6.2. *GRI combustion example: Full system (thin line) and ROM solutions (ROM₃₂: thick solid line, ROM₇₈: thick dashed line).*

TABLE 6.2
GRI combustion example: Similarity indexes.

ROM ₃₂			ROM ₇₈		
	$t_1 = 5$	$t_2 = 10$		$t_1 = 5$	$t_2 = 10$
\mathcal{S}_1	0.8991	0.6076	\mathcal{S}_1	0.3656	0.6306
\mathcal{S}_2	0.4992	$9.366 \cdot 10^{-3}$	\mathcal{S}_2	$1.870 \cdot 10^{-2}$	$7.206 \cdot 10^{-2}$

The time evolution of O_2 , CH_4 , and CO_2 , as well as T/T_0 are shown in Figure 6.2. The full system solution is depicted with a thin solid line, while the ROM solutions are depicted with thick lines (solid for ROM₃₂ and dashed for ROM₇₈).

For each ROM the similarity indexes \mathcal{S}_1 and \mathcal{S}_2 were calculated using $q = 5$ SCE unit vectors. The model parameters were defined to be scaling factors, with nominal values of 1.0, for the $m = 177$ reaction rates in the full model. The model response was taken to be the set of all species concentrations; i.e. $h(t, y, p) = c(t)$. Table 6.2 provides the similarity indexes at two different times, $t_1 = 5$ and $t_2 = 10$, before and after ignition, respectively.

As these results show, ROM₇₈ is more robust than ROM₃₂ based on the similarity indexes at $t_2 = 10$, but much less robust at the time before ignition, $t_1 = 5$. To explain this (apparently) surprising result, note first that the selection of reactions kept in the reduced-order models was based on minimizing the norm $\|e\|_{L_2} = \|y - \tilde{y}\|_{L_2}$ over the time interval $[t_0, t_2]$. For ROM₇₈ we have $\|e_{78}^*\| = 1.3736 \cdot 10^{-2}$, while for ROM₃₂, $\|e_{32}^*\| = 3.0349 \cdot 10^{-1}$. In this global measure, ROM₇₈ is a better reduced-order model, but, as can be seen in Fig. 6.2, at the price of predicting an earlier ignition. Secondly, note that a parameter perturbation along the direction of largest error growth at $t_1 = 5$ for the full model also leads to an earlier ignition time. As a consequence,

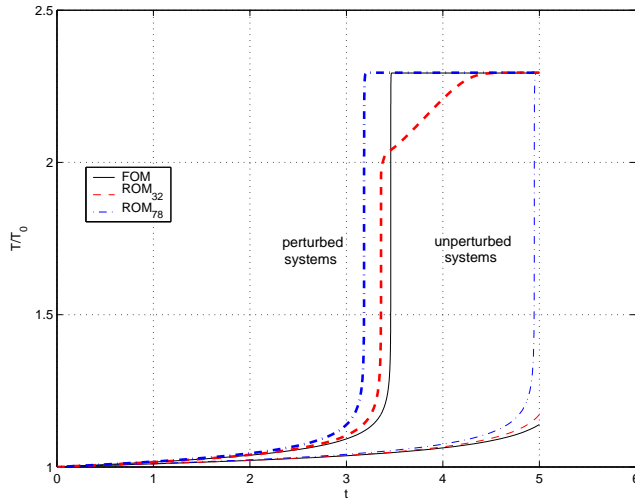


FIG. 6.3. *GRI combustion example: Evolution of the temperature over the time interval $[t_0, t_1]$ for the full model (solid line), ROM_{32} (dashed line), and ROM_{78} (dot-dashed line). Both unperturbed solutions (thin lines) and perturbed solutions (thick lines) are represented. The perturbation is in the direction of maximum error growth in the norm of the concentrations at time $t_1 = 5$ for the full model.*

ROM_{32} is better at predicting a perturbation-induced hastening of the ignition. This can also be seen in Fig. 6.3 which depicts the time evolution up to $t_1 = 5$ of the temperature for the perturbed and unperturbed, full and reduced-order models.

