

Model Reduction of the Nonlinear Complex Ginzburg–Landau Equation*

Miloš Ilak[†], Shervin Bagheri[†], Luca Brandt[†], Clarence W. Rowley[‡], and Dan S. Henningson[†]

Abstract. Reduced-order models of the nonlinear complex Ginzburg–Landau (CGL) equation are computed using a nonlinear generalization of balanced truncation. The method involves Galerkin projection of the nonlinear dynamics onto modes determined by balanced truncation of a linearized system and is compared to a standard method using projection onto proper orthogonal decomposition (POD) modes computed from snapshots of nonlinear simulations. It is found that the nonlinear reduced-order models obtained using modes from linear balanced truncation capture very well the transient dynamics of the CGL equation and outperform POD models; i.e., a higher number of POD modes than linear balancing modes is typically necessary in order to capture the dynamics of the original system correctly. In addition, we find that the performance of POD models compares well to that of balanced truncation models when the degree of nonnormality in the system, in this case determined by the streamwise extent of a disturbance amplification region, is lower. Our findings therefore indicate that the superior performance of balanced truncation compared to POD/Galerkin models in capturing the input/output dynamics of linear systems extends to the case of a nonlinear system, both for the case of significant transient growth, which represents a basic model of boundary layer instabilities, and for a limit cycle case that represents a basic model of vortex shedding past a cylinder.

Key words. model reduction, balanced truncation, Ginzburg–Landau equation

AMS subject classifications. 76E15, 78M34

DOI. 10.1137/100787350

1. Introduction. Model reduction plays an important role in developing effective strategies for practical applications in fluid flow control, since the dynamical systems that describe most flows are discretized PDEs with a very large number of degrees of freedom—on the order of 10^5 or higher. Considerable progress has been made in recent years in modeling and control of several types of fluid flows using the methods of linear control theory (see, for example, [19, 11, 15, 4, 2, 5]). While these efforts have been successful, they are dependent on linear models, meaning that the controllers are often effective only for small disturbances when applied to nonlinear flows, i.e., in the neighborhood of the equilibrium about which the flow has been linearized. Obtaining high-fidelity nonlinear reduced-order models is therefore also of interest from a control perspective. Nonlinear reduced-order models of fluid flows have been computed extensively using proper orthogonal decomposition (POD) in the past few decades in order to study the dynamics of these flows (see, for example, [3, 23, 32]). These models,

*Received by the editors March 2, 2010; accepted for publication (in revised form) by D. Barkley September 14, 2010; published electronically DATE. This work was partially supported by National Science Foundation grant CMS-0347239 and by the Swedish Research Council (VR).

<http://www.siam.org/journals/siads/x-x/78735.html>

[†]Linné FLOW Centre, Department of Mechanics, Royal Institute of Technology, Stockholm, Sweden (ilak@mech.kth.se, shervin@mech.kth.se, luca@mech.kth.se, hennin@mech.kth.se).

[‡]Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544 (cwrowley@princeton.edu).

however, have been shown to require considerable fine tuning and often do not fully capture the dynamics of the original flow. In this work, we compare the performance of reduced-order models of a simple nonlinear one-dimensional (1D) PDE model of fluid flows obtained using POD and balanced truncation, a standard method from linear control theory that has been shown to have more desirable properties than POD for control-oriented model reduction.

The complex Ginzburg–Landau (CGL) equation is a 1D nonlinear PDE that exhibits a Hopf bifurcation. Extensive work over the past few decades has established the CGL equation as a standard 1D model for some fluid flows, as the Hopf bifurcation may be thought of as a simple model for vortex shedding past a cylinder, and the spatially developing solution in the subcritical regime is a basic model of instabilities typical for boundary layer flows (see, for example, [7, 8, 6]). Some concepts from modern stability analysis of fluid flows will be reviewed in section 2, and the relevance of the CGL equation as a model for the two classes of flows will be explained further. The modeling and control of the linear part of the CGL equation has been investigated in [5]. The subject of this work is the nonlinear case, in which the CGL equation saturates to a limit cycle for supercritical values of the bifurcation parameter.

Balanced truncation, a standard method of model reduction in control theory, was introduced first for linear, stable, time-invariant systems [22]. Extensions of the method have been developed for unstable systems by [35] and periodic systems by [28]. The method has also recently been applied with success to high-dimensional flows using an approximate method called balanced proper orthogonal decomposition (BPOD) by [27], and it has been demonstrated that the performance of the method is vastly superior to POD in [15, 4, 2, 21]. The theory of balanced truncation for nonlinear systems was first developed by [29], and extensive work on the subject has been done recently (see, for example, [33, 10]); however, “first principles” approaches to nonlinear balanced truncation are not computationally feasible for large systems such as fluid flows. An alternative snapshot-based method has been proposed by [16], but the difficulty with this method is that it too becomes intractable for high-dimensional systems and the selection of initial conditions for the snapshots is ad hoc [14]. On the other hand, linear controllers have been used for the nonlinear CGL equation in [18] without model reduction. It is therefore of interest to test whether nonlinear reduced-order models using modes obtained from linear balanced truncation are useful in model reduction of nonlinear systems.

The main contribution of this work is that for the first time, at least to the best of the authors’ knowledge, a basis obtained using balanced truncation of the linear part of a nonlinear operator is used to compute nonlinear reduced-order models of the original system via Galerkin projection. We find that these models work well and outperform POD, indicating that balanced truncation may be potentially useful in other problems with similar dynamics. In the case of the CGL equation, the initial transient growth of the solution is due to a region in the domain that amplifies disturbances. This is a linear mechanism due to the nonnormality of the linear part of the operator, and the cubic nonlinearity only adds damping to the system. Balanced truncation of the linear part of the equation therefore captures very accurately that linear mechanism and reproduces well the dynamics of the original system even in the nonlinear case. We also find that POD models perform well for a set of parameters for which the nonnormality in the linear part of the CGL equation is not as pronounced as in the main cases of interest. Our findings indicate that linear balanced truncation, which is readily

computed using standard algorithms, and whose advantages over other methods in capturing the input/output behavior of linearized fluid flows have been established, may also be used for nonlinear systems, including those with a Hopf bifurcation.

This paper is organized as follows. In section 2 we describe briefly the dynamics of the nonlinear CGL equation and its relevance as a simple model of more complex flows. The model reduction methods used in this work are reviewed in section 3. Finally, the results for the two cases we consider are described in section 4, followed by conclusions and directions for future work in section 5.

2. The complex Ginzburg–Landau equation. The CGL equation is a 1D PDE of convection-diffusion type [6, 8, 5], with an extra term representing linear exponential disturbances, defined on the infinite domain $x \in \{-\infty, \infty\}$, and it is given by

$$(2.1) \quad \frac{\partial q}{\partial t} = Aq + f(q),$$

where the linear part is given by

$$(2.2) \quad Aq = \left(-\nu \frac{\partial}{\partial x} + \gamma \frac{\partial^2}{\partial x^2} + \mu \right) q.$$

Since the spatial domain is infinite, there are no boundary conditions, but the complex solution $q(x, t)$ is required to remain finite as $x \rightarrow \pm\infty$. The complex convective velocity is defined as $\nu = U + 2ic_u$, $\gamma = 1 + ic_d$ is a diffusion parameter, and the parameter μ is defined as

$$(2.3) \quad \mu(x) = (\mu_0 - c_u^2) + \mu_2 \frac{x^2}{2}.$$

The spatially varying parameter $\mu(x)$ models the presence of exponential disturbances, and in this form defines a region

$$(2.4) \quad -\sqrt{-2(\mu_0 - c_u^2)/\mu_2} < x < \sqrt{-2(\mu_0 - c_u^2)/\mu_2}$$

in which disturbances are amplified. The upstream and downstream ends of the unstable region are called branch I and branch II, respectively, following standard terminology [5]. The nonlinear term $f(q)$ is given by

$$(2.5) \quad f(q) = -a|q|^2q,$$

where a is real and positive. This nonlinear term corresponds to the normal form for a Hopf bifurcation, and a Landau amplitude equation which determines the stability of the solutions depending on the bifurcation parameter μ_0 can be derived, as shown in [6]. This form of nonlinearity leads to a limit cycle in the cases where the linear part of the problem is globally unstable; i.e., there are unstable eigenvectors of the linear operator whose growth is counteracted by the nonlinear term.

