Projection-based reduced order modeling and data-driven artificial viscosity closures for incompressible fluid flows

Aviral Prakash [*], Yongjie Jessica Zhang

Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

A R T I C L E   I N F O

A B S T R A C T

Projection-based reduced order models rely on offline–online model decomposition, where the data-based energetic spatial basis is used in the expensive offline stage to obtain equations of reduced states that evolve in time during the inexpensive online stage. The online stage requires a solution method for the dynamic evolution of the coupled system of pressure and velocity states for incompressible fluid flows. The first contribution of this article is to demonstrate the applicability of the incremental pressure correction scheme for the dynamic evolution of pressure and velocity states. The evolution of a large number of these reduced states in the online stage can be expensive. In contrast, the accuracy significantly decreases if only a few reduced states are considered while not accounting for the interactions between unresolved and resolved states. The second contribution of this article is to compare three closure model forms based on global, modal and tensor artificial viscosity approximation to account for these interactions. The unknown model parameters are determined using two calibration techniques: least squares minimization of error in energy approximation and closure term approximation. This article demonstrates that an appropriate selection of solution methods and data-driven artificial viscosity closure models is essential for consistently accurate dynamics forecasting of incompressible fluid flows.

1. Introduction

Over the past decade, improvements in computational hardware and advancements in physical simulations have pushed the envelope of complexity of scientific applications that can be modeled with adequate accuracy. High-fidelity simulations of many challenging applications demand extensive computational resources and long computational time. The expensive nature of these simulations prohibits their use in the context of multi-query applications, real-time simulation and dynamics forecasting. In such situations, reduced order models (ROMs) are an attractive alternative as they can potentially simulate engineering systems at a lower computational overhead without a significant loss in accuracy. ROMs often rely on offline–online task decomposition, where the computationally expensive component is ideally performed in the offline stage and typically carried out only once. In contrast, the inexpensive online stage is performed when the model is deployed in the actual application.

Projection-based ROMs [1,2] are the most popular among model reduction approaches. These models rely on high-fidelity data to obtain spatial basis, using modal decomposition methods such as proper orthogonal decomposition (POD) [3,4] or reduced basis [5], and optimally representing the data in terms of reduced states. Using this basis representation in the relevant physical equation and projecting the solution to a lower-dimensional subspace, a set of equations for the evolution of reduced states is obtained. The reduced states are classified as resolved states, which are the most energetic states whose dynamics are predicted, and unresolved

states, which are less energetic states whose dynamics are neglected. These equations can be solved in the online stage of ROMs to get the dynamic evolution of the solution. The cost of the online stage can scale highly with the number of resolved states; for example, in the case of Navier–Stokes equations, the cost can scale as $O(r^3)$ where r is the number of resolved states. A lower number of resolved states must be considered for such applications. This selection comes at the cost of more unresolved states, which negatively influences the dynamics of the resolved states. Several approaches have been considered in ROM literature to account for the interactions between resolved and unresolved states either directly [2,6,7] or indirectly [8,9]. Closure models are direct approaches that add additional terms to the dynamic evolution of the resolved states to account for their interaction with unresolved states. A comprehensive summary of several approaches under this category is given in [10]. Common indirect approaches include stabilization methods [8,9,11–13] that typically use Petrov–Galerkin projection, which implicitly adds terms to the dynamic evolution of reduced states to account for the interaction of unresolved states. Some of these Petrov–Galerkin projection methods can also be considered a variational multiscale method [8,14] and categorized as closure models. These Petrov–Galerkin projection ROMs often represent the dynamics of resolved states based on the optimality conditions [11,12].

In this article, we focus on the closure models to account for the influence of unresolved states on the resolved states. These models can be broadly classified into physics-driven and data-driven models. Physics-driven models [2,15,16] typically use physical insight to add terms to equations of the resolved states to account for their interaction with unresolved states. For example, artificial/eddy viscosity model forms [2,7,16,17] are based on the assumption that unresolved states are primarily responsible for dissipating energy from the system. Augmented versions of these models based on modal artificial viscosity enable energy dissipation characteristics to vary for individual resolved states [6,15,16] and provide more accurate dynamics of resolved states [18]. Alternate models based on the theory developed in the context of large eddy simulation (LES) [19], variational multiscale method [14,20] and regularization [21] have also been developed. The model parameters in these models are typically obtained using several calibration techniques like manual tuning, statistical averages [16,18], data-assimilation [22–24] and least squares minimization [16]. With the growing interest in modern machine learning algorithms, data-driven models for closure modeling in ROMs [10,25] and LES [26–29] are becoming popular. These data-driven models rely on raw data to determine the approximate operators or correction to the operators in the resulting dynamic system. Often, these models assume a model form, such as models with a linear term [30], models with both linear and quadratic terms [25,31] and neural network model forms [32]. As the development and deployment of ROMs are typically in a data-rich environment, data-driven closure models can directly leverage the data used for constructing the energetic spatial basis. These models are generally ill-posed and require regularization to work well [25,33]. The selection of an appropriate regularization scheme and parameters depends on the application scenario and model selection, thereby adding an expensive model-tuning process to obtain efficient ROMs.

In addition to the closure problem, ROMs of incompressible fluid flows require unique solution methods for coupled continuity and momentum equations in the incompressible Navier–Stokes equations. Most ROMs for incompressible flows neglect the role of pressure gradient in the momentum equations by assuming the discretely divergence free nature of the velocity data. However, in most situations, this assumption is not satisfied. For example, data from commonly used numerical methods, like finite element methods with Lagrange elements, is also not discretely divergence free. Furthermore, obtaining discretely divergence free velocities may not be possible with common experimental measurement techniques. ROMs of shear layers are sensitive to the approximation of pressure gradient; thereby, neglecting this term can reduce ROM accuracy [34]. Accurate dynamics prediction of pressure states is also essential for several applications such as fluid–structure interaction [35]. Therefore, ROMs that avoid such assumptions and accurately solve coupled equations of velocity and pressure reduced states are essential. In these coupled equations, the role of the pressure states is to enforce the incompressibility condition. Improper solution methods can result in spurious pressure modes or pressure instabilities and several studies have focused on mitigating them [35–38].

Based on the discussion above, an ideal ROM for incompressible flows needs to have two key features: (1) an accurate and stable solution method for the dynamic evolution of velocity and pressure states and (2) an adequate closure model to account for interactions between unresolved and resolved states while reducing the costly human-in-loop involvement of parameter tuning, such as parameter tuning while adding regularization or manually tuning parameters for physics-driven models. Inspired by these challenges, the main contributions of this article are two-fold:

- The incremental pressure correction scheme to solve coupled equations of velocity and pressure states while providing a second-order accurate temporal evolution.
- New artificial viscosity closure models with unknown parameters determined using two data-driven calibration techniques: least squares minimization of error in energy approximation and error in closure term approximation.

The first contribution is to demonstrate the applicability of incremental pressure correction schemes through numerical experiments. This solution method has been commonly used in FOM [39]; however, to the extent of the authors knowledge, this solution method has not been used previously in the context of ROMs. This solution method gives us velocity and pressure states without requiring the pressure and velocity states to satisfy *inf–sup* condition, thereby avoiding the need for a more expensive supremizer enrichment method [36]. The second contribution is to demonstrate the applicability of several artificial viscosity closure models for resolving the closure problem in ROMs. In addition to the two physics-driven model forms discussed above, which are global [2] and modal [15,16] artificial viscosity model forms, we also propose a tensor artificial viscosity model form. This model form is inspired by the concept of tensor eddy viscosity subgrid models for LES [40,41]. The unknown parameters for these three model forms are learned using the least squares minimization technique, typically used for data-driven closure models. In this article, we compare six closure models obtained by combining each of these three physics-driven model forms with two data-driven calibration techniques for determining unknown model parameters. These closure models lie between physics-driven and data-driven models

and we refer to them as data-driven artificial viscosity closure models. Unlike the data-driven closure models that approximate the quadratic operator [25,31], these closure models did not require additional regularization for the numerical experiments in this article as these physics-driven model forms have fewer unknown parameters to calibrate. Together with the proposed solution method for incompressible flows, we validate and compare the proposed six closure models for forecasting the flow dynamics over a 2-D cylinder at three Reynolds numbers.

The outline of this article is given below. Section 2 discusses the incompressible Navier–Stokes equations. Section 3 discusses POD for dimensionality reduction. Section 4 mentions the construction of projection-based ROMs with a POD basis while highlighting the closure problem and several models proposed in this work to overcome it. This section also discusses the incremental pressure correction scheme for solving coupled equations for resolved velocity and pressure states. Section 5 validates the solution method for ROMs and several closure models for an incompressible fluid flow case: flow over a 2-D cylinder. Section 6 concludes this article by summarizing its main contributions and highlighting possible directions for future research.

2. Incompressible Navier–Stokes equations

Incompressible Navier–Stokes equations are commonly solved to obtain discrete approximations, denoted by superscript h, of velocity $\boldsymbol{u}^h : [0, T] \times \Omega \to \mathbb{R}^d$ and pressure $p^h : [0, T] \times \Omega \to \mathbb{R}$ in an incompressible flow in a domain $\Omega \subset \mathbb{R}^d$ of dimensionality d at time $T \in \mathbb{R}^+$. The approximate velocity and pressure lie in respective function spaces, $\boldsymbol{u}^h \in \mathcal{V}^h$ and $p^h \in \mathcal{Q}^h$, which are appropriately chosen to represent the flow physics. These equations consist of the continuity equation

$$\nabla \cdot \boldsymbol{u}^h = 0, \tag{1}$$

which ensures the conservation of mass, and the momentum equation

$$\frac{\partial \boldsymbol{u}^h}{\partial t} + \boldsymbol{u}^h \cdot \nabla \boldsymbol{u}^h = -\nabla p^h + \nu \nabla^2 \boldsymbol{u}^h, \tag{2}$$

which ensures the conservation of flow momentum. From here onwards, we will remove superscript h and use $\boldsymbol{u}$ and p to refer to discrete approximations of velocity and pressure, respectively. The solution to incompressible Navier–Stokes equations exhibits a saddle-point structure, which is accompanied by the role of the pressure to enforce the incompressibility constraint. For this scenario, the equation that governs the dynamic evolution of pressure is unavailable, unlike compressible Navier–Stokes equations, where such an equation can be obtained. The saddle point nature of these equations often restricts selecting appropriate function spaces for velocity and pressure. For such saddle point problems, the Ladyzhenskaya–Babuška–Brezzi (LBB) or *inf–sup* condition [42,43] is a sufficient condition to ensure the uniqueness of the solution.