6.3. Heat shock model. This problem models the production of heat shock proteins in *Escherichia coli* cells as a response to extreme heat. These special proteins refold other cell proteins, essential to cell survival, that have become unfolded due to excessive heat. The optimization problem is to find a balance between the effect of the protective proteins and the cost of their production. The mathematical model of the heat shock response is reported in [14]. The model employs first order kinetics (law of mass-action) to describe the various components of the heat shock system, namely the synthesis of heat shock proteins and their regulator σ_{32} , in addition to protein folding and the association/dissociation activity of molecules. Reactions involving association/dissociation of molecules are often faster than synthesis and degradation of new molecules. Therefore the model exhibits a wide range of time scales and is subject to numerical stiffness. The differential equations that describe the fast states were transformed into algebraic constraints by means of a partial equilibrium approximation, yielding an index-one DAE system. This system describes the evolution of 11 differential variables, coupled to the evolution of 20 algebraic variables, and includes 31 parameters:

$$\begin{aligned} y' &= F(t, y, x, p^1, p^2) = 0 \\ G(t, y, x, p^2) &= 0 \\ y(0) &= y_0, \quad x(0) = x_0, \end{aligned}$$

where we have split the model parameters $p \in \mathbb{R}^{31}$ into two subsets, with the parameters $p^1 \in \mathbb{R}^{11}$ appearing only in the differential equations. The integration time

TABLE 6.3

Heat shock example: SV analysis results. Errors induced by parameter perturbations along the direction of maximum error growth are compared against errors induced by perturbations in random directions.

Full model			Reduced-order model		
δ in:	p^1	p^1 and p^2	δ in:	p^1	p^1 and p^2
\mathcal{E}	$8.2541 \cdot 10^4$	$5.6598 \cdot 10^4$	\mathcal{E}	$5.2887 \cdot 10^3$	$7.4958 \cdot 10^3$
$\ e(t, u_1)\ $	$6.0817 \cdot 10^4$	$6.7626 \cdot 10^4$	$\ \tilde{e}(t, \tilde{u}_1)\ $	$6.6292 \cdot 10^3$	$6.6292 \cdot 10^3$
$\ e(t, u_{rnd})\ $	$1.7648 \cdot 10^4$	$4.2042 \cdot 10^3$	$\ \tilde{e}(t, \tilde{u}_{rnd})\ $	$1.8461 \cdot 10^3$	$4.7902 \cdot 10^2$
	$1.9156 \cdot 10^4$	$2.2488 \cdot 10^4$		$1.3622 \cdot 10^3$	$1.3374 \cdot 10^3$
	$1.8371 \cdot 10^4$	$7.3998 \cdot 10^3$		$2.2759 \cdot 10^3$	$1.1942 \cdot 10^4$
	$1.9269 \cdot 10^3$	$9.5864 \cdot 10^3$		$1.4381 \cdot 10^3$	$1.2135 \cdot 10^3$

interval was taken such that the heat shock response occurs exactly at the middle of the time interval.

The reduced model is obtained in two steps. During the first step we reduce the number of differential variables from 11 to 5 by using an approximation which relates 6 variables (through a proportionality constant) to 2 other variables. The second step reduces further the number of variables by 15, taking advantage of the fact that the reduced variables can be well approximated by constants before and after the heat shock response. In the heat shock response region their behavior is highly nonlinear, but the length of this region is much smaller than the length of any neighboring region, allowing us to make the approximation.

We analyze the full and reduced-order models under perturbations in the model parameters. The following two cases are considered:

- Perturbation only in the subset p^1 of model parameters representing those parameters that do *not* appear in the algebraic equations.
- Perturbation in all model parameters p^1 and p^2 .