With appropriate inputs and outputs and appropriate discretization of the infinite-dimensional operator A in space, the full nonlinear system we consider in this work is given by

$$(2.6) \quad \begin{aligned} \dot{q} &= Aq + f(q) + Bu, \\ y &= Cq. \end{aligned}$$

This system is in a standard linear state-space form (see, for example, [25] or other standard texts on linear systems or control) with the additional nonlinear term. The matrix B represents an input, which is given in time by $u(t)$. The output $y(t)$ represents a vector of measurements, extracted from the full solution $q(x, t)$ using a measurement matrix C . In this work we consider only a single output; i.e., the C matrix is a row vector.

2.1. The CGL equation as a model for spatially developing flows. Results on reduced-order modeling and control of the linear part of the CGL equation are given in [5] along with a review of the fundamental dynamics of fluid flows that the equation models. Here we outline some main features of the dynamics of the CGL equation that make it a suitable, although very simple, model for two important classes of fluid flows. More details can be found in [7, 8, 6].

Without the second term in the expression for the coefficient $\mu(x)$, defined in (2.3), the flow would be known as parallel; i.e., there would be no spatial dependence of the coefficients of the governing equation. The spatial dependence of $\mu(x)$ makes the flow *nonparallel*, and in that respect the CGL equation models many important fluid flows, including a spatially developing boundary layer. An important difference between the CGL equation and a true boundary layer flow is that no new wavenumbers are introduced by the nonlinear term, which can easily be shown. Rather, the nonlinear term reproduces the saturation of the disturbance level, where the damping is proportional to the magnitude of the solution. Nonetheless, the existence of the x -dependent amplification region allows for the spatial evolution of perturbations that grow within the region and then convect out of the domain of interest, which is the fundamental instability mechanism in boundary layer flows. A flow where only such perturbations are present is also known as *convectively unstable* [6].

The *transient* behavior of a disturbance to a base flow (steady-state solution) in fluid systems plays an important role in the dynamics. For many fluid systems, the phenomenon of *transient growth*, which is due to the nonnormality of the system (see, for example, [30]), is observed independently of whether the disturbance is ultimately exponentially unstable from a linear point of view (i.e., whether there are eigenvalues of the linearized operator A with positive real part). Therefore, there are often two competing mechanisms in transition—that of nonnormal growth and that of exponential growth. The CGL equation exhibits both of these mechanisms for some parameter sets, which makes it a relevant simple model for fluid flows.

Transition in boundary layer flows in the absence of exponentially growing perturbations is explained by the large transient growth of linear disturbances, up to the point where nonlinearities become significant and transition occurs. In the regime that we will define as *subcritical*, which is characterized by the absence of unstable eigenvectors of the operator A , the CGL equation is an example of such a flow and models the convective instability mechanism of boundary layer flows, as mentioned above.

Above a certain value of the critical parameter μ_0 , the CGL equation enters a limit cycle, where the dynamics is dominated by one or more linearly unstable modes. We will refer to this case as *supercritical* hereinafter, as opposed to the convectively unstable case that exhibits transient growth only. This supercritical case can also be categorized as *absolutely unstable* and is analogous to vortex shedding behind a cylinder in crossflow [26, 13], although the 1D CGL equation dynamics is much simpler. From the perspective of reduced-order

modeling, it is desirable to capture the transient behavior where both nonnormal growth and an exponential instability are present, and it is particularly important in this case to obtain nonlinear reduced-order models. Therefore, we will compute reduced-order models for both the subcritical, convectively unstable case and the supercritical case, in which a limit cycle is observed.

3. Model reduction via Galerkin projection. In this section we review the two methods used for the computation of the reduced-order models in this work. These methods are standard, and the references should be consulted for more information. Galerkin projection is a standard method, where reduced-order models are obtained via a projection of the dynamics of a high-dimensional system onto a lower-dimensional space. The choice of this low-dimensional space (basis) is key for model performance. In practice one often uses *empirical* bases obtained from simulations of the studied system or experimental measurements, as opposed to complete bases such as Fourier series, or eigenvectors of the linearization of the governing operator. Two common choices for empirical low-dimensional bases are modes obtained from POD and balanced truncation, and we review these methods next.

3.1. Proper orthogonal decomposition. POD is a standard method for model reduction, used in both linear and nonlinear systems. We refer the reader to standard references, such as [31, 12], for more details. Mathematically, POD modes are the eigenfunctions of the autocorrelation matrix of the states $q(x, t)$ of a system such as the one in (2.6) integrated over time [12]. The modes can be computed by stacking snapshots $q(t_k)$ at some times t_k into a matrix X ,

$$(3.1) \quad X = [q(t_1) \ q(t_2) \ \dots \ q(t_m)],$$

where m is the total number of snapshots, and solving the $m \times m$ eigenvalue problem,

$$(3.2) \quad X^T X U = U \Lambda,$$

the orthonormal POD modes being the columns of the matrix:

$$(3.3) \quad \Theta = X U \Lambda^{-1/2}.$$

This method is also known as the method of snapshots [31], and it is numerically tractable for very large systems, since the number of snapshots is typically much smaller than the dimension of the system. POD modes have a particularly intuitive meaning for fluid flows, since they represent the most energetic structures in a given simulation if velocity snapshots are taken; the set of POD modes is the optimal solution to the problem of finding a low-order basis of given dimension that captures the largest fraction of the kinetic energy in the simulation snapshots [12]. We emphasize here that, while POD is optimal at capturing a given dataset, it may not be optimal in capturing the dynamics of the underlying system when reduced-order models are obtained using Galerkin projection [32, 27, 15].

3.2. Balanced truncation of linear systems. Balanced truncation is a standard method in control theory, introduced by Moore [22] for linear, stable, time-invariant systems. From a physical point of view, it may be thought of as a balance between capturing the sensitivity of

the system to inputs with the potential of system states to affect the output. Thus, a reduced-order model computed using balanced truncation captures well the input-output behavior of the original system. Balanced truncation is computed by simultaneously diagonalizing the controllability and observability Gramians of a state-space system, which are given by

$$(3.4) \quad W_c = \int_0^\infty e^{At} B B^T e^{A^T t} dt, \quad W_o = \int_0^\infty e^{A^T t} C^T C e^{At} dt,$$

where in our case A , B , and C are the matrices from the system given in (2.6). A standard method for computing the coordinate transformation that simultaneously diagonalizes the Gramians is given by [17]. In a similar method, described in [27], this diagonalization is achieved using a set of *balancing modes* Φ (also referred to as *direct modes*) and a set of the corresponding *adjoint modes*, Ψ , which can also be computed from the snapshots of the impulse responses of the linear system and its corresponding adjoint system when the method of [17] is intractable, i.e., if the original system is very large. If the Gramians are factorized as

$$(3.5) \quad W_c = X X^T, \quad W_o = Y Y^T,$$

these sets of modes are obtained using the singular value decomposition (SVD)

$$(3.6) \quad Y^T X = U \Sigma V^T,$$

where U and V are unitary matrices and Σ is a diagonal matrix. The balancing transformation Φ and its inverse Ψ^T are then found by

$$(3.7) \quad \Phi = X V \Sigma^{-1/2}, \quad \Psi = Y U \Sigma^{-1/2}.$$

As described in [27], the two bases of modes

$$(3.8) \quad \Phi = [\phi_1 \ \phi_2 \ \dots \ \phi_n], \quad \Psi = [\psi_1 \ \psi_2 \ \dots \ \psi_n],$$

where n is the dimension of the original system,¹ are biorthogonal; i.e.,

$$(3.9) \quad \langle \psi_i, \phi_j \rangle = \delta_{ij},$$

with the inner product $\langle v_1, v_2 \rangle$ defined as

$$(3.10) \quad \langle v_1, v_2 \rangle = \int_{\Omega} v_1 v_2 d\Omega,$$

where Ω is the spatial domain in which the original system evolves. Thus, even though balanced truncation and POD may result in modes that span the same subspace, the resulting models may have quite different performances, since in balanced truncation the projection

¹This is the case only if the Gramians have full rank, as discussed in [27]; otherwise only the leading p modes are obtained, where p is the rank of $Y^T X$, and the higher modes obtained are not true balancing modes. However, typically p is sufficiently large; i.e., the rank r of the reduced-order model we want to compute is smaller than p .

is done along a different direction—the direction orthogonal to the space spanned by the adjoint modes. We note that, for linear systems, the controllability Gramian is exactly the autocorrelation tensor whose eigenvectors are the POD modes of the systems. However, POD does not take into account the observability of the system, which is captured by balanced truncation due to the inclusion of the observability Gramian. In the balanced coordinates, the controllability and observability Gramians of the new system are equal and diagonal, which means that the most controllable states (the states responding best to forcing) are also the most observable states (the states that affect the system outputs the most).