3. Proper orthogonal decomposition

An effective ROM relies on constructing the basis and determining reduced states that can effectively represent the flow dynamics. POD is one of the most common methods for dimensionality reduction and obtaining an effective spatial basis for the dynamic system [3,4]. Velocity and pressure fields are decomposed using POD as

$$\boldsymbol{u}(\boldsymbol{x}, t) \approx \boldsymbol{u}_m(\boldsymbol{x}) + \sum_{i=1}^{n} a_i(t) \boldsymbol{\phi}_i(\boldsymbol{x}) \qquad \text{and} \qquad p(\boldsymbol{x}, t) \approx p_m(\boldsymbol{x}) + \sum_{i=1}^{n} b_i(t) \psi_i(\boldsymbol{x}), \tag{3}$$

where n is the number of modes/states, $a_i : [0, T] \to \mathbb{R}$ and $b_i : [0, T] \to \mathbb{R}$ are the temporally evolving reduced states of velocity and pressure respectively, $\boldsymbol{\phi}_i : \Omega \to \mathbb{R}^d$ and $\psi_i : \Omega \to \mathbb{R}$ are spatial modes of velocity and pressure respectively and $\boldsymbol{u}_m$ and p_m are the reference velocity and pressure respectively which are taken to be respective mean values in this article. When finite element methods are used to solve PDEs, the spatial modes are unknown and determined by assembling a large system of equations and solving them. Meanwhile, ROMs use high-fidelity simulation or experimental data as spatial modes to reduce the dimensionality of the problem and study the temporal evolution of these reduced states. POD basis optimally represents the energy of incompressible flows. Through numerical experiments, we observed that the coupled approach, where velocity and pressure POD basis are computed together [8], did not result in a monotonic decrease in energy error with increased modes. Therefore, this article follows a decoupled approach where velocity and pressure POD basis are computed separately [35].

For generating velocity and pressure POD basis, the velocity data $\boldsymbol{X}_u(\boldsymbol{x})$ and pressure data $\boldsymbol{X}_p(\boldsymbol{x})$ at each spatial coordinate location $\boldsymbol{x} \in \Omega$ are represented as

$$\boldsymbol{X}_u(\boldsymbol{x}) := [\boldsymbol{u}^1(\boldsymbol{x}) \ \boldsymbol{u}^2(\boldsymbol{x}) \ \cdots \ \boldsymbol{u}^n(\boldsymbol{x})] \qquad \text{and} \qquad \boldsymbol{X}_p(\boldsymbol{x}) := [p^1(\boldsymbol{x}) \ p^2(\boldsymbol{x}) \ \cdots \ p^n(\boldsymbol{x})], \tag{4}$$

where $\boldsymbol{u}^i(\boldsymbol{x}) = \boldsymbol{u}(\boldsymbol{x}, t_i) - \boldsymbol{u}_m$ and $p^i(\boldsymbol{x}) = p(\boldsymbol{x}, t_i) - p_m$ for $t_i \in \{t_1, t_2, \ldots, t_n\}$. The selection of the POD velocity modes is based on minimizing the approximation error

$$\boldsymbol{\Phi} = \underset{\hat{\boldsymbol{\Phi}} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_n]}{\arg \min} \sum_{i=1}^{n} \left\| \boldsymbol{u}^i(\boldsymbol{x}) - \sum_{j=1}^{n} a_j(t_i) \boldsymbol{\phi}_j(\boldsymbol{x}) \right\|_{L^2}^2, \tag{5}$$

such that

$$(\boldsymbol{\phi}_i, \boldsymbol{\phi}_j) = \delta_{ij}, \text{ for } i, j = 1, 2, \ldots, n, \tag{6}$$

where $||\boldsymbol{v}(\boldsymbol{x})||_{L^2}^2 = \int_\Omega \boldsymbol{v}(\boldsymbol{x}) \cdot \boldsymbol{v}(\boldsymbol{x}) d\boldsymbol{x}$ is the norm introduced by the inner product $(\cdot, \cdot)$ and δ_{ij} is the Kronecker delta. Other weighed norms and inner products have been considered in the past [44,45] for better stability properties. However, these are not considered in this article. The velocity POD modes are obtained using the relation

$$\boldsymbol{\Phi} = \boldsymbol{X}_u \hat{\boldsymbol{\Phi}} \sqrt{\boldsymbol{\Lambda}_u^{-1}}. \tag{7}$$

The components of $\hat{\boldsymbol{\Phi}}(= [\hat{\boldsymbol{\phi}}_1 \ \hat{\boldsymbol{\phi}}_2 \ \cdots \ \hat{\boldsymbol{\phi}}_n])$ and components of $\boldsymbol{\Lambda}_u$, which is a diagonal matrix ($\boldsymbol{\Lambda}_u = \mathrm{diag}(\lambda_1^u, \lambda_2^u, \ldots, \lambda_n^u)$), are obtained by solving the eigenproblem

$$\boldsymbol{R}^u \hat{\boldsymbol{\phi}}_i = \lambda_i^u \hat{\boldsymbol{\phi}}_i, \tag{8}$$

where the ijth component of the velocity correlation matrix $\boldsymbol{R}^u$ is given as

$$R_{ij}^u = \int_\Omega \boldsymbol{u}^i(\boldsymbol{x}) \cdot \boldsymbol{u}^j(\boldsymbol{x}) d\boldsymbol{x}. \tag{9}$$

Similarly, the selection of POD pressure modes is based on minimizing the approximation error

$$\boldsymbol{\Psi} = \underset{\hat{\boldsymbol{\Psi}} = [\psi_1, \psi_2, \ldots, \psi_n]}{\arg\min} \sum_{i=1}^n \left|\left| p^i(\boldsymbol{x}) - \sum_{j=1}^n b_j(t_i) \psi_j(\boldsymbol{x}) \right|\right|_{L^2}^2, \tag{10}$$

such that

$$(\psi_i, \psi_j) = \delta_{ij}, \ \text{for} \quad i, j = 1, 2, \ldots, n . \tag{11}$$

The pressure POD modes are obtained using the relation

$$\boldsymbol{\Psi} = \boldsymbol{X}_p \hat{\boldsymbol{\Psi}} \sqrt{\boldsymbol{\Lambda}_p^{-1}}, \tag{12}$$

where components of $\hat{\boldsymbol{\Psi}}(= [\hat{\psi}_1 \ \hat{\psi}_2 \ \cdots \ \hat{\psi}_n])$ and components of $\boldsymbol{\Lambda}_p$, which is a diagonal matrix ($\boldsymbol{\Lambda}_p = \mathrm{diag}(\lambda_1^p, \lambda_2^p, \ldots, \lambda_n^p)$), are obtained by solving the eigenproblem

$$\boldsymbol{R}^p \hat{\psi}_i = \lambda_i^p \hat{\psi}_i. \tag{13}$$

The ijth component of the pressure correlation matrix $\boldsymbol{R}^p$ is determined using

$$R_{ij}^p = \int_\Omega p^i(\boldsymbol{x}) \cdot p^j(\boldsymbol{x}) d\boldsymbol{x}. \tag{14}$$

The orthonormality of velocity and pressure modes ensures that the reduced states at time t_j are obtained using

$$a_i(t_j) = (\boldsymbol{\phi}_i, \boldsymbol{u}(t_j)) \quad \text{and} \quad b_i(t_j) = (\psi_i, p(t_j)). \tag{15}$$

The temporal evolution of reduced states governs the dynamics of the solution. Depending on the availability of computational resources, it may be computationally unaffordable to determine the dynamics of all n reduced states. Therefore, these reduced states are further decomposed as resolved (denoted by subscript c) and unresolved states (denoted by subscript f). Keeping this in mind, POD for velocity can be rewritten as

$$\boldsymbol{u}(\boldsymbol{x}, t) \approx \boldsymbol{u}_c(\boldsymbol{x}, t) + \boldsymbol{u}_f(\boldsymbol{x}, t), \tag{16}$$

where

$$\boldsymbol{u}_c(\boldsymbol{x}, t) = \boldsymbol{u}_m(\boldsymbol{x}) + \sum_{i=1}^r a_i(t) \boldsymbol{\phi}_i(\boldsymbol{x}) \quad \text{and} \quad \boldsymbol{u}_f(\boldsymbol{x}, t) = \sum_{i=r+1}^n a_i(t) \boldsymbol{\phi}_i(\boldsymbol{x}). \tag{17}$$

The reduced states a_i for $i = 1, 2, \ldots, r$ are the resolved velocity states and a_i for $i = r+1, r+2, \ldots, n$ are the unresolved velocity states. Similarly, POD for pressure is rewritten as

$$p(\boldsymbol{x}, t) \approx p_c(\boldsymbol{x}, t) + p_f(\boldsymbol{x}, t), \tag{18}$$

where

$$p_c(\boldsymbol{x}, t) = p_m(\boldsymbol{x}) + \sum_{i=1}^r b_i(t) \psi_i(\boldsymbol{x}) \quad \text{and} \quad p_f(\boldsymbol{x}, t) = \sum_{i=r+1}^n b_i(t) \psi_i(\boldsymbol{x}). \tag{19}$$

The reduced states b_i for $i = 1, 2, \ldots, r$ are the resolved pressure states and b_i for $i = r+1, r+2, \ldots, n$ are the unresolved pressure states. Even though different number of velocity and pressure states can be used, we use the same number of states in this article.

4. Reduced order modeling

As the introduction suggests, ROMs are critical for multi-query applications, such as optimization and uncertainty quantification, real-time control, and dynamics forecasting. ROMs have two main components: representation of the high-dimensional solution in terms of reduced states and dynamical evolution of these reduced states. In Section 3, we explored using POD to obtain the set

of energy-dominant spatial bases and corresponding reduced states. This section focuses on the evolution of these reduced states for dynamics forecasting. Recently, there has been a growing trend towards fully data-based ROMs involving machine learning to represent the reduced states and evolve them in time [10]. Unfortunately, these fully data-based ROMs often have limited applicability for dynamics forecasting problems, where the goal is to predict future states that are not part of the high-fidelity training data used to learn these models. This article restricts our discussion to traditional ROMs involving physical equations to evolve the reduced states. These latter ROMs are better suited for forecasting problems where the model is evaluated for future time instances that do not inform the generation of the energy-dominant basis.

