A New Closure Strategy for Proper Orthogonal Decomposition Reduced-Order Models

Imran Akhtar
Department of Mechanical Engineering, NUST College of Electrical & Mechanical Engineering, National University of Sciences & Technology (NUST), Islamabad, Pakistan
e-mail: imran.akhtar@cem.e.nust.edu.pk

Zhu Wang
Jeff Borggaard
Traian Iliescu
Interdisciplinary Center for Applied Mathematics, MC 0531, Virginia Tech, Blacksburg, VA 24061

Proper orthogonal decomposition (POD) is one of the most significant reduced-order modeling (ROM) techniques in fluid mechanics. However, the application of POD based reduced-order models (POD-ROMs) is primarily limited to laminar flows due to the decay of physical accuracy. A few nonlinear closure models have been developed for improving the accuracy and stability of the POD-ROMs, which are generally computationally expensive. In this paper we propose a new closure strategy for POD-ROMs that is both accurate and effective. In the new closure model, the Frobenius norm of the Jacobian of the POD-ROM is introduced as the eddy viscosity coefficient. As a first step, the new method has been tested on a one-dimensional Burgers equation with a small dissipation coefficient $\nu = 10^{-3}$. Numerical results show that the Jacobian based closure model greatly improves the physical accuracy of the POD-ROM, while maintaining a low computational cost. [DOI: 10.1115/1.4005928]

1 Introduction

The proper orthogonal decomposition (POD) is widely used to derive reduced-order models (ROM) of large and complex systems. POD has been successfully applied to many scientific and engineering problems, such as fluid flow control [1–6], image processing [3], and pattern recognition [7]. It was introduced to the field of turbulence by Lumley [8] to identify coherent structures in the flow. POD starts with a collection of instantaneous flow field data from an accurate numerical simulation or physical experiment (often referred to as snapshots in the literature [21–23]), extracts the most energetic eigenfunctions (modes), and utilizes a Galerkin procedure to produce a low-dimensional dynamical system that models the flow field, known as a POD-ROM.

The accuracy of a POD-ROM crucially depends on how well a reduced basis represents the state variables. For turbulent flows, due to the inherent nonlinearity in the system, a large set of POD basis functions is typically required. However, it would compromise the computational efficiency which is often the prime objective in reduced-order modeling. To balance the low computational cost required by a ROM and the complexity of the targeted turbulent flows, appropriate closure modeling strategies need to be employed.

Relatively few reports on the closure of POD-ROMs for complex flows have been published [11–18]. Most of these closure models are inspired by large eddy simulation (LES) of turbulent flows [19,20] (diagrams in Fig. 1 draw a parallel between LES and POD). Since energy dissipation in turbulent flows takes place in small eddies that are neglected (correspond to higher POD modes), the conventional ROM is not able to dissipate enough energy and leads to erroneous and unstable solutions. It is thus necessary to model the interaction between retained and neglected modes. This interaction involves nonlinear terms that pose a challenge in the POD-ROM framework. In this report we put forth a novel closure model that achieves the same physical accuracy as the Smagorinsky closure POD-ROM proposed in [16–18] without increasing the computation cost. As a first step of investigations, we address this model in the Burgers equation. While Burgers equation is being commonly used as a one-dimensional approximation of the Navier-Stokes equations, it does not model turbulence. However, we use it here for simplicity and clarity of exposition.

The rest of this paper is organized as follows. In Sec. 2 we give an overview of the POD and present the conventional POD-ROM. The new nonlinear closure model is developed in Sec. 3. In Sec. 4 numerical results are presented.

2 Conventional POD-ROM

In this section we briefly describe the proper orthogonal decomposition method and present POD-ROM for one-dimensional (1D)
Given the time instances the computational domain, ... the ensemble of snapshots

\[ V := \text{span}\{\mathbf{u}(\cdot, t_1), \ldots, \mathbf{u}(\cdot, t_N)\} \]  

with \( d \). In the context of this paper, snapshots are a collection of flow field data obtained by experiments or numerical simulations at various instants of time. The POD seeks a low-dimensional basis \( \{\phi_1, \ldots, \phi_M\} \), with \( M \ll d \), which optimally approximates the input collection. Specifically, the POD basis satisfies