For systems with a relatively small number of degrees of freedom, the balancing transformation and its inverse may easily be computed using standard algorithms, such as, for example, that of [17]. An extension of the method for unstable systems was introduced in [35], and we use this method. This method has also been extended to very large systems [2]. For the nonlinear CGL equation, we compute balanced truncation of the *linear* part of the CGL equation and use the resulting basis for Galerkin projection.

3.3. Nonlinear models of the CGL equation. The first step in model reduction via Galerkin projection is the expansion of the field q onto a set of modes as $q = \Phi z$. Then the *rank* of the model, defined as r , is chosen, where typically $r \ll n$, n being the dimension of the original system. Then, taking the inner product of the equations with the reduced basis consisting of the first r adjoint modes, Ψ_r , the CGL equation is projected as follows:

$$(3.11) \quad \begin{aligned} \dot{z} &= \Psi_r^T A \Phi_r z - a(\Psi_r^T |\Phi_r z|^2 (\Phi_r z)) + \Psi_r^T B u, \\ y &= C \Phi_r z, \end{aligned}$$

where we have used (2.5) to expand the nonlinear term in terms of the basis Φ_r . We thus obtain a set of r ordinary differential equations. We note that the basis corresponding to the adjoint modes Ψ_r for POD is exactly the basis of the direct modes ($\Phi_r = \Psi_r$ for POD), and we will refer to the resulting models as projections onto POD modes hereinafter. When using the direct and adjoint modes from balanced truncation of the linear part of the system in (2.5), we will refer to the procedure as a projection onto balancing modes hereinafter.²

4. Results. The parameter sets for the linear part of the CGL equation used in this work are the same as those in [5] for the two cases we consider, and they are given in Table 1. As discussed in section 2, the subcritical case is a model of the convective instability exhibited by boundary layers, and the supercritical case is a model of vortex shedding past a cylinder in crossflow. The amplitude of the nonlinear term is $a = 0.1$ for both cases. The CGL equation and the reduced-order models are discretized in space using Hermite polynomials and integrated in time using the Crank–Nicholson scheme, and the spatial resolution is $n_x = 220$ in each case. The dimension of the state of the original system is thus 440, since the equation is complex. All computations were done in MATLAB, and some routines from the differentiation toolbox in [34] were utilized. The reduced-order models were integrated in time using the same numerical scheme as the full CGL equation.

²We choose the term for brevity, and it is implied that we are projecting the original nonlinear equation onto the set of projection onto balancing modes obtained using linear balanced truncation.

Table 1*The parameters of the CGL equation used in the two cases.*

Case	ν	γ	μ_0	μ_2
Subcritical	$2 + 0.4i$	$1 - i$	0.38	-0.01
Supercritical	$2 + 0.4i$	$1 - i$	0.41	-0.01

The choices of system inputs and outputs are the same as those in [5]. In both cases the input to the system (the B matrix in (2.6)) is an optimal perturbation for the linear CGL equation, i.e., an initial condition that will result in the largest energy amplification of the solution field q in a given time interval. Such perturbations are treated in detail, for example, in [30] and computed for the particular case of the linear CGL equation in [5]. The measurement (output) matrix C is a Gaussian function centered at the branch II (the downstream end of the unstable region). The particular form of the measurement here is

$$(4.1) \quad C = \exp \left[- \left(\frac{x - x_2}{s} \right)^2 \right]^T M,$$

where M is the matrix of quadrature weights for the discretization, and x_2 is the location of branch II. The width of the Gaussian was chosen to be $s = 1.6$. The choice of an optimal perturbation ensures that the nonnormality in the system is exhibited in the response due to the transient energy growth. The measurement is located at branch II because this is the location in the domain where the perturbation has achieved maximum growth in space, which can be seen in Figure 2(a). The motivation for these inputs and outputs is the potential use of models for control design—in a full control system setup, we would attempt to control the optimal perturbation using an actuator, and the output we would try to minimize would be the measurement at branch II.

4.1. Convectively unstable case. We begin by considering the convectively unstable case, where an initial perturbation is amplified by the unstable region defined by (2.4) and convects out of the simulation domain downstream. The optimal perturbation, which is computed using the method given in [5], and used as the system input (B matrix), is shown in Figure 1(a).

Figure 2(a) shows the evolution of the disturbance shown in Figure 1. The disturbance initially convects downstream to branch II of the unstable region, after which the oscillating wave around that point in the flow slowly decays, as shown in Figure 2(a). The energy norm exhibits initial nonnormal growth typical of the linear mechanism that governs the flow, as shown in Figure 2(b), where the energy growth is compared for the linear and nonlinear evolutions of the optimal perturbation. It has been demonstrated that in the case of the linear convectively unstable CGL equation balanced truncation outperforms POD significantly [5]. In this work, although overall the dynamics of the convectively unstable case is very similar to that of the linear convectively unstable CGL equation, a nonlinear term is present, and we thus investigate if a projection onto balancing modes outperforms a projection onto POD modes.

Next, the POD modes, as well as the direct and adjoint modes for balanced truncation, are computed. POD modes are computed using the method of snapshots from 1000 snapshots

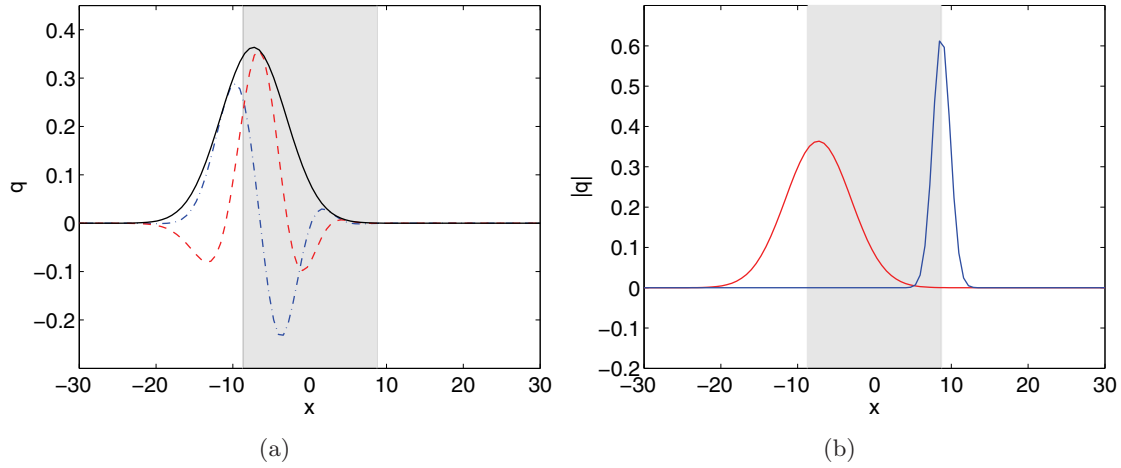


Figure 1. (a) The real part (dash-dotted line), imaginary part (dashed line), and the absolute value (solid line) of the optimal disturbance used as the input (B matrix in (2.6)) in the convectively unstable case. (b) The absolute value of the B matrix (red) and C matrix (blue). In both (a) and (b), the grey region indicates the region defined by (2.4) where disturbances are amplified.