4.1. Projection-based reduced order modeling and closure problem

Projection-based ROMs are a model reduction approach where high-fidelity data is used to generate the spatial modes and the temporal evolution of reduced states is governed by physical equations. For incompressible flows, the reduced states can be obtained by substituting POD of velocity (Eq. (16)) and pressure (Eq. (18)) in Eq. (1) and Eq. (2) to obtain

$$\nabla \cdot (\boldsymbol{u}_c + \boldsymbol{u}_f) = 0 \tag{20}$$

and

$$\frac{\partial (\boldsymbol{u}_c + \boldsymbol{u}_f)}{\partial t} + (\boldsymbol{u}_c + \boldsymbol{u}_f) \cdot \nabla (\boldsymbol{u}_c + \boldsymbol{u}_f) = -\nabla(p_c + p_f) + \nu \nabla^2 (\boldsymbol{u}_c + \boldsymbol{u}_f). \tag{21}$$

Using Galerkin projection [2], the continuity equation is projected onto a linear subspace spanned by ψ_k for $k = 1, 2, \ldots, r$. The resulting continuity equation is

$$\left(\psi_k, \nabla \cdot \boldsymbol{u}_m\right) + \sum_{i=1}^{r} a_i \left(\psi_k, \nabla \cdot \boldsymbol{\phi}_i\right) + D_k^c(\boldsymbol{a}) = 0, \tag{22}$$

where $D_k^c(\boldsymbol{a}) = \sum_{i=r+1}^{n} a_i \left(\psi_k, \nabla \cdot \boldsymbol{\phi}_i\right)$ is the closure term for continuity equation. On the other hand, the momentum equation is projected onto a linear subspace spanned by $\boldsymbol{\phi}_k$ for $k = 1, 2, \ldots, r$. The resulting momentum equation is

$$\dot{a}_k + \sum_{i=1}^{r} L_{ki} a_i + \sum_{i=1}^{r} \sum_{j=1}^{r} B_{kij} a_i a_j + \sum_{i=1}^{r} L_{ki}^P b_i + C_k + D_k^m(\boldsymbol{a}, \boldsymbol{b}) = 0, \tag{23}$$

where

$$L_{ki} = -\nu \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i\right) + \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{u}_m\right) + \left(\boldsymbol{\phi}_k, \boldsymbol{u}_m \cdot \nabla \boldsymbol{\phi}_i\right), \tag{24}$$

$$B_{kij} = \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{\phi}_j\right), \tag{25}$$

$$L_{ki}^P = \left(\boldsymbol{\phi}_k, \nabla \psi_i\right), \tag{26}$$

$$C_k = \left(\boldsymbol{\phi}_k, \boldsymbol{u}_m \cdot \nabla \boldsymbol{u}_m\right) + \left(\boldsymbol{\phi}_k, \nabla p_m\right) - \nu \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{u}_m\right) \tag{27}$$

and

$$\begin{aligned} D_k^m(\boldsymbol{a}, \boldsymbol{b}) =\ & \sum_{i=1}^{r} \sum_{j=r+1}^{n} a_i a_j \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{\phi}_j\right) + \sum_{i=r+1}^{n} \sum_{j=1}^{r} a_i a_j \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{\phi}_j\right) + \sum_{i=r+1}^{n} \sum_{j=r+1}^{n} a_i a_j \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{\phi}_j\right) \\ & - \nu \sum_{i=r+1}^{n} a_i \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i\right) + \sum_{i=r+1}^{n} a_i \left(\boldsymbol{\phi}_k, \boldsymbol{u}_m \cdot \nabla \boldsymbol{\phi}_i\right) + \sum_{i=r+1}^{n} a_i \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{u}_m\right) + \sum_{i=r+1}^{n} b_i \left(\boldsymbol{\phi}_k, \nabla \psi_i\right) \end{aligned} \tag{28}$$

is the closure term for the momentum equation. This approach is often referred to as the variational multiscale (VMS) method for ROMs and D_k^m can be thought of as the closure term that accounts for coarse and fine-scale interactions defined in VMS methods [14,17]. This article neglects the effect of unresolved pressure states in the closure term. With this approximation, the resulting closure modeling becomes

$$\begin{aligned} D_k^m(\boldsymbol{a}, \boldsymbol{b}) \approx \bar{D}_k^m(\boldsymbol{a}) =\ & \sum_{i=1}^{r} \sum_{j=r+1}^{n} a_i a_j \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{\phi}_j\right) + \sum_{i=r+1}^{n} \sum_{j=1}^{r} a_i a_j \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{\phi}_j\right) + \sum_{i=r+1}^{n} \sum_{j=r+1}^{n} a_i a_j \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{\phi}_j\right) \\ & - \nu \sum_{i=r+1}^{n} a_i \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i\right) + \sum_{i=r+1}^{n} a_i \left(\boldsymbol{\phi}_k, \boldsymbol{u}_m \cdot \nabla \boldsymbol{\phi}_i\right) + \sum_{i=r+1}^{n} a_i \left(\boldsymbol{\phi}_k, \boldsymbol{\phi}_i \cdot \nabla \boldsymbol{u}_m\right). \end{aligned} \tag{29}$$

For practical near real-time prediction, it is desirable to have $r \ll n$, especially for non-linear equations where the cost of the non-linear terms can scale with a high exponent. For example, the cost of the online stage for the incompressible Navier–Stokes equations is $O(r^3)$, which is very expensive for large values of r. Unfortunately, this also implies that a cost-effective approach involves a relatively higher number of unresolved and fewer resolved states. As the dynamic evolution of unresolved states is not computed, the influence of these unresolved states on the resolved states are neglected. Therefore, if $r \ll n$, the accuracy of the

dynamics prediction is affected, especially for strongly advective flows where the decay of Kolmogorov n-width is slow [46]. If the influence of unresolved states is ignored by setting $a_i = 0$ and $b_i = 0$ for $i = r + 1, r + 2, \ldots, n$, the closure terms D_k^c and D_k^m become zero and the equations Eq. (22) and Eq. (23) reduce to

$$\left(\psi_k, \nabla \cdot \boldsymbol{u}_m\right) + \sum_{i=1}^{r} a_i \left(\psi_k, \nabla \cdot \boldsymbol{\phi}_i\right) = 0 \tag{30}$$

and

$$\dot{a}_k + \sum_{i=1}^{r} L_{ki} a_i + \sum_{i=1}^{r} \sum_{j=1}^{r} B_{kij} a_i a_j + \sum_{i=1}^{r} L_{ki}^P b_i + C_k = 0. \tag{31}$$

We refer to this as the Galerkin projection-based ROMs without closure modeling. Suppose the influence of unresolved states on the dynamics of resolved states is significant, the accuracy and stability of the predicted resolved states can be severely affected in these ROMs. Therefore, accurate models that approximate D_k^m without incurring a significant evaluation cost are essential. For incompressible flows, D_k^c may not severely affect the dynamics of a_i and b_i and is set to zero in this article.

4.2. Closure models for ROMs

As discussed in the introduction, several closure models have been proposed over the years that account for the interaction of the unresolved states with the resolved states. These models can be classified as physics-driven or data-driven. Physics-driven models involve analysis of equations and physical insights to determine a model form and calibrate the model parameters [2,6,7,18]. These model forms typically have fewer parameters to ease the process of calibrating them. Most common physics-driven models involve the assumption of a model form as a linear term in Eq. (23), that is

$$\bar{D}_k^m(\boldsymbol{a}) \approx D_k^{m-model}(\boldsymbol{a}) = \sum_{i=1}^{r} a_i \tilde{L}_{ki}, \tag{32}$$

where $\tilde{L}_{ki}$ is the unknown closure term. Three model forms can approximate $\tilde{L}_{ki}$. The first model form, which we refer to as the global artificial viscosity, approximates $\tilde{L}_{ki}$ as

$$\tilde{L}_{ki} = v^t \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i\right), \tag{33}$$

where v^t is the unknown global artificial viscosity. The second model form, which we refer to as the modal artificial viscosity, approximates $\tilde{L}_{ki}$ as

$$\tilde{L}_{ki} = v_k^t \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i\right), \tag{34}$$

where v_k^t is the unknown modal artificial viscosity. These two model forms are common in the literature [2,15,18]. In addition to these model forms, we propose a third model form, which we refer to as the tensor artificial viscosity, that approximates $\tilde{L}_{ki}$ as

$$\tilde{L}_{ki} = v_{ki}^t \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i\right), \tag{35}$$

where v_{ki}^t is the kith component of the unknown tensor artificial viscosity. Several strategies have been used to determine the unknown parameters in Eq. (33) and Eq. (34). The mean energy balance was used in [18] to obtain v^t. On the other hand, the mean of modal energy was used in [18,34] to obtain v_k^t. Data-assimilation techniques were used in [22,23] to obtain the relevant unknown closure term parameters. Recently, data-driven models that use high-fidelity simulation data and machine learning techniques have been proposed. These models approximate D_k^m using linear [30] or quadratic operators [25,31] instead of assuming a physical model form like artificial viscosity-based closure models. As projection-based ROMs are data-intensive and rely on data to obtain the energy-dominant spatial modes, substantial data needed for determining unknown parameters in closure models is readily available. Sensitivity analysis can also be used to identify the role of each component of tensor artificial viscosity on closure model performance and reduce the number of unknown closure model parameters. Such analysis can reduce the offline stage cost of determining model parameters, although the online stage cost for the closure models proposed in this article will remain the same.

Data-driven models [25,30,31] often used least squares minimization to obtain the unknown parameters in the model form. However, there is limited work on utilizing this calibration technique for determining the unknown parameters in physics-driven closure models [16]. Even in [16], only the modal artificial viscosity was determined using least squares minimization, but not used as a closure model in ROMs. Therefore, one of the main contributions of this article is to provide a systematic feasibility study on using least squares minimization for learning unknown parameters in three artificial viscosity closure model forms. The least squares minimization for obtaining unknown parameters can be applied in two ways. The first and more common calibration technique is to approximate the closure term $\bar{D}_k^m$ in Eq. (29) as accurately as possible. We refer to this technique as the closure term approximation. This technique is typically followed in the data-driven closure model literature. [25,30,31].

The alternate calibration technique accurately captures the energetic interactions between resolved and unresolved states. We refer to this technique as the energy approximation. The influence of closure terms on the energy equation can be determined by examining the energy equation. This equation can be derived from the Navier–Stokes equations by taking the inner product of Eq. (2) with velocity to obtain

$$\frac{\partial}{\partial t}(\boldsymbol{u}, \boldsymbol{u}) + \left(\boldsymbol{u}, \boldsymbol{u} \cdot \nabla \boldsymbol{u}\right) + \left(\boldsymbol{u}, \nabla p\right) - v \left(\boldsymbol{u}, \nabla^2 \boldsymbol{u}\right) = 0. \tag{36}$$

Table 1
List of the six closure models considered in this article.