For both the full and reduced-order models, the direction of maximum error growth was estimated with the method described in §3 using 4 adjoint systems, i.e., 4 unit vectors in the SCE estimates. We considered perturbations either only in the subset p^1 of model parameters, or in all the model parameters. In both cases, all the states of the corresponding model were included in the perturbation-induced errors. Table 6.3 provides the estimated value \mathcal{E} compared against the norm of the error along the estimated direction of maximum error growth computed by integrating the forward error equation, $\|e(t, u_1)\|$. We also provide the norms of errors due to perturbations along a few random directions, $\|e(t, u_{rnd})\|$, to emphasize the fact that it is very unlikely to get close to the “worst-case” scenario by only using random perturbations. All perturbations were $0.01 \cdot u$ (with u either $u^* = u_1$ or some random unit vector).

Next, we computed the similarity indexes of §4 for this example. If we base the errors on all states of the ROM, the two indexes are:

	Perturbation in p^1	Perturbation in p^1 and p^2
\mathcal{S}_1	0.9788	0.4844
\mathcal{S}_2	0.1089	0.0882

Considering only the differential variables in the ROM in the error estimates, the reduced-order model is found to be closer (in the sense of these similarity indexes) to the full model:

TABLE 6.4

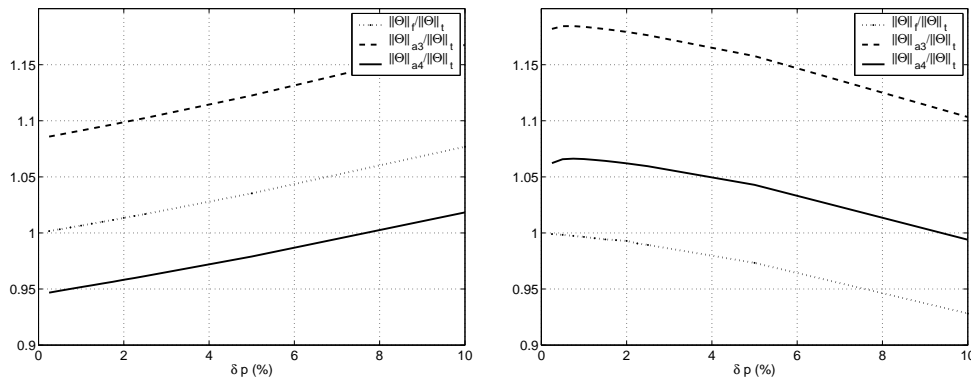
Heat shock example: Cumulative error due to model reduction and perturbations in parameters.