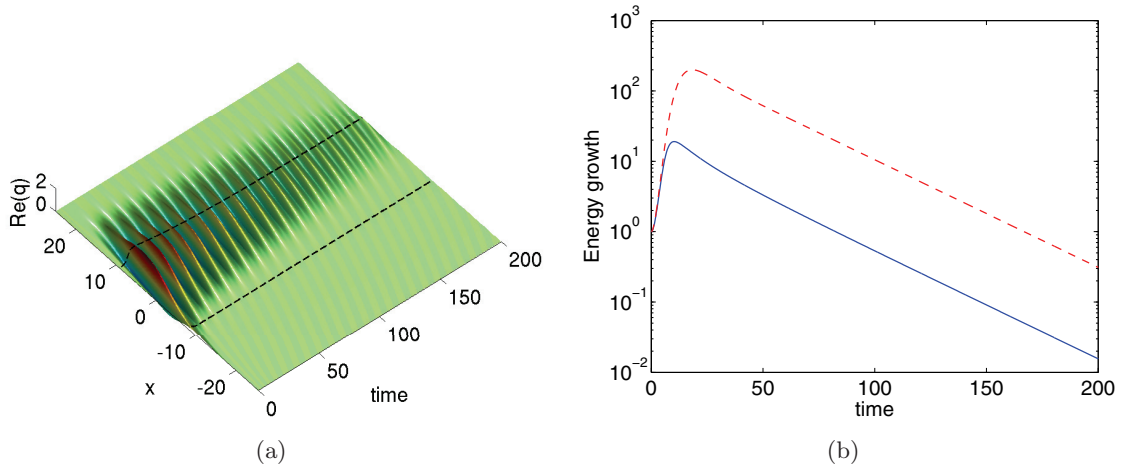


Figure 2. (a) The spatiotemporal nonlinear evolution of the optimal disturbance in the subcritical case (real part). (b) The energy norm for the linear (red dashed line) and nonlinear evolution of the CGL equation (solid line) for the same case. The transient nonnormal growth of the nonlinear solution is considerably smaller compared to the linear case due to the nonlinear damping.

equally spaced in time, the time interval being from $t = 0$ until $t = 800$, at which point in time the energy of the perturbation has decayed to negligible levels compared to the initial energy. The method of [17] was used to compute balanced truncation of the linear part of the CGL equation and obtain the modes for the Galerkin projection of the nonlinear system given by (3.11).

Figure 3 shows the performance of projection onto balancing modes and POD reduced-order models of ranks 5 and 10, corresponding to 10 and 20 degrees of freedom (respectively,

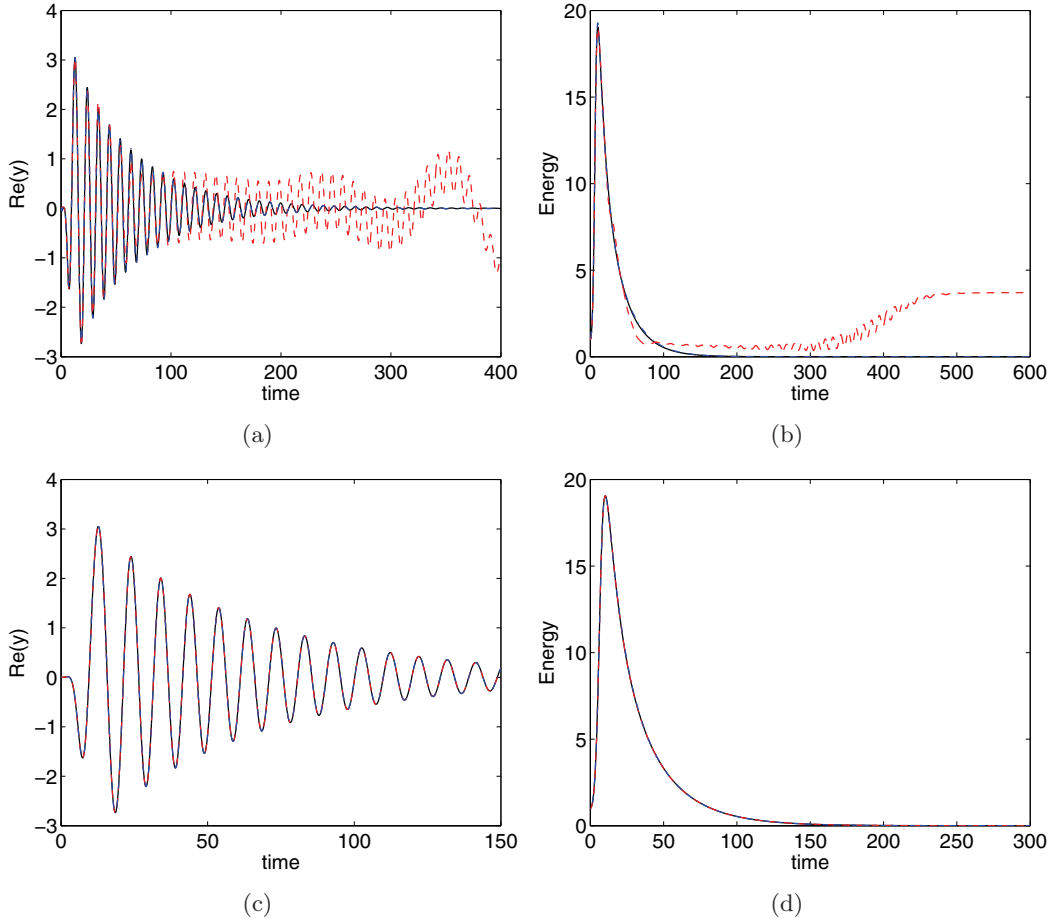


Figure 3. (a) The real part of the output y for models of rank 5 from projection onto balancing modes (blue dash-dotted line) and POD (red dashed line), compared to the output of the full system (solid black line). (b) A comparison of the energy norm $\|q\|^2$ for the full solution and these models (same legend). In both (a) and (b), the lines representing projection onto balancing modes are almost indistinguishable visually from the true solution. The performance of both projection POD modes and projection onto balancing modes at rank 10, shown in (c) and (d), where all three lines are indistinguishable, is excellent. Note that the time intervals shown in different plots differ in order to illustrate the performance clearly.

five and ten complex modes), as compared to the 440 degrees of freedom in the original system. We see that at rank 5 the projection onto balancing modes captures very well the time evolution of both the single output y and the total energy of the flow field. On the other hand, the projection onto POD modes captures only the initial transient, but the solution does not decay and enters a limit cycle instead, as shown by Figure 3(b). At rank 10, there is no appreciable difference between the performance of the projection onto balancing modes and the projection onto POD modes. Thus, as expected for this system with fairly simple dynamics, the projection onto POD modes performs well if a sufficient number of modes (in this particular case, 10) is included. This number is typically much higher than the number of balanced truncation/BPOD modes required for good performance, which has also been