Closure model	Model form	Calibration technique
GV-GE	Global artificial viscosity: Eq. (33)	Global energy: $\sum_{k=1}^{r} a_k \bar{D}_k^m$
MV-ME	Modal artificial viscosity: Eq. (34)	Modal energy: $a_k \bar{D}_k^m$
TV-ME	Tensor artificial viscosity: Eq. (35)	Modal energy: $a_k \bar{D}_k^m$
GV-C	Global artificial viscosity: Eq. (33)	
MV-C	Modal artificial viscosity: Eq. (34)	Closure term: $\bar{D}_k^m$, Eq. (29)
TV-C	Tensor artificial viscosity: Eq. (35)	

Substituting POD for velocity and pressure in this equation and simplifying it, we get

$$\sum_{k=1}^{r} \frac{da_k^2}{dt} + \sum_{i=1}^{r}\sum_{k=1}^{r} a_i a_k L_{ki} + \sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} a_i a_j a_k B_{kij} + \sum_{i=1}^{r}\sum_{k=1}^{r} b_i a_k L_{ki}^p + \sum_{k=1}^{r} a_k (C_k + \bar{D}_k^m(\boldsymbol{a})) = 0. \tag{37}$$

The energetic interactions of the unresolved states and their contribution towards the energy of all resolved states are accounted for in the last term: $\sum_{k=1}^{r} a_k \bar{D}_k^m$. Each component of this term, $a_k \bar{D}_k^m$, accounts for energetic interactions of unresolved states to the energy of the resolved state a_k. These interactions comprise dyadic energetic interactions due to linear terms and triadic energetic interactions due to non-linear terms. Similarly, the energy equation for the kth resolved state is

$$\frac{da_k^2}{dt} + a_k \sum_{i=1}^{r} a_i L_{ki} + a_k \sum_{i=1}^{r}\sum_{j=1}^{r} a_i a_j B_{kij} + a_k \sum_{i=1}^{r} b_i L_{ki}^p + a_k C_k + a_k \bar{D}_k^m(\boldsymbol{a}) = 0. \tag{38}$$

The modal energy equation determines the temporal evolution of energy for each resolved state. Based on these two equations, the energy approximation calibration technique for determining unknown parameters can be further classified as global and modal energy approximation techniques. The global energy approximation aims to ensure that the global energy balance (in Eq. (37)) is accurate. Therefore, this technique estimates unknown model parameters such that $\sum_{k=1}^{r} a_k \bar{D}_k^m$ is accurately calculated. On the other hand, the modal energy approximation involves ensuring that the modal energy balance is accurate. Therefore, this technique estimates unknown model parameters such that $a_k \bar{D}_k^m$ is accurately calculated. Determining parameters using global energy approximation is appealing for a global artificial viscosity approximation. Meanwhile, determining parameters using modal energy approximation is a better-posed problem for modal and tensor artificial viscosity models.

The two alternate pathways for calibrating closure models, accurate closure term and accurate energy contribution, will yield different model parameters, and it is important to study which calibration technique is better suited. Combining three closure model forms, namely global, modal, and tensor eddy artificial viscosity model forms, with these two least squares minimization-based calibration techniques, namely minimization of error in energy approximation and closure term approximation, results in six possible combinations of data-driven artificial viscosity closure models as shown in Table 1. In this article, we pose the closure model calibration process as a linear regression problem which provides a deterministic model without uncertainty estimates. A Bayesian regression-based approach, such as the one used in [47], can also be used to calibrate closure models while providing uncertainty estimates. Only the former approach is considered in this article, although the latter one is an ideal extension that will be considered in the future.

4.2.1. GV-GE model

GV-GE model assumes the model form in Eq. (33), where the optimal value of the global artificial viscosity (ν^t) ensures that the contribution from the closure term on global energy $\sum_{k=1}^{r} a_k \bar{D}_k^m(\boldsymbol{a})$ is accurately estimated. Mathematically, this model is posed as

$$\nu^t = \arg\min_{\hat{\nu}} \left\| \hat{\nu} \sum_{i=1}^{r}\sum_{k=1}^{r} a_i a_k \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i \right) - \sum_{k=1}^{r} a_k \bar{D}_k^m(\boldsymbol{a}) \right\|^2. \tag{39}$$

The same data used for obtaining POD modes is used to obtain projected reduced states $a_i(t)$ at time $t = t_l$ where $l = 1, 2, \ldots, n$. Using this data, this problem transforms to

$$\nu^t = \arg\min_{\hat{\nu}} \left\| \hat{\nu}\boldsymbol{z} - \boldsymbol{b} \right\|_2^2, \tag{40}$$

where the lth component of the vector $\boldsymbol{z}$ is

$$z_l = \sum_{i=1}^{r}\sum_{k=1}^{r} a_i(t_l) a_k(t_l) \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i \right) \tag{41}$$

and the lth component of the vector $\boldsymbol{b}$ is

$$b_l = \sum_{k=1}^{r} a_k(t_l) \bar{D}_k^m(\boldsymbol{a}(t_l)). \tag{42}$$

The solution of this regression problem yields the following expression for artificial viscosity:

$$v^t = \left(z^T z\right)^{-1} z^T b. \tag{43}$$

For this model, we obtain a single artificial velocity that best approximates the energy at all sampled timesteps. If the flow is not stationary, the optimal global viscosity may not be the best approximation of the energetic interactions at different time instances. Therefore, this closure model is more suited for stationary flows. Despite the model form appearing very similar to the model in [18], this model differs in the calibration technique to determine the value of v^t. In [18], the mean value of global energy is used to determine v^t, whereas raw data of global energy with least squares minimization is used in this article.

4.2.2. MV-ME model

The MV-ME model assumes the model form in Eq. (34), where the optimal value of the modal artificial viscosity (v_k^t) ensures that the contribution of closure term on modal energy $a_k \bar{D}_k^m(a)$ is accurately estimated. Mathematically, this model is posed as

$$v_k^t = \arg \min_{\hat{v}} \left\| \hat{v} a_k \sum_{i=1}^{r} a_i \left(\phi_k, \nabla^2 \phi_i\right) - a_k \bar{D}_k^m(a) \right\|^2. \tag{44}$$

Following the procedure for other models, we use the sampled data. This problem transforms to

$$v_k^t = \arg \min_{\hat{v}} \left\| \hat{v} z - b \right\|_2^2, \tag{45}$$

where the lth component of the vector z is

$$z_l = a_k(t_l) \sum_{i=1}^{r} a_i(t_l) \left(\phi_k, \nabla^2 \phi_i\right) \tag{46}$$

and the lth component of the vector b is

$$b_l = a_k(t_l) \bar{D}_k^m(a(t_l)). \tag{47}$$

The solution of this regression problem yields the following expression for artificial viscosity for the kth mode:

$$v_k^t = \left(z^T z\right)^{-1} z^T b. \tag{48}$$

This model can better approximate the dynamic interaction between unresolved and resolved states than the GV-GE model, as it is more expressive. Although the model form appears very similar to the model in [18], this model differs in the calibration technique used to determine the v_k^t value. In [18], the mean value of modal energy is used to determine v_k^t, whereas least squares minimization with raw data of modal energy is directly used in this article. This model is similar to the one mentioned in [16] where the least squares minimization is used to obtain v_k^t. However, the applicability of this closure model for ROMs was not demonstrated in [16].

4.2.3. TV-ME model

The TV-ME model assumes the model form in Eq. (35), where the optimal value of the tensor artificial viscosity (v_{ki}^t) is obtained to ensure that the contribution of closure term on modal energy $a_k \bar{D}_k^m(a)$ is accurately estimated. Mathematically, this model is posed as

$$v_k^t = \arg \min_{\hat{v} = [\hat{v}_1, \hat{v}_2, \dots, \hat{v}_r]} \left\| a_k \sum_{i=1}^{r} \hat{v}_i a_i \left(\phi_k, \nabla^2 \phi_i\right) - a_k \bar{D}_k^m(a) \right\|^2 \tag{49}$$

with v_{ki}^t as the ith element of vector v_k^t. Following the procedure for other models, we use the sampled data. This problem transforms to

$$v_k^t = \arg \min_{\hat{v}} \left\| \hat{v} Z - b \right\|_2^2, \tag{50}$$

where the lith component of the matrix Z is

$$Z_{li} = a_k(t_l) a_i(t_l) \left(\phi_k, \nabla^2 \phi_i\right) \tag{51}$$

and the lth component of the vector b is given in Eq. (47). The solution of this regression problem yields the following expression for vector-valued artificial viscosity for the kth mode:

$$v_k^t = \left(Z^T Z\right)^{-1} Z^T b. \tag{52}$$

This model can better approximate the dynamic interaction between unresolved and resolved states than GV-GE and MV-ME models as it is more expressive.

4.2.4. GV-C model

The GV-C model assumes the model form in Eq. (33), where the optimal value of the artificial viscosity (v^t) ensures that the closure term $\bar{D}_k^m$ is accurately estimated. Mathematically, this model is posed as

$$v^t = \arg\min_{\hat{v}} \left\| \hat{v} \sum_{i=1}^{r} a_i \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i \right) - \bar{D}_k^m(\boldsymbol{a}) \right\|^2, \tag{53}$$

where data from all r modes is used to learn the artificial viscosity. Following the procedure for other models, we use the sampled data. This problem transforms to Eq. (40), where the oth component of the vector $\boldsymbol{z}$ is

$$z_o = \sum_{i=1}^{r} a_i(t_l) \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i \right) \tag{54}$$

and the oth component of the vector $\boldsymbol{b}$ is

$$b_o = \bar{D}_k^m(\boldsymbol{a}(t_l)), \tag{55}$$

where o is a unique combination of k and l. The solution to this regression problem, shown in Eq. (43), provides the value of v^t. For this model, we identify a single artificial velocity that best approximates the closure term for all the modes at every time step. Therefore, if the flow is not stationary, the optimal global viscosity may not best approximate the closure term at different time instances.

4.2.5. MV-C model

The MV-C model assumes the model form in Eq. (34) where the optimal value of the modal artificial viscosity (v_k^t) ensures that the closure term $\bar{D}_k^m$ is accurately estimated. Mathematically, this model is posed as

$$v_k^t = \arg\min_{\hat{v}} \left\| \hat{v} \sum_{i=1}^{r} a_i \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i \right) - \bar{D}_k^m(\boldsymbol{a}) \right\|^2. \tag{56}$$

Following the procedure for other models, we use the sampled data. This problem transforms to Eq. (45), where the lth component of the vector $\boldsymbol{z}$ is

$$z_l = \sum_{i=1}^{r} a_i(t_l) \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i \right) \tag{57}$$

and the lth component of the vector $\boldsymbol{b}$ is

$$b_l = \bar{D}_k^m(\boldsymbol{a}(t_l)). \tag{58}$$

The solution to this regression problem, shown in Eq. (48), provides the value of v_k^t. This model better approximates the dynamic interaction between unresolved and resolved states than the GV-C model, as it is more expressive. This model is similar to the one mentioned in [16], although the applicability of the closure model for ROM was not demonstrated in that article.