\[ \min_{\phi} \frac{1}{N} \sum_{i=1}^{N} \| \mathbf{u}(\cdot, t_i) - \sum_{j=1}^{M} (\mathbf{u}(\cdot, t_i), \phi_j(\cdot))_{H^1} \phi_j(\cdot) \|_H^2 \]  

subject to the conditions that \( (\phi_i, \phi_j)_{H^1} = \delta_{ij} \), \( 1 \leq i, j \leq M \). In order to solve Eq. (2) we consider the eigenvalue problem

\[ K \theta = \lambda \phi \]  

where \( K \in \mathbb{R}^{N \times N} \), with \( K_{ij} = \frac{1}{N} (\mathbf{u}(\cdot, t), \mathbf{u}(\cdot, t_i))_{H^1} \) is the snapshot correlation matrix \( \mathbf{K} \), for \( k = 1, \ldots, N \) are the eigenvectors, and \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N > 0 \) are the positive eigenvalues. It can then be shown (see \([3,21,23]\)) that the solution of Eq. (2) is given by

\[ \phi_k(\cdot) = \frac{1}{\sqrt{\lambda_k}} \sum_{j=1}^{N} (v_k, \mathbf{u}(\cdot, t_j)) \quad 1 \leq k \leq M \]  

where \( (v_k, \cdot) \) is the \( k \)-th component of the eigenvector \( v_k \). If the snapshots comprise the velocity field then \( \lambda_k \) is a measure of energy associated with each eigenfunction.

### 2.1 Proper Orthogonal Decomposition

The Burgers equation

\[ \begin{cases} u_t - \nu u_{xx} + uu_x = f & \text{in } \Omega \\ u(x, 0) = u_0(x) & \text{in } \Omega \\ u(x, t) = g(x, t) & \text{on } \partial \Omega \end{cases} \]  

where \( \nu \) is the diffusion parameter, \( f \) is the forcing term, \( \Omega \subset \mathbb{R} \) is the computational domain, \( t \in [0, T] \), with \( T \) the final time, \( u_0(\cdot) \) is the initial condition, and \( g(\cdot) \) is the boundary conditions. Without loss of generality, we assume that \( g \equiv 0 \) in the sequel. We emphasize that Burgers equation (5) does not model turbulence. However, we use it for simplicity and clarity of exposition.

The standard Galerkin finite element weak formulation is as follows: Find \( u \in H^1_0(\Omega) \), such that

\[ \begin{cases} (u_t, \phi) + (uu_x, \phi_x) + (uu_x, \phi) = (f, \phi) & \forall \phi \in H^1_0(\Omega) \\ (u(0), \chi) = (u_0, \chi) & \forall \chi \in L^2(\Omega) \end{cases} \]  

Let the time step \( \Delta t := T/N \) and the time instances \( t_k = k\Delta t, k = 0, 1, \ldots, N \). A truncated POD basis \( \mathbf{S}^M = \text{span} \{\phi_1, \phi_2, \ldots, \phi_M\} \) is obtained from the snapshots \( \{u(\cdot, t_1), u(\cdot, t_2), \ldots, u(\cdot, t_N)\} \) based on Eq. (4). A POD-ROM is developed by utilizing a Galerkin projection on the space \( \mathbf{S}^M \). Specifically, we first consider the approximation \( u \approx \sum_{m=1}^{M} q_m(t)\phi_m(x) \), then substitute this approximation in the Burgers equation (5) and employ the Galerkin method.

Let \( u_0 \in L^2(\Omega) \) and \( \mathbf{C} \in \mathbb{R} \) be the computational domain, \( \Omega \subset \mathbb{R} \). The standard Galerkin POD (POD-G) ROM of the Burgers equation (5) is to find \( u(\cdot, t) \in \mathbf{S}^M \), such that

\[ \begin{cases} (u_t, \phi) + (uu_x, \phi_x) + (uu_x, \phi) = (f, \phi) & \forall \phi \in \mathbf{S}^M \\ (u(0), \chi) = (u_0, \chi) \end{cases} \]  

The system can be written componentwise as follows: For all \( k = 1, 2, \ldots, M \),

\[ q_k(t) = b_k + \sum_{m=1}^{M} A_{km} q_m(t) + \sum_{m=1}^{M} \sum_{n=1}^{M} B_{kmn} q_n(t) q_m(t) \]  

where

\[ \begin{align*} b_k &= (f, \phi_k) \\ A_{km} &= -\phi(\phi_{m,k}, \phi_k) \\ B_{kmn} &= -\phi_{m}\phi_{n,k} \phi_k \end{align*} \]  

### 3 Closure Modeling in ROM

The dynamical system (8) is often accurate for laminar flows. From the earliest stages of POD for turbulent flows \([10]\), it was recognized that a simple Galerkin truncation will generally produce inaccurate results, even if the retained modes capture most of the system’s energy \([25,26]\). Thus, closure modeling (i.e., modeling the effect of the discarded modes onto the retained modes in the ROM \( \{\phi_1, \ldots, \phi_M\} \)) has always played a central role in POD. Reduced-order model strategies.