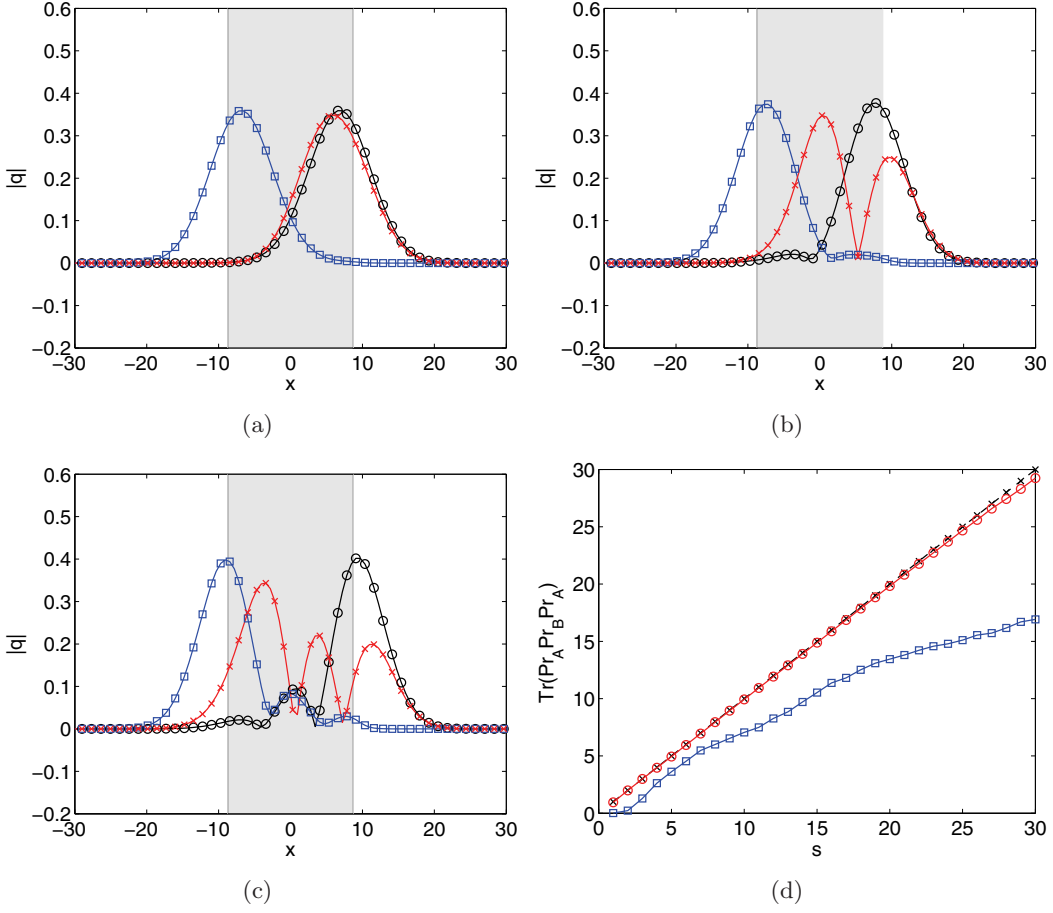


Figure 4. The absolute values of the first (a), second (b), and third (c) POD mode (red $-X-$), and the direct (black $-O-$) and adjoint (blue $-□-$) modes from balanced truncation. (d) Subspace comparison using the method of traces, comparing the subspaces of POD and balancing modes (red $-O-$) and POD and adjoint balancing modes (blue $-□-$) (see text for explanation).

observed in [15] for channel flow and in [2] for flow past a flat plate at an angle of attack, as well as for the linear CGL equation in [5]. What is more important is that the projection onto balancing modes performs well even at low rank; i.e., the superior performance in terms of the input-output dynamics typical for linear balanced truncation models is also observed for the projection of the nonlinear CGL equation onto linear balancing modes.

In order to explain the difference in model performance at low rank, we next compare the two bases of modes. Figure 4(a)–(c) shows the absolute values of the first three balanced truncation modes, the corresponding adjoint modes, and the POD modes. The first POD mode is quite similar to the first balancing mode, and, while the higher modes look different, they have approximately the same spatial support. On the other hand, the peaks of the adjoint modes resulting from balancing are located near branch I and have quite different spatial support. Since these adjoint modes are used for the nonorthogonal projection, the performance of the models is different.

It is of interest to make an overall comparison of the subspaces resulting from the model reduction procedure in order to explain the difference in model performance. Using the method of computing the trace of the product of projection operators [9, 15], the POD basis is compared to the direct and adjoint modes from the balanced truncation in Figure 4(d). The idea of a comparison between two subspaces, as opposed to individual basis vectors in these subspaces, arises from the field of iterative subspace methods where the convergence of an algorithm is determined by comparing the result to the previous iteration. In [9], the subspaces resulting from different POD calculations are compared using the method of traces. In particular, if Pr_A and Pr_B are the projectors corresponding to two subspaces, A and B , both with dimension s , the subspaces spanned by two sets of basis functions can be compared by computing the trace of the matrix $Pr_A Pr_B Pr_A$. If the subspaces are completely different (say, in three dimensions the basis vectors e_1 and e_2), the value of the trace will be zero, while if the subspaces are identical, the value will be exactly the dimension of the subspaces s .

It turns out that about 20 of the leading balancing modes span the same subspace as the corresponding POD modes: even though the individual modes do not look identical in Figure 4, the trace calculation shows that they span nearly identical subspaces. (Recall that the POD modes must be orthogonal to each other, while the balancing modes need not be.)

It has been observed in [15] that although POD modes and direct modes from balanced truncation may span similar subspaces, the dynamics of the resulting models can be quite different due to the nonorthogonal projection using adjoint modes in balanced truncation. Figure 4(d) also shows that the subspace spanned by the adjoint modes is quite different from the one spanned by the POD modes, which illustrates this difference. Thus, we may conclude that in this case the characteristic superior performance of linear balancing, which may be attributed to the nonorthogonal projection, also extends to model reduction of the nonlinear CGL equation.

4.2. Limit cycle case. Next, we consider the supercritical case, where the value of the parameter μ_c is above its critical value, and a limit cycle develops after the initial transient. Since the linear part of the CGL equation is unstable here, the method of [35] is used in order to compute balanced truncation. In addition to the POD modes computed in the same time interval as in the subcritical case, in this case we also compute POD only from snapshots of the limit cycle, i.e., when the effects of the transient are not visible anymore. This POD basis consists of only one complex mode, which we use in order to attempt to recreate the limit cycle in a model with a single complex mode. We note that in this case, after the transients have decayed, the flow is simple enough to be described by only one complex mode, in contrast to, for example, the cylinder flow investigated in [23], where eight POD modes are needed to capture the limit cycle due to the harmonics of spatial frequencies and symmetries in the two-dimensional flow.

Figure 5(a) shows the limit cycle that develops in the case of a supercritical value of the parameter μ_0 . Figure 5(b) also shows the response of the linear part of the CGL equation for the same set of parameters; there is initial fast linear transient growth, followed by unbounded exponential growth of an unstable mode at slower rate. It is this unstable mode that sustains the limit cycle in the nonlinear case.

Figure 6 shows a comparison of the first POD mode from the basis computed over the entire duration of the impulse response (from $t = 0$ until $t = 800$), the first balancing mode

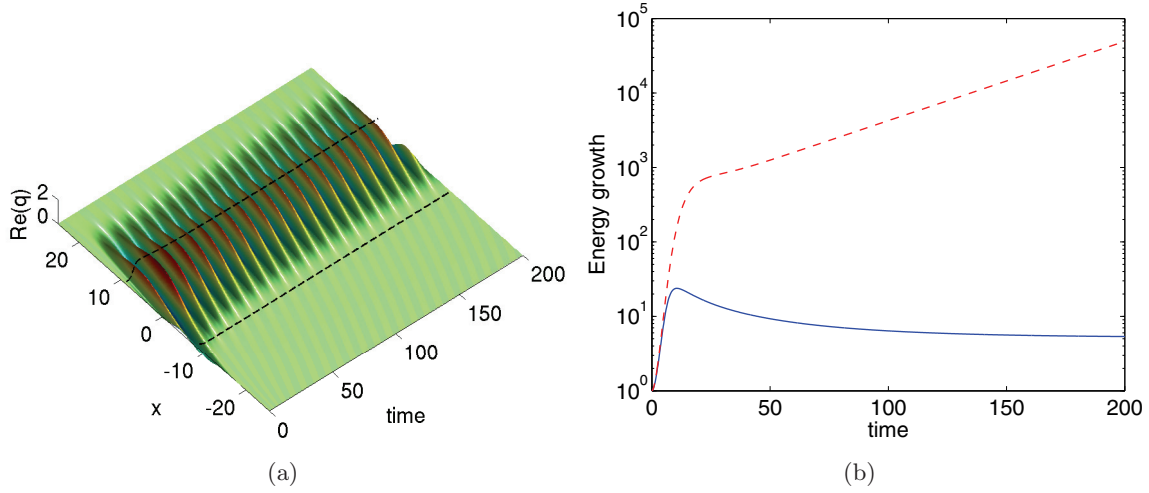


Figure 5. (a) The spatiotemporal nonlinear evolution of the optimal disturbance in the limit cycle case (real part). (b) The energy norm for the linear (red dashed line) and nonlinear (solid line) evolutions of the CGL equation for the same case. In the linear case, the energy grows exponentially due to the unstable mode after the initial transient, while the limit cycle develops in the nonlinear case.