4.2.6. TV-C model

The TV-C model assumes the model form in Eq. (35), where the optimal value of the artificial viscosity (v_k^t) ensures that the closure term $\bar{D}_k^m$ is accurately estimated. Mathematically, this model is posed as

$$v_k^t = \arg\min_{\hat{\boldsymbol{v}}=[\hat{v}_1, \ \hat{v}_2,...,\hat{v}_r]} \left\| \sum_{i=1}^{r} \hat{v}_i a_i \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i \right) - \bar{D}_k^m(\boldsymbol{a}) \right\|^2 \tag{59}$$

with v_{ki}^t as the ith element of vector v_k^t. Following the procedure for other models, we use the sampled data. This problem transforms to Eq. (50), where the lith component of the matrix $\boldsymbol{Z}$ is

$$Z_{li} = a_i(t_l) \left(\boldsymbol{\phi}_k, \nabla^2 \boldsymbol{\phi}_i \right) \tag{60}$$

and the lth component of the vector $\boldsymbol{b}$ is given in Eq. (58). The solution to this regression problem, shown in Eq. (52), provides the value of v_k^t. This model better approximates the dynamic interaction between unresolved and resolved states than GV-C and MV-C models as it is more expressive.

4.2.7. Differences compared to other closure modeling techniques

As discussed earlier, several other closure models exist for ROMs. Physics-driven models are better posed as the model form is determined using the physical relationship between variables; for example, global and modal artificial viscosity models have been demonstrated to tackle the closure problem adequately [10,17,18]. In contrast to the conventional physics-driven models like Eq. (33) and Eq. (34), the tensor artificial viscosity model form in Eq. (35) is more expressive and can potentially give better accuracy. On the other hand, data-driven closure models are naturally suited for ROMs due to their data-intensive nature. However, common data-driven closure models that rely on least squares regression for closure modeling require regularization to remedy the ill-conditioning in the model learning problem [25]. The selection of the ideal regularization strategies and constants varies with the problem definitions and introduces an additional expensive tuning process. For example, the models proposed in [25,31] have more

unknown parameters; thereby, they are more expressive but also suffer from ill-conditioning and require regularization. The linear operator data-driven model in [30] has fewer parameters and does not require regularization for flows considered in this article. Numerical tests not shown in this article for brevity indicated that this linear operator model yielded similar results to the TV-C model and better than models with quadratic operator model form [25,31]. The closure models discussed in this article consider several physics-driven model forms discussed above and determine unknown parameters using least squares regression, which is more commonly used for data-driven closure models.

4.3. Incremental pressure correction schemes for projection-based reduced order modeling

After including the closure model, the final ROM equations to be solved are

$$\left(\psi_k, \nabla \cdot \boldsymbol{u}_m\right) + \sum_{i=1}^{r} a_i \left(\psi_k, \nabla \cdot \boldsymbol{\phi}_i\right) = 0 \tag{61}$$

and

$$\dot{a}_k + \sum_{i=1}^{r} L_{ki} a_i + \sum_{i=1}^{r} \sum_{j=1}^{r} B_{kij} a_i a_j + \sum_{i=1}^{r} L_{ki}^{P} b_i + C_k + \bar{D}_k^{m-model}(\boldsymbol{a}) = 0, \tag{62}$$

where $\bar{D}_k^{m-model}$ is determined using one of the proposed closure models. The solution of these equations also exhibits a saddle-point form like those exhibited by incompressible Navier–Stokes equations [36]. No equation dictates the evolution of b_i, and hence, several alternate solution methods have been proposed to solve these problems [39,48] in the context of FOM. Most studies on ROMs for incompressible Navier–Stokes equations assume the discrete divergence-free nature of the velocity and thereby neglect the pressure gradient term in the momentum equation [35]. Unfortunately, most of the practical data sources, both computational and experimental, do not satisfy the discrete divergence-free condition of velocity. Furthermore, the pressure gradient term in the momentum equation significantly influences flow behavior for several fluid flows, such as confined flows [34]. In such scenarios, appropriate treatment of pressure–velocity coupling is essential.

Our solution method for these equations is inspired by the pressure-projection method used to solve incompressible Navier–Stokes equations. These approaches are preferred for efficiently solving large-scale incompressible flow problems [39]. The most commonly used solution method is the classical pressure correction scheme in [49,50], which uncouples pressure and velocity to avoid the solution of the saddle point problem. However, this method is first order in time and our implementation required a small time step for convergence of results. Therefore, we use incremental pressure correction that is second order in time [51]. For incompressible Navier–Stokes equations, this method involves sub-steps

$$\frac{3\tilde{\boldsymbol{u}}^{n+1} - 4\boldsymbol{u}^n + \boldsymbol{u}^{n-1}}{2\Delta t} + \tilde{\boldsymbol{u}}^{n+1} \cdot \nabla \tilde{\boldsymbol{u}}^{n+1} + \nabla p^n - \nu \nabla^2 \tilde{\boldsymbol{u}}^{n+1} = 0, \tag{63}$$

$$\frac{3\boldsymbol{u}^{n+1} - 3\tilde{\boldsymbol{u}}^{n+1}}{2\Delta t} + \nabla p^{n+1} - \nabla p^n = 0 \tag{64}$$

and

$$\nabla \cdot \boldsymbol{u}^{n+1} = 0. \tag{65}$$

A similar solution method can be used for Galerkin projection-based ROMs by substituting Eq. (3) in Eq. (63), Eq. (64) and Eq. (65). The resulting momentum equation, Eq. (63), is projected onto the linear subspace spanned by the $\boldsymbol{\phi}_i$ for $i = 1, 2, \ldots, r$. We first take the divergence of Eq. (64) and then project the resulting equation along with the continuity equation Eq. (65) onto the linear subspace spanned by the ψ_i for $i = 1, 2, \ldots, r$. The resulting equations are

$$\frac{3\tilde{a}_k^{n+1} - 4a_k^n + a_k^{n-1}}{2\Delta t} + \sum_{i=1}^{r} \tilde{a}_i^{n+1} L_{ki} + \sum_{i=1}^{r} \sum_{j=1}^{r} \tilde{a}_i^{n+1} \tilde{a}_j^{n+1} B_{kij} + \sum_{i=1}^{r} b_i^n L_{ki}^{P} + C_k + \bar{D}_k^{m-model}(\tilde{\boldsymbol{a}}^{n+1}) = 0, \quad \text{for} \quad k = 1, 2, \ldots, r, \tag{66}$$

$$\sum_{i=1}^{r} (a_i^{n+1} - \tilde{a}_i^{n+1})(\psi_k, \nabla \cdot \boldsymbol{\phi}_i) - \frac{2\Delta t}{3} \sum_{i=1}^{r} (b_i^{n+1} - b_i^n)(\nabla \psi_k, \nabla \psi_i) = 0, \quad \text{for} \quad k = 1, 2, \ldots, r \tag{67}$$

and

$$\left(\psi_k, \nabla \cdot \boldsymbol{u}_m\right) + \sum_{i=1}^{r} a_i^{n+1}(\psi_k, \nabla \cdot \boldsymbol{\phi}_i) = 0, \tag{68}$$

where $a_k^n = a_k(t_n)$ and $b_k^n = b_k(t_n)$. As discussed in Section 2, *inf–sup* condition restricts the choice of function spaces for pressure and velocity. For example, mixed elements such as Taylor–Hood elements [52] are often used to overcome this restriction. On the other hand, if equal order interpolation functions are chosen for velocity and pressure, additional stabilization techniques, such as the pressure-stabilizing Petrov–Galerkin (PSPG) method [48], may be needed. In the context of ROMs, the spatial basis for both pressure and velocity is determined from data, and there is limited literature on identifying adequate function spaces for velocity and pressure POD basis. The literature in this area highlights using stabilization techniques [8,35] and supremizers [36,53]. In the context of FOMs, such as those using finite element methods, pressure-projection schemes have been shown to overcome the

restriction posed by *inf–sup* condition on selecting velocity and pressure spaces [54]. Therefore, these schemes are also expected to work well for ROMs and this article demonstrates this through numerical simulations. Furthermore, decoupling velocity and pressure states make the solution process efficient especially where a large number of reduced states are considered. As pressure correction schemes [49,50] are shown to have a similar form as the PSPG scheme [48], pressure correction schemes could implicitly act as a model for unresolved pressure states [55] and thereby improve pressure stability. Detailed numerical analysis of the stability and convergence of this method in the context of ROMs will be considered a future study.

Despite good predictions of a_i using the incremental pressure correction scheme for ROMs, the accurate temporal evolution of b_i was not observed for the flow under consideration in this article. This observation could be attributed to the role of pressure to enforce the incompressibility condition which appears to be moderately satisfied for continuous ROMs for the flow under consideration. As the temporal update of reduced pressure states depends on the divergence of velocity basis, as seen in Eq. (67), a near divergence free velocity basis implies that pressure is not updated correctly. Therefore, the evolution of b_i is instead obtained in the post-processing stage by utilizing the a_i and solving the pressure-Poisson equation as also done in [34,56]. These equations are obtained by substituting Eq. (3) in the pressure-Poisson equation

$$\nabla^2 p = -\nabla \cdot (\boldsymbol{u} \cdot \nabla \boldsymbol{u}) \tag{69}$$

and then projecting it on a linear subspace formed by ψ_i for $i = 1, 2, \ldots, r$, resulting in

$$\left(\psi_k, \nabla^2 p_m\right) + \sum_{i=1}^{r} b_i \left(\psi_k, \nabla^2 \psi_i\right) + \sum_{i=1}^{r} \sum_{j=1}^{r} a_i a_j \left(\psi_k, \nabla \cdot (\boldsymbol{\phi}_i \cdot \nabla \boldsymbol{\phi}_j)\right) + \sum_{i=1}^{r} a_i \left(\psi_k, \nabla \cdot (\boldsymbol{\phi}_i \cdot \nabla \boldsymbol{u}_m)\right)$$

$$+ \sum_{i=1}^{r} a_i \left(\psi_k, \nabla \cdot (\boldsymbol{u}_m \cdot \nabla \boldsymbol{\phi}_i)\right) + \left(\psi_k, \nabla \cdot (\boldsymbol{u}_m \cdot \nabla \boldsymbol{u}_m)\right) = 0, \quad \text{for} \quad k = 1, 2, \ldots, r. \tag{70}$$

Note that the values of b_i obtained through this post-processing step are solely used to compute the pressure distribution and not used in the evolution of velocity states. The solution of the pressure-Poisson equation costs $O(r^3)$. Therefore, this step can be avoided if only post-processing of velocity is needed. The use of pressure-Poisson for post-processing to obtain pressure is also common when a discretely divergence free condition is used to simulate incompressible flows, as pressure is not computed during the solution stage. On the other hand, other common solution methods for resolved pressure states [35] typically involve the solution of the pressure-Poisson equation during the solution step.

5. Numerical experiments

This section aims to validate the performance of the ROMs with proposed closure models and compare the results to ROMs without any closure model. First, we describe the details of the FOM simulation setup for the validation case: flow over a 2-D cylinder [57]. This flow has been frequently considered for validating ROM solution methods and closure models [25,31]. Second, we give details for the construction of the operators and time evolution of reduced states. Third, we compare the performance of proposed closure models at three Reynolds numbers: $Re = 200$, $Re = 500$ and $Re = 1000$.