δp (%)	Perturbation in p^1			Perturbation in p^1 and p^2		
	$\ \Theta\ _{true}$	$\ \Theta\ _{fwd}$	$\ \Theta\ _{SCE}$	$\ \Theta\ _{true}$	$\ \Theta\ _{fwd}$	$\ \Theta\ _{SCE}$
0.25	$7.498 \cdot 10^2$	$7.510 \cdot 10^2$	$7.098 \cdot 10^2$	$8.437 \cdot 10^1$	$8.428 \cdot 10^1$	$8.962 \cdot 10^1$
0.50	$1.497 \cdot 10^3$	$1.502 \cdot 10^3$	$1.420 \cdot 10^3$	$1.722 \cdot 10^2$	$1.719 \cdot 10^2$	$1.835 \cdot 10^2$
0.75	$2.243 \cdot 10^3$	$2.254 \cdot 10^3$	$2.131 \cdot 10^3$	$2.601 \cdot 10^2$	$2.595 \cdot 10^2$	$2.774 \cdot 10^2$
1.00	$2.985 \cdot 10^3$	$3.005 \cdot 10^3$	$2.841 \cdot 10^3$	$3.493 \cdot 10^2$	$3.481 \cdot 10^2$	$3.723 \cdot 10^2$
1.25	$3.726 \cdot 10^3$	$3.756 \cdot 10^3$	$3.552 \cdot 10^3$	$4.367 \cdot 10^2$	$4.347 \cdot 10^2$	$4.652 \cdot 10^2$
1.50	$4.464 \cdot 10^3$	$4.508 \cdot 10^3$	$4.262 \cdot 10^3$	$5.253 \cdot 10^2$	$5.223 \cdot 10^2$	$5.590 \cdot 10^2$
1.75	$5.199 \cdot 10^3$	$5.259 \cdot 10^3$	$4.973 \cdot 10^3$	$6.141 \cdot 10^2$	$6.102 \cdot 10^2$	$6.529 \cdot 10^2$
2.00	$5.931 \cdot 10^3$	$6.011 \cdot 10^3$	$5.684 \cdot 10^3$	$7.032 \cdot 10^2$	$6.982 \cdot 10^2$	$7.468 \cdot 10^2$
2.25	$6.662 \cdot 10^3$	$6.762 \cdot 10^3$	$6.394 \cdot 10^3$	$7.920 \cdot 10^2$	$7.846 \cdot 10^2$	$8.402 \cdot 10^2$
2.50	$7.389 \cdot 10^3$	$7.514 \cdot 10^3$	$7.105 \cdot 10^3$	$8.822 \cdot 10^2$	$8.727 \cdot 10^2$	$9.346 \cdot 10^2$
5.00	$1.451 \cdot 10^4$	$1.502 \cdot 10^4$	$1.421 \cdot 10^4$	$1.796 \cdot 10^3$	$1.748 \cdot 10^3$	$1.873 \cdot 10^3$
10.00	$2.791 \cdot 10^4$	$3.005 \cdot 10^4$	$2.842 \cdot 10^4$	$3.774 \cdot 10^3$	$3.503 \cdot 10^3$	$3.751 \cdot 10^3$

	Perturbation in p^1	Perturbation in p^1 and p^2
\mathcal{S}_1	0.9998	0.7340
\mathcal{S}_2	0.1499	0.1213

Table 6.4 summarizes the analysis of the effect of parameter perturbations on the ROM through the estimation of the norm of the cumulative error due to model reduction and perturbations in parameters. The same information (in a slightly different form) is shown in Fig. 6.4. The norms of the following quantities were computed: the true (nonlinear) error ($\|\Theta\|_{true}$), the estimation using integration of the forward error equation ($\|\Theta\|_{fwd}$), and the SCE estimate based on $q = 4$ unit vectors ($\|\Theta\|_{SCE}$). All the ROM variables (which are a subset of the full model variables) were considered in these errors. The calculations were done for different perturbation magnitudes, given as a percentage of the norm of the perturbed parameters. The perturbation itself was along some random direction.

7. Conclusions. Our research extends the work presented in [10], focusing on estimating validity concepts for general nonlinear reduced order models (ROMs) of dynamical models, with application to systems described by ordinary differential (ODE) and differential algebraic equations (DAE). Our estimates, based on a combination of adjoint sensitivity analysis and the small sample statistical condition estimate (SCE), quantify how well a given ROM captures the changes in the output functional of the full system that result from perturbations in the problem parameters. We measure the extent to which the ROM preserves the value of maximum perturbation-induced error (and/or the direction of the perturbations that induce it) by considering two types of similarity indexes. The direction in parameter space of maximum error growth in the output functional is efficiently computed within the framework of singular vector analysis, using the SCE estimate for the error norm. Numerical examples (for both ODE and DAE systems) validate our approach. They also show that the similarity index concept is very effective for detecting situations where the ROM fails to capture the behavior of the full model when the parameters are slightly perturbed.



(a) Only those parameters *not* affecting the algebraic equations were perturbed.

(b) All model parameters were perturbed.