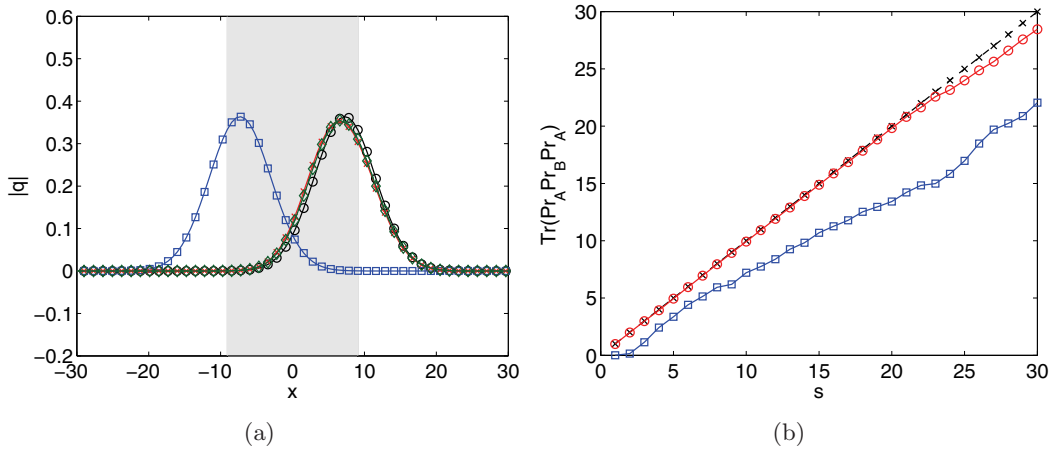


Figure 6. (a) The absolute values of the first POD mode (red $-X-$), as well as the direct (black $-O-$) and adjoint (blue $-□-$) modes from balanced truncation for the limit cycle case. The absolute value of the leading POD mode computed only from snapshots of the limit cycle is also shown (green $-◇-$). (b) Subspace comparison using the method of traces, comparing the subspaces of POD and balancing modes (red $-O-$) and POD and adjoint balancing modes (blue $-□-$).

and the first adjoint mode for the limit cycle case, and the leading POD mode that is computed only from snapshots of the limit cycle, taken from $t = 400$ until $t = 800$. The traces shown in Figure 6(b) are very similar to those shown in Figure 4(d), indicating that similar structures are responsible for the transient in both the convectively unstable case and the case with the limit cycle.

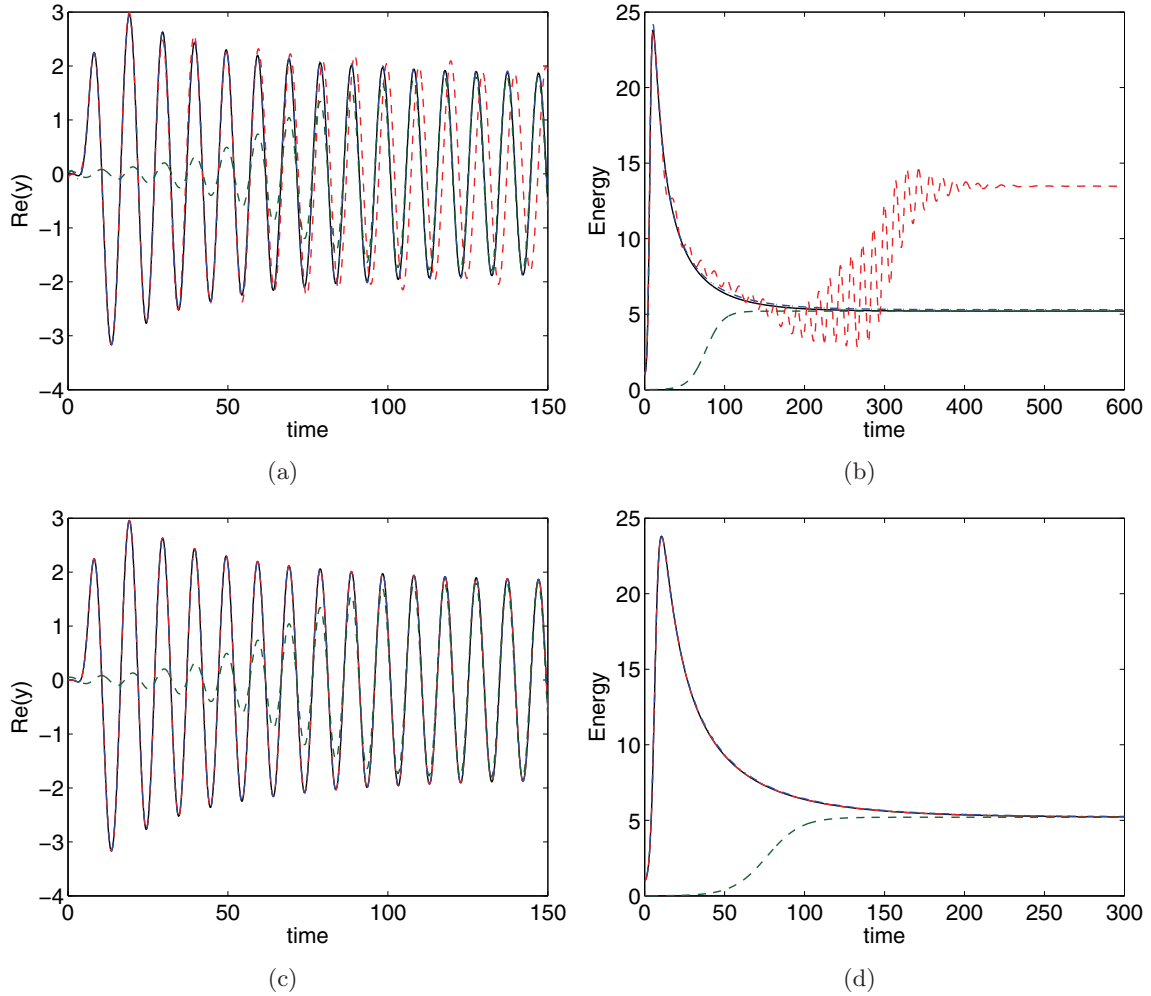


Figure 7. (a) The real part of the output y for models of rank 5 using projection onto balancing modes (blue dash-dotted line) and projection onto POD modes (red dashed line), as well as the real part of the output of the single-mode POD model (green dashed line), compared to the real part of output of the full system (solid black line), in the limit cycle case. (b) A comparison of the corresponding energy norm $\|q\|^2$ for the full solution and these models. The performance of both projection onto POD and projection onto balancing modes at rank 10, shown in (c) and (d), is excellent, as in the convectively unstable case. The performance of the single-mode POD model is shown again in (c) and (d) for comparison. Note that the time intervals shown in different plots differ in order to illustrate the performance clearly.