5.1. Full order model details

FOM simulations use a finite element code developed using the DOLFINx library [58]. The finite element code closely follows openly available tutorials in [59,60]. The commonly used Taylor–Hood elements (P2-Q1) are chosen as the spatial discretization for the FOM simulation. The problem can be posed as:

Find $\boldsymbol{u} \in \mathcal{V}^h \subset \mathcal{V} = (H^1(\Omega))^2$ and $p \in \mathcal{Q}^h \subset \mathcal{Q} = \{q \in L^2(\Omega) : \int_{\Omega} q dx = 0\}$, such that

$$\left(\boldsymbol{v}, \frac{\partial \boldsymbol{u}}{\partial t}\right) + \left(\boldsymbol{v}, \boldsymbol{u} \cdot \nabla \boldsymbol{u}\right) = -\left(\boldsymbol{v}, \nabla p\right) + \left(\boldsymbol{v}, \nu \nabla^2 \boldsymbol{u}\right), \quad \forall \quad \boldsymbol{v} \in \mathcal{V}^h \tag{71}$$

and

$$\left(q, \nabla \cdot \boldsymbol{u}\right) = 0 \quad \forall \quad q \in \mathcal{Q}^h \tag{72}$$

subject to boundary conditions

$$\boldsymbol{u}|_{\partial \Omega_D} = g \quad \text{and} \quad p|_{\partial \Omega_N} = h, \tag{73}$$

where $\partial \Omega_D$ is the Dirichlet boundary and $\partial \Omega_N$ is the Neumann boundary. The Dirichlet boundary for the flow over a 2-D cylinder that encompasses of wall boundary (Ω_w), inlet boundary (Ω_{in}), and outlet boundary (Ω_{ou}), that is $\partial \Omega_D = \partial \Omega_w \bigoplus \partial \Omega_{in} \bigoplus \partial \Omega_{ou}$. The domain and locations of the boundaries are shown in Fig. 1. The wall boundary condition is

$$\boldsymbol{u}|_{\partial \Omega_w} = 0, \tag{74}$$

the inlet boundary condition is

$$\boldsymbol{u}|_{\partial \Omega_{in}} = \left[4U_0 \frac{y(0.41 - y)}{0.41^2}, 0\right], \quad \text{where} \quad U_0 = 1.5 \text{ m/s}, \tag{75}$$
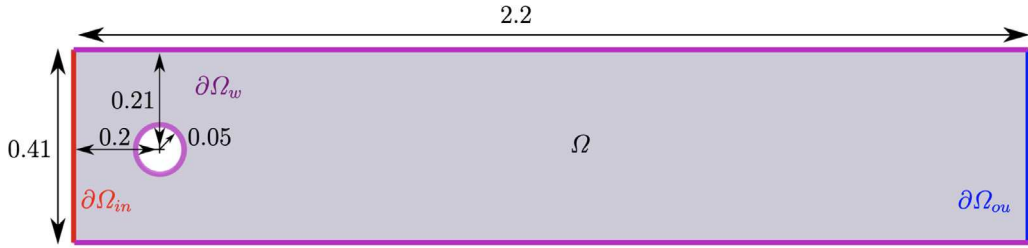
Fig. 1. Domain for the flow over a 2-D cylinder.

and the outlet boundary condition is

$$p|_{\partial \Omega_{ou}} = 0. \tag{76}$$

The simulation setup is similar to the well-known benchmark [57]. Unlike the benchmark, which uses a sinusoidally varying velocity, the flow considered in this article has a constant inflow velocity. These equations are solved using the Crank–Nicholson and semi-implicit Adams–Bashforth time integration scheme with the above boundary conditions. The mesh for $Re = 200$ was composed of quadrilateral elements formed using approximately 14,000 nodes. Meanwhile, the $Re = 500$ and 1000 meshes comprised around 22,000 nodes. This mesh resolution possibly underresolves the flow at the given Reynolds numbers. However, this mesh works well to demonstrate the applicability of the closure model and ROM solution method for incompressible fluid flows. The FOM simulation is started from a zero velocity and pressure initial condition. The simulation is integrated in time until $T_f = 10s$ with a timestep of $1/1600$. For $Re = 500$ and 1000, a lower timestep was also tested to ensure that the lower Courant–Friedrichs–Lewy (CFL) number remains similar between simulations of different Reynolds numbers. However, the behavior of FOM simulations and ROMs remained insensitive to this selection.

5.2. Reduced order modeling details

The offline step of ROMs involves the computation of operators as these are expensive to compute, for example, the non-linear term of Navier–Stokes equations in our case. ROMs can be categorized as discrete ROMs or continuous ROMs following the approach taken to construct operators in Eq. (66), Eq. (67) and Eq. (68). The key differences between these ROMs are discussed in [13,61]. In this article, we work with continuous ROMs as it is mathematically more consistent with FOM than discrete ROMs [13,62]. Continuous ROMs leverage interpolation and numerical integration routines used in the corresponding FOM solver, favoring the consistency between ROMs and FOMs. The first step for assembling these ROMs involves interpolating the POD basis, ϕ and ψ, to a mesh with a suitable polynomial degree. The polynomial degree for the spatial discretization can be arbitrarily chosen but could introduce consistency issues if not matched with the spatial discretization in the FOM [62]. In this article, a more consistent approach is undertaken by selecting function spaces for ϕ and ψ to be the same as those for $\boldsymbol{u}$ and p, that is, $\mathcal{V}^h$ and $\mathcal{Q}^h$ respectively. Additionally, we use the same quadrature rule as FOMs for operator integrals.

The online step of ROMs involves the time integration of the dynamical equations of reduced states. This article uses a second-order implicit time stepping scheme as discussed in Section 4.3. As this scheme is not the same as the time discretization used for the FOM, some inconsistency in time discretization may be introduced. This decision was still made to demonstrate the applicability of pressure correction schemes for the temporal evolution of reduced states in ROMs. The equations Eq. (66), Eq. (67) and Eq. (68) are solved using SNES non-linear solver from PETSc [63]. The time step for ROMs was chosen to be the same as the one for FOMs after a systematic timestep convergence study, which is not included in this article for brevity. The initial transient of the FOM simulation, that is, $t \in [0, 5]$ s, is ignored. The data generated by FOM is uniformly sampled in a time interval of $T_{pred} = t \in [5, 10]$ s and used for testing the ROMs. The POD basis, equation operators and closure models are obtained using uniformly sampled data in $t \in [5, 6]$ s. Data at every 20th timestep is selected in this time interval, resulting in 80 data points. Other time intervals and data selection frequencies were also considered and yielded similar conclusions. The goal of these computational tests is to demonstrate the ability of ROM for forecasting problems. For all Reynolds numbers, we are interested in answering the question: Do ROMs obtained using data from $t \in [5, 6]$ s give good predictions at a future time: $t \in [5, 10]$ s (renamed as $t \in [0, 5]$ s)?

5.3. Cylinder flow at $Re = 200$

We compare the energy (E), drag coefficient (C_D), and lift coefficient (C_L) for different ROMs with corresponding values of FOM. These are defined as

$$E(\boldsymbol{u}, t) = \Big(\boldsymbol{u}(\boldsymbol{x}, t), \boldsymbol{u}(\boldsymbol{x}, t) \Big), \tag{77}$$

$$C_D(\boldsymbol{u}, p, t) = \frac{2}{LU_0^2} \int_{\partial \Omega_S} \nu n \cdot \nabla u_{t_s}(t) n_y - p(t) n_x \, ds \tag{78}$$
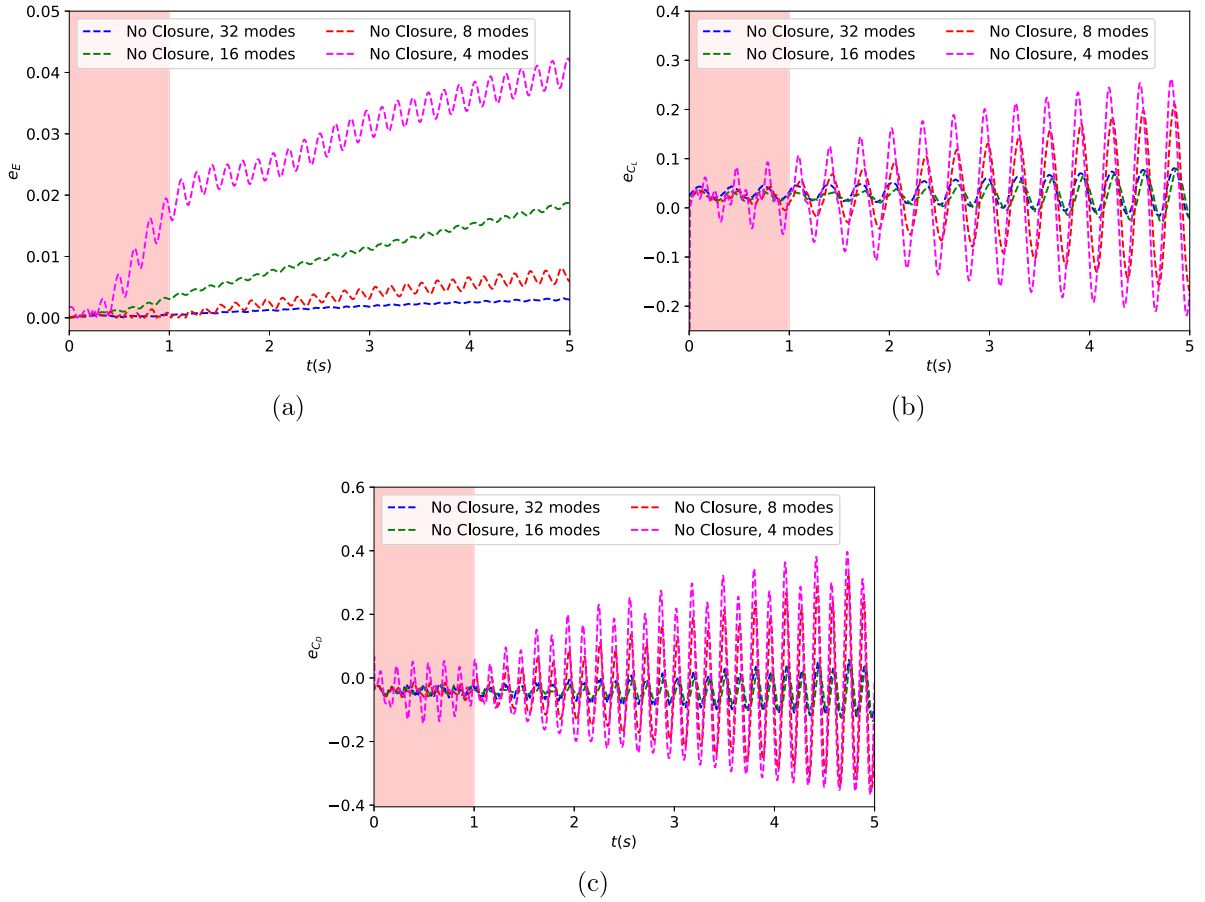
Fig. 2. Cylinder flow at Re = 200: Relative errors in (a) energy, (b) C_L and (c) C_D for Galerkin ROM without a closure model. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.