LES-inspired POD closure modeling generally aims to improve physical accuracy by employing the concept of an energy cascade. In \([13,16,17]\) a constant eddy viscosity was introduced in the POD-ROM. More realistic nonlinear closures for POD-ROM were developed in \([18–20]\), but introduce significant computational cost. In this section we propose a novel nonlinear closure models that is both accurate and efficient.

To illustrate the new POD-ROM closure model approach, we first present the Smagorinsky POD-ROM closure model proposed in \([18–20]\), which introduces a Smagorinsky-type eddy viscosity term in the POD-G model (8). The model is named POD-L here and reads

\[ q_k(t) = b_k + \sum_{m=1}^{M} A_{km} q_m(t) + \sum_{m=1}^{M} \sum_{n=1}^{M} B_{kmn} q_n(t) q_m(t) \]

where \( b_k, A_{km}, \) and \( B_{kmn} \) are same as those in Eq. (8), and

\[ D_{km} = (C_s |^2 \phi_{m,k} \phi_{k}) \]

In Eq. (11) \( |^2 \phi \equiv |u_\phi| \) is the Frobenius norm of \( u_\phi \) and \( C_s \) is a constant.

Although physically accurate, the closure model term \( D_{km} \) significantly increases the computational complexity of the ROM. Indeed, \( D_{km} \) involves assembling derivatives of \( u_\phi \) to compute the Frobenius norm, and needs to be calculated at every time step. This is in contrast to the other terms in Eq. (9), which are computed only once at the beginning of the numerical simulation. Since the POD basis is global, assembling \( D_{km} \) is time consuming. To lower the computational cost of the POD-L model, the authors proposed a two-level discretization algorithm in \([18]\), which computed the closure term on a much coarser grid instead of the same fine mesh as used in the direct numerical simulation (DNS). Numerical tests showed that the algorithm can reduce the computational time by an order of magnitude, without compromising the accuracy. However, the computational time of the proposed algorithm was significantly larger than POD-G.

Our objective in this study is to put forth a new POD-ROM closure model that possesses the same spirit as the POD-L ROM. It also employs the concept of energy cascade by introducing
additional, but measured, eddy viscosity into the POD-ROM, while it achieves the same accuracy as the POD-L ROM. However, the computational overhead is negligible.

The new POD-ROM closure model, POD-J ROM reads

\[ \dot{q}_k(t) = b_k + \sum_{m=1}^{M} A_{km} q_m(t) + \sum_{m=1}^{M} \sum_{n=1}^{M} B_{kmn} q_n(t) q_m(t) + \sum_{m=1}^{M} D_{km} q_m(t) \]  

(11)

where \( b_k, A_{km}, \) and \( B_{kmn} \) are same as those in Eq. (8), and the new closure term \( D_{km} \) is defined as

\[ D_{km} = (C_i, S_j, \phi_m, \phi_k) \]  

(12)

where \( C_i \) is a constant, \( |S_j| = \left| \frac{\partial q_l}{\partial q_j} \right| \), and \( \frac{\partial q_l}{\partial q_j} \) is the Jacobian of the POD-G ROM (8), which satisfies

\[ \frac{\partial \dot{q}_l}{\partial q_j} = A_{lj} + \sum_{m=1}^{M} B_{lkm} q_m + \sum_{n=1}^{M} B_{lnk} q_n \]  

(13)

POD-J ROM also requires repeated computations of the nonlinear closure term. However, since \( |S_j| = \left| \frac{\partial q_l}{\partial q_j} \right| \) can be evaluated at a negligible computational cost, the new closure model is computationally efficient. Next, we will test the numerical behavior of the POD-J ROM.