In Figure 7 we show the performance of reduced-order models of rank 5 and 10 (respectively, 10 and 20 degrees of freedom), similarly to Figure 3, except that here we also include the performance of the POD model consisting only of the single complex POD mode (two degrees of freedom). The results are similar to the convectively unstable case; the POD models computed for the entire duration of the impulse response capture the correct transient at low rank, but they do not capture the limit cycle correctly, while balanced truncation does capture both the transient and the limit cycle even at a low rank. The single-mode POD

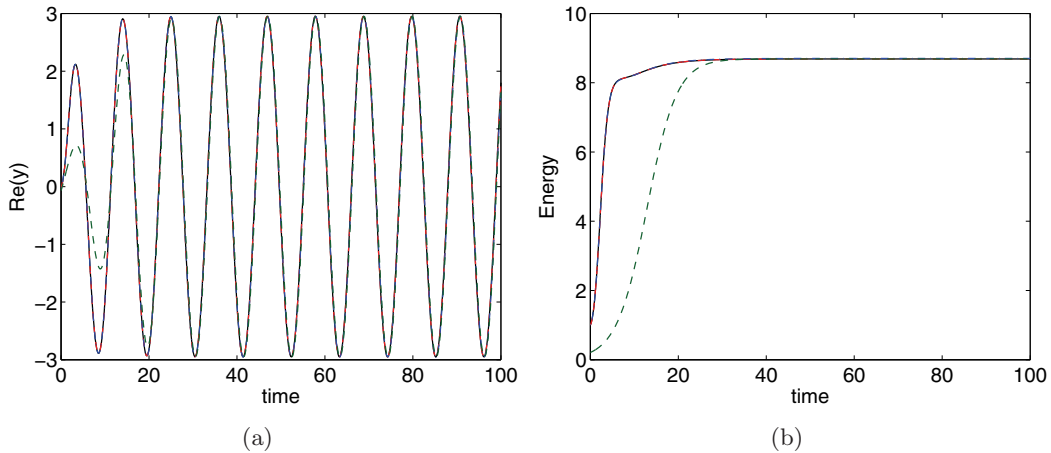


Figure 8. (a) The real part of the output y for models of rank 5 from projection onto balancing modes (blue dash-dotted line) and projection onto POD modes (red dashed line), as well as the real part of the output of the single-mode POD model (green dashed line), compared to the real part of output of the full system (solid black line), in the limit cycle case with low nonnormality. The single-mode POD model still does not capture the transient. (b) A comparison of the corresponding energy norm $\|q\|^2$ for the full solution and these models (same legend). Both the projection onto POD modes and the projection onto balancing modes capture very well the transient energy growth, while the single-mode POD model misses it.

model captures the amplitude of the limit cycle quite well, although, as may be expected, it completely misses the transient. This result is not surprising, since transient dynamics needs to be included in a POD model through taking a sufficient number of snapshots during the transient.

4.3. The effect of decreasing nonnormality. In order to further explore the difference in the performance between projection onto balancing modes and projection onto POD modes, we look into the effect of nonnormality. In the case of the CGL equation, as well as in many spatially developing flows, the nonnormality is due to the nonparallel nature of the flow, which is clearly demonstrated by the streamwise separation of the direct and adjoint eigenmodes,³ as well as of the balancing modes and their corresponding adjoint modes. In this section we consider another case, where the parameters of the CGL equation have been changed so that the unstable region is much smaller, and thus the nonnormality of the operator is not as highly exhibited as in the other cases. The parameter values are the same as in Table 1, except that $\mu_0 = 0.69$ and $\mu_2 = -0.12$.

The performance of the models is shown in Figure 8, and here a five-mode POD model performs as well as the balanced truncation model in capturing both the total energy and the single output. The simple 1D model using only a POD mode of the limit cycle again does not capture the transient, as expected, but does capture the dynamics on the limit cycle. A likely cause of this improvement in performance of projection onto POD modes is the lower nonnormality in this case. In Figure 9 we show the absolute values of the first mode from

³See, for example, [6], for a discussion of the streamwise spatial separation of direct and adjoint modes and its relation to nonnormality.

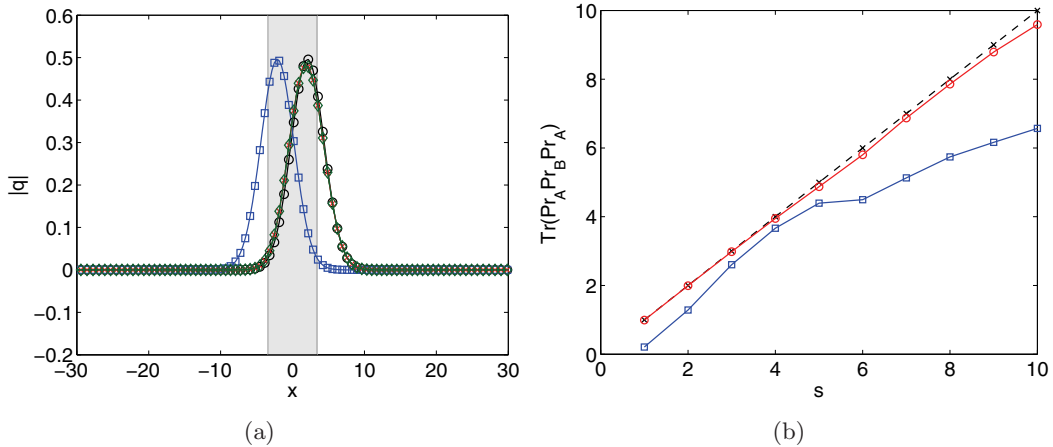


Figure 9. (a) The absolute values of the first POD mode (red $-X-$), as well as the direct (black $-O-$) and adjoint (blue $-□-$) modes from balanced truncation for the case with low nonnormality. The absolute value of the leading POD mode computed only from snapshots of the limit cycle is also shown (green $-◇-$). Compare to Figure 6(a). (b) Subspace comparison using the method of traces, comparing the subspaces of POD and balancing modes (red $-O-$) and POD and adjoint balancing modes (blue $-□-$).

each basis, as in Figure 6. The unstable region clearly has a much smaller spatial extent. When the subspaces are compared, we see that the POD and balanced truncation modes are similar, as before, but now the subspace spanned by the first few adjoint modes (at least the first five) is much closer to that spanned by the POD modes. Thus, we may expect the POD modes, which now have spatial support over more of the region than in the case discussed in the previous section, to be able to capture the dynamics of the original system.

We also note that the nonnormality of a nonparallel flow is also captured using global eigenmodes (eigenvectors of the matrix A) and the corresponding adjoint eigenmodes [1]. In this method, however, inputs and outputs are not captured in a systematic way, and the resulting reduced-order models thus typically have inferior performance. The results of our work and previous other work demonstrate that balanced truncation shows superior performance due to the capturing of both the nonnormality in the system dynamics and the input/output behavior, which is dependent on the inputs and outputs.

5. Discussion and conclusions. We have demonstrated that reduced-order models obtained from a nonlinear generalization of balanced truncation capture well the dynamics of the nonlinear CGL equation, with only a small number of the leading modes included in the models. The model reduction method was proposed in [27] and involves a linear coordinate transformation of the full nonlinear system into a frame in which a linearized system is a balanced realization, followed by truncation of the modes that are least controllable/observable. On the other hand, the performance of models computed using POD where the snapshots include the transient is comparable to the performance of the projection onto balancing modes only if a larger number of modes is included. We also see that, as may be expected, a single POD mode that describes the limit cycle with almost complete accuracy fails to capture the transient.

As discussed in section 2, the CGL equation is a model for various phenomena in fluids, such as vortex shedding in the flow past a cylinder, in which a Hopf bifurcation occurs as a parameter such as Reynolds number is increased beyond a critical value. The limit cycle for both the cylinder and the nonlinear CGL equation is due to a linear instability mechanism that is counteracted by the nonlinear damping, and this linear mechanism is captured by balanced truncation. We must note, however, that there is a major difference between the nonlinear CGL equation and the cylinder in crossflow. The dynamics of the linear and nonlinear CGL equations evolves in similar regions of the state space, since the oscillation in the limit cycle case is about the equilibrium solution $q(x) = 0$. This feature of the dynamics may be key to the success of balanced truncation of the system that has been linearized about that equilibrium. In the cylinder flow the mean of a limit cycle and the (unstable) equilibrium solution are different. It was found to be necessary to include a so-called *shift mode*, which is the difference between the equilibrium and the mean, to reproduce the correct dynamics in reduced-order models of cylinder flow [23]. A topic of further investigation is the use of our generalization of balanced truncation for cases where the equilibrium is different from the mean of the limit cycle, such as in the case of the cylinder. For the spatially developing boundary layer, however, there is no shift mode, since only a convective instability is exhibited, and so we may expect the method to be more successful for this flow.