and

$$C_L(\boldsymbol{u}, p, t) = -\frac{2}{LU_0^2} \int_{\partial\Omega_S} \nu n \cdot \nabla u_{t_s}(t) n_x + p(t) n_y \, ds, \qquad (79)$$

where $\partial\Omega_S$ is the surface of the cylinder and $u_{t_s}(= \boldsymbol{u} \cdot (n_y, -n_x))$ is the tangential component of the velocity with n_x and n_y as the components of the normal to the cylinder surface. In particular, we compare relative errors in energy (e_E), relative errors in C_D (e_{C_L}) and relative errors in C_L (e_{C_L}) which are defined as

$$e_E(t) = \frac{E(\boldsymbol{u}^{FOM}, t) - E(\boldsymbol{u}^{ROM}, t)}{E(\boldsymbol{u}^{FOM}, t)}, \qquad (80)$$

$$e_{C_L}(t) = \frac{C_L(\boldsymbol{u}^{FOM}, p^{FOM}, t) - C_L(\boldsymbol{u}^{ROM}, p^{ROM}, t)}{\max_t(C_L(\boldsymbol{u}^{FOM}, p^{FOM}, t)) - \min_t(C_L(\boldsymbol{u}^{FOM}, p^{FOM}, t))} \qquad (81)$$

and

$$e_{C_D}(t) = \frac{C_D(\boldsymbol{u}^{FOM}, p^{FOM}, t) - C_D(\boldsymbol{u}^{ROM}, p^{ROM}, t)}{\max_t(C_D(\boldsymbol{u}^{FOM}, p^{FOM}, t)) - \min_t(C_D(\boldsymbol{u}^{FOM}, p^{FOM}, t))}. \qquad (82)$$

Note that relative errors in C_L and C_D have a mean bias introduced due to the difference in pressure computation between ROMs and FOM simulations. ROM determines pressure by solving pressure-Poisson equation Eq. (70) without closure models, whereas no such equations are solved in FOM simulations.

We first demonstrate the performance of the ROM without using a closure model. The evolution of relative errors in energy, C_L and C_D for ROMs without a closure model for different numbers of modes is shown in Fig. 2. For all selection of modes, we observe an increase in the error with an increase in time. Using eight modes for the ROM gives better energy prediction than using 16 modes. This result implies that increasing the number of modes for representing the solution does not improve the errors in
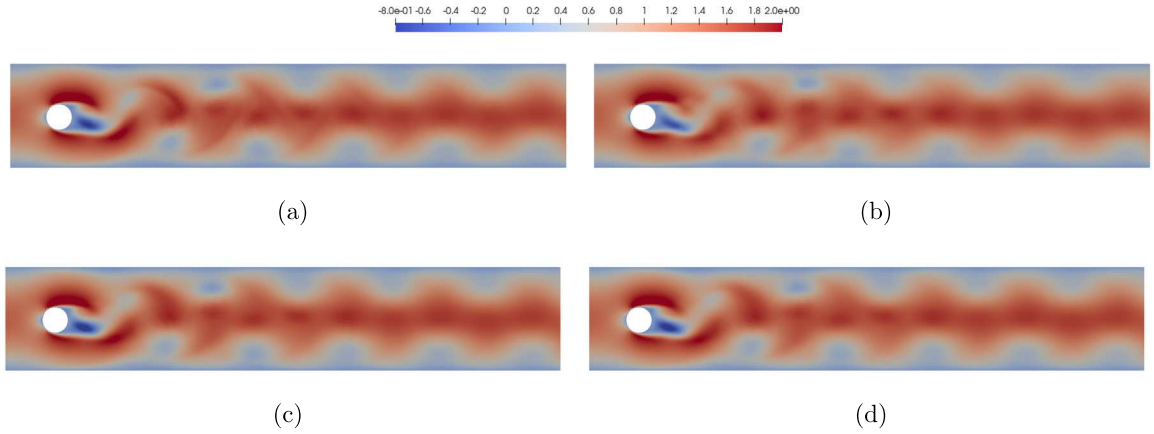
Fig. 3. Cylinder flow at Re = 200: Streamwise velocity component (u_1) field at t = 5s for (a) FOM, (b) Galerkin ROM with no closure model, (c) Galerkin ROM with the TV-ME closure model and (d) Galerkin ROM with the TV-C closure model when $r = 4$ modes are used.
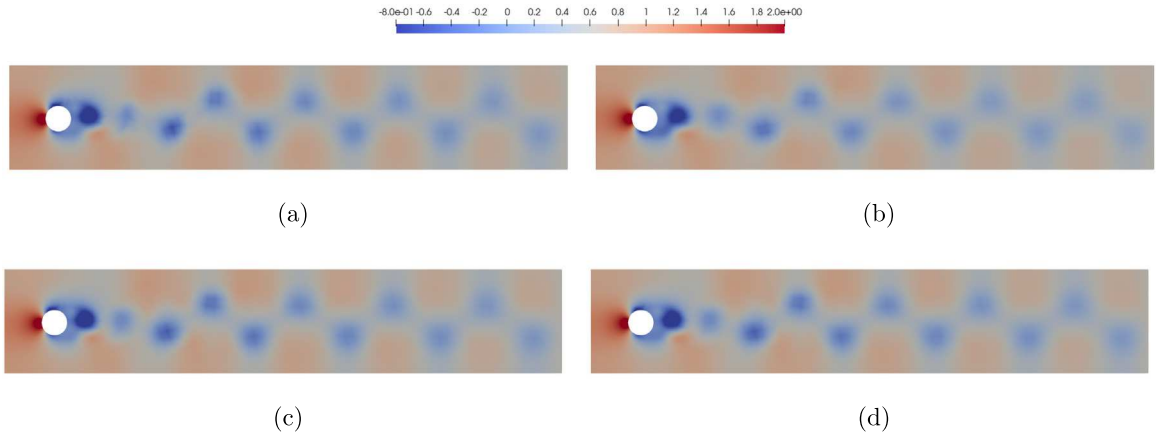

Fig. 4. Cylinder flow at Re = 200: Pressure field at t = 5s for (a) FOM, (b) Galerkin ROM with no closure model, (c) Galerkin ROM with the TV-ME closure model and (d) Galerkin ROM with the TV-C closure model when $r = 4$ modes are used.

energy approximation in time. On the other hand, results for C_D and C_L give a more consistent picture as the error for eight modes is much higher than the error for 16 modes. The error in these aerodynamic coefficients reduces significantly with an increase in the number of modes.

The numerical results indicate that the incremental pressure correction scheme gives adequately accurate velocity and pressure predictions for the incompressible fluid flow under consideration in this article. These results also indicate that ROMs would require many modes to achieve accurate predictions without using any closure model. As the online cost of ROMs scales as r^3, where r is the number of modes, using many modes to achieve more accurate results becomes computationally expensive and is not ideal. These observations and drawbacks of Galerkin ROMs are well-known in the literature. This discussion motivates the need for closure models that ensure higher accuracy without increasing the online cost of ROMs.

The streamwise velocity component and pressure fields at $t = 5$ s are shown in Fig. 3 and Fig. 4 respectively. The pressure distribution for several ROMs tested in this study is similar to those for FOM. These observations highlight the robustness of pressure prediction even without closure models. As discussed in Section 4.3, the numerical solution scheme may implicitly act as a model for unresolved pressure states, thereby providing good pressure predictions. ROMs without a closure model lead to a slightly different prediction of the streamwise velocity in the wake region closer to the cylinder. ROMs with TV-ME and TV-C closure models in this region yield a closer wake structure to FOM. In the far-wake region, the flow structures predicted by FOMs and ROMs slightly differ with TV-ME and TV-C closure models; however, these are still in better agreement compared to ROMs without a closure model used. These results demonstrate the applicability of the incremental pressure correction scheme-based ROMs to predict similar flow structures as FOMs.

The temporal evolution of relative errors in energy for four and six mode ROMs with different closure models is shown in Fig. 5. A significant error growth is observed when no closure model is used for four mode ROMs. ROMs with GV-GE and GV-C closure models provide even higher errors than those without a closure model. Similarly, ROMs with MV-ME and MV-C closure models
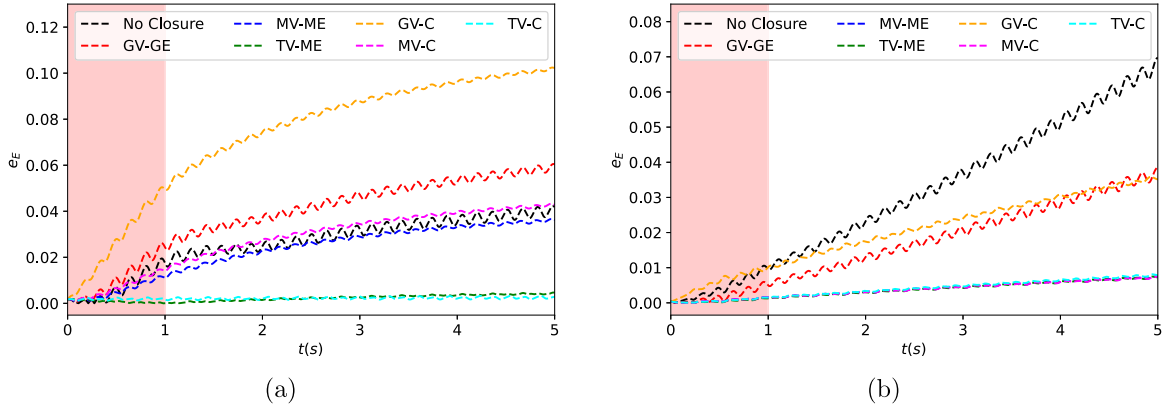
Fig. 5. Cylinder flow at Re = 200: Relative errors in energy (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.
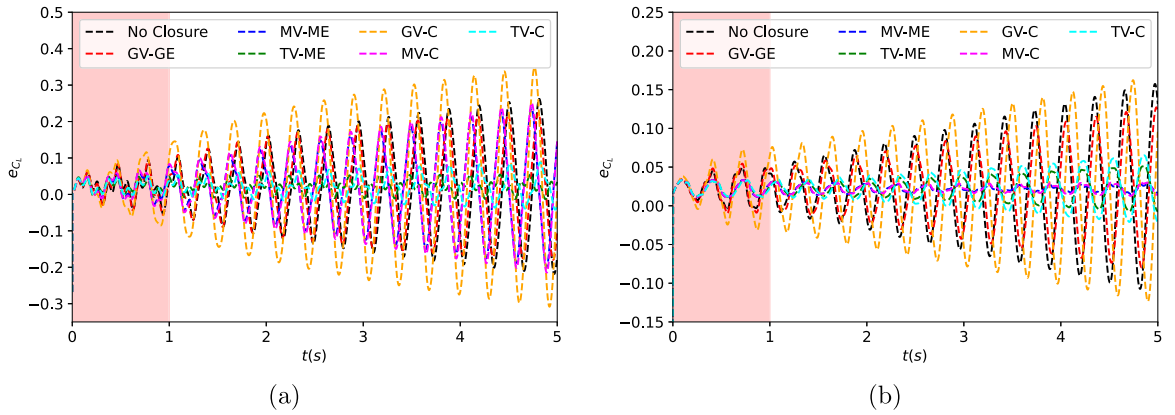