FIG. 6.4. *Heat shock example: Cumulative error due to model reduction and perturbations in parameters. The plots show the ratio of the estimated norm of the cumulative error and the norm of the true (nonlinear) cumulative error for different magnitudes of parameter perturbations along some random direction.*

REFERENCES

- [1] A. ANTOUNAS AND D. SORENSEN, *Approximation of large-scale dynamical systems: An overview*, Tech. Report TR0101, Rice University, 2001.
- [2] U. M. ASCHER AND L. R. PETZOLD, *Stability of computational methods for constrained dynamics systems*, SIAM J. Sci. Comput., 14(1) (1993), pp. 95–120.
- [3] K. E. BRENNAN, S. L. CAMPBELL, AND L. R. PETZOLD, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, SIAM, Philadelphia, Pa, 1996.
- [4] Y. CAO, S. LI, L. R. PETZOLD, AND R. SERBAN, *Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution*, SIAM J. Sci. Comput., 24(3) (2003), pp. 1076–1089.
- [5] Y. CHEN, D. BATTISTI, T. PALMER, J. BARSUGLI, AND E. SARACHIK, *Study of the predictability of tropical Pacific SST in a coupled atmosphere-ocean model using singular vector analysis: The role of the annual cycle and the ENSO cycle*, Mon. Weather Rev., 125 (1997), pp. 831–845.
- [6] P. FAUCHER, M. GAVART, AND P. D. MEY, *Isopycnal empirical orthogonal functions (EOFs) in the North and tropical Atlantic and their use in estimation problems*, J. Geophys. Res.-Oceans, 107 (2002), p. 3107.
- [7] T. GUDMUNDSSON, C. KENNEY, AND A. LAUB, *Small-sample statistical estimates for matrix norms*, SIAM J. Matrix Anal. Appl., 16(3) (1995), pp. 776–792.
- [8] S. GUGERCIN AND A. C. ANTOUNAS, *A survey of model reduction by balanced truncation and some new results*, Int. J. Control, 77(8) (2004), pp. 748–766.
- [9] E. HAIRER AND G. WANNER, *Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems*, Springer-Verlag, Berlin, 1991.
- [10] C. HOMESCU, L. PETZOLD, AND R. SERBAN, *Error estimation for reduced order models of dynamical systems*, SIAM J. Numer. Anal., submitted (2004).
- [11] I. T. JOLLIFFE, *Principal Component Analysis*, Springer-Verlag, 2002.
- [12] C. KENNEY AND A. LAUB, *Small-sample statistical condition estimates for general matrix functions*, SIAM J. Sci. Comp., 15(1) (1994), pp. 36–61.
- [13] K. KUNISCH AND S. VOLKWEIN, *Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition*, J. Optimiz. Theory App., 102(2) (1999), pp. 345–371.
- [14] H. KURATA, H. EL-SAMAD, T. YI, M. KHAMMASH, AND J. DOYLE, *Feedback regulation of the Heat shock response in E. Coli*, in 40th IEEE Conference on Decision and Control, 2001, pp. 837–842.
- [15] H. KURODA AND M. KISHI, *A data assimilation technique applied to estimate parameters for the NEMURO marine ecosystem model*, Ecological Modelling, 172(1) (2004), pp. 69–85.
- [16] Z. LI, I. NAVON, AND M. HUSSAINI, *Analysis of the singular vectors of the full-physics fsu global*

- spectral model*, Accepted for publication in *Tellus*, (2004).
- [17] E. LORENZ, *A study of the predictability of a 28 variable atmospheric model*, *Tellus*, 17 (1965), pp. 321–333.
 - [18] L. PETZOLD AND W. ZHU, *Model reduction for chemical kinetics: An optimization approach*, *AIChE J.*, 45(4) (1999), pp. 869–886.
 - [19] M. RATHINAM AND L. PETZOLD, *A new look at proper orthogonal decomposition*, *SIAM J. Numer. Anal.*, 41(5) (2003), pp. 1893–1925.
 - [20] P. TUPPER, *Adaptive model reduction for chemical kinetics*, *BIT*, 42(2) (2002), pp. 447–465.