The significance of our findings is that high-fidelity nonlinear models that outperform POD models may be obtained using a projection of the nonlinear system onto a linear projection onto balancing modes, avoiding some of the challenges in computing models using POD, such as noisy and/or unavailable transient data in experiments, or the computational challenges associated with snapshot-based methods for nonlinear balanced truncation such as that of [16]. For systems with a small to moderate number of degrees of freedom (most 1D systems), only a linear model of the system is necessary for computing balanced truncation. The recent success in the use of methods for computing balanced truncation for two-dimensional and three-dimensional flows such as BPOD and the Eigensystem Realization Algorithm (ERA) [20] has made balanced truncation available for modeling and linear control of these flows as well. It remains to be investigated whether our results may be extended to these cases, in particular for systems with interaction among modes with different frequencies, such as in [24]. The effect of nonnormality on the model performance also warrants further research, as our initial investigation shows that balanced truncation captures the nonnormality in the original system very well, while POD does well for less nonnormal systems, which may be used as a guideline in picking the control strategy for a particular problem based on insights from the basic physics.

REFERENCES

- [1] E. ÅKERVIK, J. HØPFFNER, U. EHRENSTEIN, AND D. S. HENNINGSON, *Optimal growth, model reduction and control in a separated boundary-layer flow using global eigenmodes*, J. Fluid Mech., 579 (2007), pp. 305–314.
- [2] S. AHUJA AND C. W. ROWLEY, *Feedback control of unstable steady states of flow past a flat plate using reduced-order estimators*, J. Fluid Mech., 645 (2010), pp. 447–478.
- [3] N. AUBRY, P. HOLMES, J. L. LUMLEY, AND E. STONE, *The dynamics of coherent structures in the wall region of a turbulent boundary layer*, J. Fluid Mech., 192 (1988), pp. 115–173.
- [4] S. BAGHERI, L. BRANDT, AND D. S. HENNINGSON, *Input-output analysis, model reduction and control of the flat-plate boundary layer*, J. Fluid Mech., 620 (2009), pp. 263–298.

- [5] S. BAGHERI, D. S. HENNINGSON, J. HÖPFFNER, AND P. J. SCHMID, *Input-output analysis and control design applied to a linear model of spatially developing flows*, Applied Mechanics Reviews, 62 (2009).
- [6] J. CHOMAZ, *Global instabilities in spatially developing flows: Non-normality and nonlinearity*, in Annual Review of Fluid Mechanics, Vol. 37, Annual Reviews, Palo Alto, CA, 2005, pp. 357–392.
- [7] J. M. CHOMAZ, P. HUERRE, AND L. G. REDEKOPP, *Bifurcations to local and global modes in spatially developing flows*, Phys. Rev. Lett., 60 (1988), pp. 25–28.
- [8] C. COSSU AND J. CHOMAZ, *Global measures of local convective instabilities*, Phys. Rev. Lett., 78 (1997), pp. 4387–4390.
- [9] R. EVERSON AND L. SIROVICH, *Karhunen-Loève procedure for gappy data*, J. Opt. Soc. Amer. A, 12 (1995), pp. 1657–1664.
- [10] K. FUJIMOTO AND D. TSUBAKINO, *Computation of nonlinear balanced realization and model reduction based on Taylor series expansion*, Systems Control Lett., 57 (2008), pp. 283–289.
- [11] M. HÖGBERG, T. R. BEWLEY, AND D. S. HENNINGSON, *Linear feedback control and estimation of transition in plane channel flow*, J. Fluid Mech., 481 (2003), pp. 149–175.
- [12] P. HOLMES, J. L. LUMLEY, AND G. BERKOOZ, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, UK, 1996.
- [13] P. HUERRE AND P. A. MONKEWITZ, *Local and global instabilities in spatially developing flows*, in Annual Review of Fluid Mechanics, Vol. 22, Annual Reviews, Palo Alto, CA, 1990, pp. 473–537.
- [14] M. ILAK, *Model Reduction and Feedback Control of Transitional Channel Flow*, Ph.D. thesis, Princeton University, Princeton, NJ, 2009.
- [15] M. ILAK AND C. W. ROWLEY, *Modeling of transitional channel flow using balanced proper orthogonal decomposition*, Phys. Fluids, 20 (2008).
- [16] S. LALL, J. E. MARSDEN, AND S. GLAVAŠKI, *A subspace approach to balanced truncation for model reduction of nonlinear control systems*, Internat. J. Robust Nonlinear Control, 12 (2002), pp. 519–535.
- [17] A. J. LAUB, M. T. HEATH, C. C. PAGE, AND R. C. WARD, *Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms*, IEEE Trans. Automat. Control, 32 (1987), pp. 115–122.
- [18] E. LAUGA AND T. R. BEWLEY, *Performance of a linear robust control strategy on a nonlinear model of spatially developing flows*, J. Fluid Mech., 512 (2004), pp. 343–374.
- [19] K. H. LEE, L. CORTELEZZI, J. KIM, AND J. SPEYER, *Application of reduced-order controller to turbulent flows for drag reduction*, Phys. Fluids, 13 (2001), pp. 1321–1330.
- [20] Z. MA, S. AHUJA, AND C. W. ROWLEY, *Reduced order models for control of fluids using the eigensystem realization algorithm*, Theoret. Comput. Fluid Dynam., (2010).
- [21] Z. MA, C. W. ROWLEY, AND G. TADMOR, *Snapshot-based balanced truncation for linear time-periodic systems*, IEEE Trans. Automat. Control, 55 (2010), pp. 469–473.
- [22] B. C. MOORE, *Principal component analysis in linear systems: Controllability, observability, and model reduction*, IEEE Trans. Automat. Control, 26 (1981), pp. 17–32.
- [23] B. NOACK, K. AFANASIEV, M. MORZYŃSKI, G. TADMOR, AND F. THIELE, *A hierarchy of low-dimensional models for the transient and post-transient cylinder wake*, J. Fluid Mech., 497 (2003), pp. 335–363.
- [24] B. R. NOACK, M. SCHLEGEL, B. AHLBORN, G. MUTSCHKE, M. MORZYŃSKI, P. COMTE, AND G. TADMOR, *A finite-time thermodynamics of unsteady fluid flows*, J. Non-Equilibrium Thermodynamics, 33 (2008), pp. 103–148.
- [25] K. OGATA, *Modern Control Engineering*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ, 1997.
- [26] M. PROVANSAL, C. MATHIS, AND L. BOYER, *Bénard-von Kármán instability: Transient and forced regimes*, J. Fluid Mech., 182 (1987), pp. 1–22.
- [27] C. W. ROWLEY, *Model reduction for fluids using balanced proper orthogonal decomposition*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (2005), pp. 997–1013.
- [28] H. SANDBERG AND A. RANTZER, *Balanced truncation of linear time-varying systems*, IEEE Trans. Automat. Control, 49 (2004), pp. 217–229.
- [29] J. M. A. SCHERPEN, *Balancing for nonlinear systems*, Systems Control Lett., 21 (1993), pp. 143–153.
- [30] P. J. SCHMID AND D. S. HENNINGSON, *Stability and Transition in Shear Flows*, Appl. Math. Sci. 142, Springer-Verlag, New York, 2001.
- [31] L. SIROVICH, *Turbulence and the dynamics of coherent structures, parts I–III*, Quart. Appl. Math., 45 (1987), pp. 561–590.

-
- [32] T. R. SMITH, J. MOEHLIS, AND P. HOLMES, *Low-dimensional models for turbulent plane Couette flow in a minimal flow unit*, J. Fluid Mech., 538 (2005), pp. 71–110.
 - [33] E. I. VERRIEST AND W. GRAY, *Nonlinear balanced realizations*, in Proceedings of the 43rd IEEE Conference on Decision and Control, IEEE Control Systems Society, Piscataway, NJ, 2004.
 - [34] J. A. C. WEIDEMAN AND S. C. REDDY, *A MATLAB differentiation matrix suite*, ACM Trans. Math. Software, 26 (2000), pp. 465–519.
 - [35] K. ZHOU, G. SALOMON, AND E. WU, *Balanced realization and model reduction for unstable systems*, Internat. J. Robust Nonlinear Control, 9 (1999), pp. 183–198.