4 Numerical Results

The POD-J ROM aims at an efficient and accurate simulation of turbulence. However, as a first step, we test this novel closure model in the context of the Burgers equation with a small diffusion coefficient \( \nu = 10^{-3} \) (corresponding to a large Reynolds number in realistic complex flows). In this case, the conventional POD-G model yields inaccurate results.

We use a computational setting that is similar to that used by Kunisch and Volkwein in [23]. The initial condition on domain \( \Omega = [0, 1] \) is defined as \( u_0(x) = 1 \) if \( x \in \{0, 1/2\} \) and \( u_0(x) = 0 \) otherwise. The forcing term is \( f = 0 \) and the time interval is \( [0, T] = [0, 1] \). The boundary conditions are homogeneous zero Dirichlet.

We utilize the standard Galerkin finite-element approach with piecewise linear polynomials to discretize the problem in space, the implicit Euler method for the time integration, and Newton’s
method to solve the nonlinear system. We implement the simulation on a fine mesh with $\Delta x = 1/8192$ and a time step $\Delta t = 10^{-3}$. We regard the results as the benchmark solution, and denote it by “DNS.” The CPU time for DNS is 4304 s. The time evolution of the state variable $u$ is plotted in Fig. 2.

To obtain POD modes, we collect $S = 1001$ snapshots from the DNS and employ the procedure mentioned in Sec. 2. The cumulative energy $E_M = \sum_{i=1}^{M} \sum_{s=1}^{S} k_i^s$ is shown in Fig. 3, which represents the percent energy captured by the first $M$ POD modes in the system. It is seen that, due to the optimality of the POD modes, the first few modes contain the most energy.

We compare the following POD-ROMs: (i) the POD-G model (8); (ii) the POD-L model (9); and (iii) the POD-J model [15]. As an efficiency criterion, we use the CPU time required to perform the simulation. To measure the accuracy we define a relative error as

$$\text{error} = \frac{1}{N} \sum_{i=1}^{N} \left\| u^{\text{POD-ROM}}(x, t_k) - u^{\text{DNS}}(x, t_k) \right\|_0^2$$

(14)

Fig. 6 POD-J ROM simulation of the Burgers equation with $\nu = 10^{-3}$. Note that the results have great improvement comparing with those of POD-G ROM.

Fig. 7 A comparison among DNS (blue), POD-G (green dash), POD-L (red triangle), and POD-J (cyan square) at $t = 1$. Note that POD-J and POD-L are much closer to DNS than POD-G ROM. However, POD-J is computationally more efficient than POD-L.

We compare the following POD-ROMs: (i) the POD-G model (8); (ii) the POD-L model (9); and (iii) the POD-J model [15]. As an efficiency criterion, we use the CPU time required to perform the simulation. To measure the accuracy we define a relative error as

$$\text{error} = \frac{1}{N} \sum_{i=1}^{N} \left\| u^{\text{POD-ROM}}(x, t_k) - u^{\text{DNS}}(x, t_k) \right\|_0^2$$

(14)

Fig. 8. Time evolution of POD basis coefficient $q_i, i = 1, 2, \ldots, 10$ for DNS projection (blue), POD-G (green dash), POD-L (red triangle), and POD-J (cyan square). Note that the POD-L and POD-J models have close behaviors to DNS, and perform much better than the POD-G model.
where \( \mathbf{u}^{\text{POD-ROM}} \) is a POD-ROM solution, \( \mathbf{u}^{\text{DNS}} \) is the benchmark solution, and \( N = 1000 \) in this particular case.

We first use \( M = 10 \) POD modes to generate all three POD-ROMs. Although the first 10 POD modes capture 99.44% of the system’s energy, the standard POD-G ROM generates inaccurate results as shown in Fig. 4, which implies the necessity of closure models. We then employ POD-L ROM and POD-J ROM, respectively. Note that the artificial viscosity coefficients in POD-L and POD-J ROMs, \( C_s \) and \( \tilde{C}_s \), respectively, should be determined in advance. But, to our knowledge, there is no exact formula or relation that describes how to choose them. Therefore, in our implementations, we use \( \nu \) as the initial guess and seek an appropriate value of \( C_s \) by running the corresponding ROM on a short interval \([0, 0.1]\) and minimizing the error over this time interval.

We obtained \( C_s = 3 \times 10^{-4} \) for the POD-L ROM and \( \tilde{C}_s = 1.05 \times 10^{-4} \) for the POD-J ROM. The time evolution of POD-L ROM solutions is shown in Fig. 5 and that of POD-J ROM solutions is shown in Fig. 6. For a clarity of comparison, we also include POD-ROMs results with the DNS at the final time \( t = 1 \) in Fig. 7. It is clear that the closure models POD-J and POD-L are quantitatively close to the DNS and much more stable than the standard POD-G model.

In Fig. 8 we plot the time evolution of the POD basis coefficients \( q_i \) (\( i = 1, 2, \ldots, 10 \)) for DNS projections, POD-G, POD-L, and POD-J models. Among the ROMs, POD-L and POD-J models perform much better than the POD-G model, especially for the higher modes.

To quantify the computational efficiency of the POD-ROMs, we compare their relative errors defined in (14) in Table 1. The relative error of the POD-L and POD-J models is significantly lower than the that of the POD-G model.

The key advantage of the POD-J model over the POD-L model lies in the low computational cost. In Table 2 we list the CPU time of the three POD-ROMs. Despite its inappropriate behavior, the POD-G model has the minimum CPU time. This is natural since the POD-G model does not employ any turbulence modeling, wherein all matrices can be precomputed. The CPU time of the POD-J model, however, is only less than twice the CPU time of the POD-G model but 36 times less than that of the POD-L model. The reason is that the closure term in the POD-J model only requires the computation of the Jacobian of the reduced-order model at each time step, which is dramatically more efficient than the computation of the Frobenius norm of \( \mathbf{u}_i \) in the POD-L model.

Thus, for this test problem, the POD-J model outperforms the POD-G and the POD-L models in both relative error and CPU time.

We then increase the number of POD modes to \( M = 30 \) (99.91% of the snapshot energy has been captured in this basis), and observe the same phenomena as in the \( M = 10 \) case. Here, \( C_s = 1 \times 10^{-3} \) and \( \tilde{C}_s = 1.15 \times 10^{-5} \) are used in the POD-L and POD-J ROMs, respectively. The relative error of the POD-G model decreases, yet the POD-L and POD-J models still perform better than the POD-G model. We also notice that the relative error of the POD-J model is less than that of the POD-L model. The CPU time of the POD-L model is 23 times larger for \( M = 30 \).

5 Conclusions

This report has put forth a new closure model for POD-ROMs. This new strategy is similar in spirit to the LES closure models previously investigated in a POD context. But using the norm of the Jacobian as the eddy viscosity coefficient in the closure term produces a more computationally efficient model.

As a first step, the new POD-J model has been developed and tested numerically in the 1D Burgers equation, which is considered as a simplified model of the Navier-Stokes equations. By comparing the results of the new POD-J with those of the POD-L ROM and the conventional POD-G ROM, we conclude that the novel POD-J model achieves the same accuracy as POD-L ROM and outperforms the POD-G ROM. The POD-J achieves this accuracy at a low computational cost. Indeed, the CPU time for the POD-J ROM near the computational complexity of POD-G.

Although in its infancy, this new POD-ROM closure model has shown promising results. We plan to apply this closure modeling strategy to more complex dynamical systems, such as the Navier-Stokes equations.

Acknowledgment

This research was partially supported by the National Science Foundation under Contract DMS-1016450 and by the Department of Energy Contract DE-EE0004261 under Subcontract No. 4345-VT-DOE-4261 from Penn State University.

References


Table 1 Relative errors of POD-ROMs with \( \nu = 10^{-3} \)

<table>
<thead>
<tr>
<th>( M )</th>
<th>POD-G</th>
<th>POD-L</th>
<th>POD-J</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>( 7.54 \times 10^{-2} )</td>
<td>( 1.73 \times 10^{-2} )</td>
<td>( 1.14 \times 10^{-2} )</td>
</tr>
<tr>
<td>30</td>
<td>( 2.38 \times 10^{-3} )</td>
<td>( 1.91 \times 10^{-3} )</td>
<td>( 1.36 \times 10^{-3} )</td>
</tr>
</tbody>
</table>

Table 2 CPU time of POD-ROMs with \( \nu = 10^{-3} \)

<table>
<thead>
<tr>
<th>( M )</th>
<th>POD-G</th>
<th>POD-L</th>
<th>POD-J</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>9 s</td>
<td>587 s</td>
<td>16 s</td>
</tr>
<tr>
<td>30</td>
<td>26 s</td>
<td>612 s</td>
<td>42 s</td>
</tr>
</tbody>
</table>