Fig. 6. Cylinder flow at Re = 200: Relative errors in C_L (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.

give errors similar to ROMs without a closure model. ROMs with TV-ME and TV-C closure models provide significantly lower errors compared to ROMs with other and without closure models. ROMs with any of the proposed closure models give better energy prediction and lower error than ROMs without a closure model for six mode ROMs. ROMs with GV-GE and GV-C closure models exhibit higher errors than ROMs with other closure models. ROMs with MV-ME, MV-C, TV-ME and TV-C closure models show very low and almost overlapping errors, thereby demonstrating the promise of these closure models over ROMs without closure models.

The temporal evolution of relative errors in C_L and C_D for ROMs with different closure models is shown in Fig. 6 and Fig. 7 respectively. The errors in C_L and C_D exhibit a similar trend to the error in energy. For four mode ROMs, closure models TV-ME and TV-C provide low errors in C_L and C_D compared to other closure models and ROMs without closure models. For six mode ROMs, MV-ME and MV-C closure models exhibit better performance with the lowest error in time, which is even smaller than the error for ROMs with TV-ME and TV-C closure models. The error in these aerodynamic coefficients remains high when GV-GE and GV-C closure models are used despite the slightly lower error in energy approximation. These results indicate that ROM without a closure model is insufficient for dynamics prediction as the error rapidly increases outside the interval for which the data is used to obtain the energetic spatial basis. However, this drawback of Galerkin ROMs can be rectified by selecting an appropriate closure model. As the error does not grow significantly in time when MV-ME, MV-C, TV-ME and TV-C closure models are used, Galerkin ROM with these closure models appear to be well suited for time dynamics prediction for this Reynolds number.

A more complete comparison of the model requires a comparison of these errors for other selections of the number of modes. Some efficient metrics for comparing the model performance of the different number of modes are the relative L_2-integrated error in energy (η_E), relative L_2-integrated error in C_L (η_{C_L}) and relative L_2-integrated error in C_D (η_{C_D}) which are defined as follows:

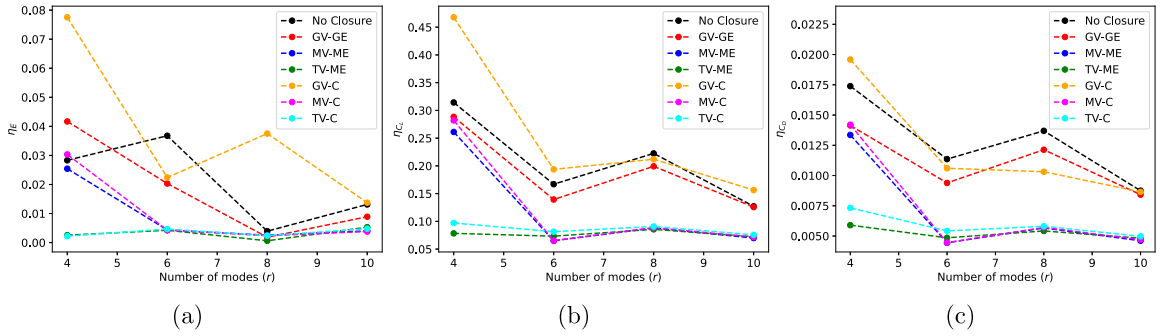$$\eta_E = \frac{||E(\boldsymbol{u}^{ROM}, \cdot) - E(\boldsymbol{u}^{FOM}, \cdot)||_2}{||E(\boldsymbol{u}^{FOM}, \cdot)||_2}, \tag{83}$$

$$\eta_{C_L} = \frac{||C_L(\boldsymbol{u}^{ROM}, p^{ROM}, \cdot) - C_L(\boldsymbol{u}^{FOM}, p^{FOM}, \cdot)||_2}{||C_L(\boldsymbol{u}^{FOM}, p^{FOM}, \cdot)||_2} \tag{84}$$
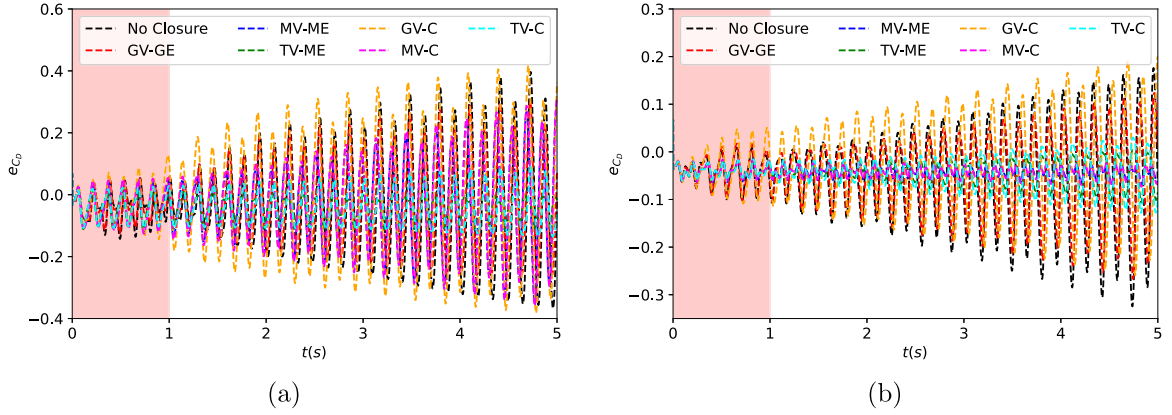
Fig. 7. Cylinder flow at Re = 200: Relative error in C_D (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.

Fig. 8. Cylinder flow at Re = 200: Relative L_2-integrated errors in (a) energy, (b) C_L and (c) C_D.

and

$$\eta_{C_D} = \frac{||C_D(\boldsymbol{u}^{ROM}, p^{ROM}, \cdot) - C_D(\boldsymbol{u}^{FOM}, p^{FOM}, \cdot)||_2}{||C_D(\boldsymbol{u}^{FOM}, p^{FOM}, \cdot)||_2}. \tag{85}$$

The variation of the L_2-integrated error in energy, C_L and C_D with the number of modes for several closure models is shown in Fig. 8. A non-monotonic change in the error is observed with increased modes for ROMs with and without closure models. The results indicate a higher error for ROMs with the GV-C closure model for several modes than when ROMs without a closure model are used, indicating that this closure model does not perform well. For a lower number of modes, ROMs with GV-GE, MV-ME and MV-C closure models appear to have a higher error in energy but lower errors in C_L and C_D compared to ROMs without a closure model. These results indicate that these models perform similarly to the ROMs without closure model for fewer modes. This behavior is drastically different when a higher number of modes are used. The results indicate that ROMs with GV-GE, MV-ME, and MV-C closure models give lower errors in energy, C_L, and C_D for a higher number of modes. The error for ROMs with the GV-GE closure model is slightly lower than those without a closure model. In contrast, ROMs with MV-ME and MV-C closure models give similar results with significantly low errors in C_L and C_D compared to ROMs without closure models. At a higher number of modes, ROMs with MV-ME and MV-C closure models give similar predictions to those from ROMs with TV-ME and TV-C closure models. The results also indicate that ROMs with MV-ME and MV-C closure models do not yield the best results at fewer modes. ROMs with TV-ME and TV-C closure models give similar results and the lowest errors for all modes. The predictions using these models exhibit a much lower error in energy, C_L and C_D compared to ROMs without closure models. These results indicate that ROMs with TV-ME and TV-C closure models consistently yield the best results for this Reynolds number.

5.4. Cylinder flow at Re = 500

The temporal evolution of relative errors in energy for four and six mode ROMs with different closure models is shown in Fig. 9. As the error in energy for the four mode ROM without a closure model is surprisingly low, ROMs with GV-GE, MV-ME, GV-C and MV-C closure models give comparatively worse errors. This lower energy error for the four mode ROM without a closure model is an anomaly, as we will observe by assessing other mode ROMs and looking at different quantities of interest. For the four mode scenario, the ROM with the TV-ME closure model gives results similar to the ROM without a closure model. In contrast, the ROM
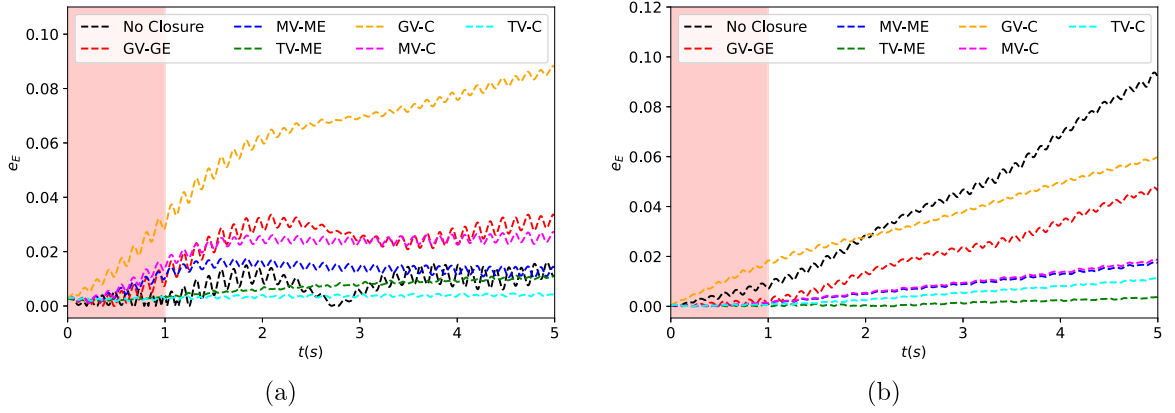
Fig. 9. Cylinder flow at Re = 500: Relative errors in energy for (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.
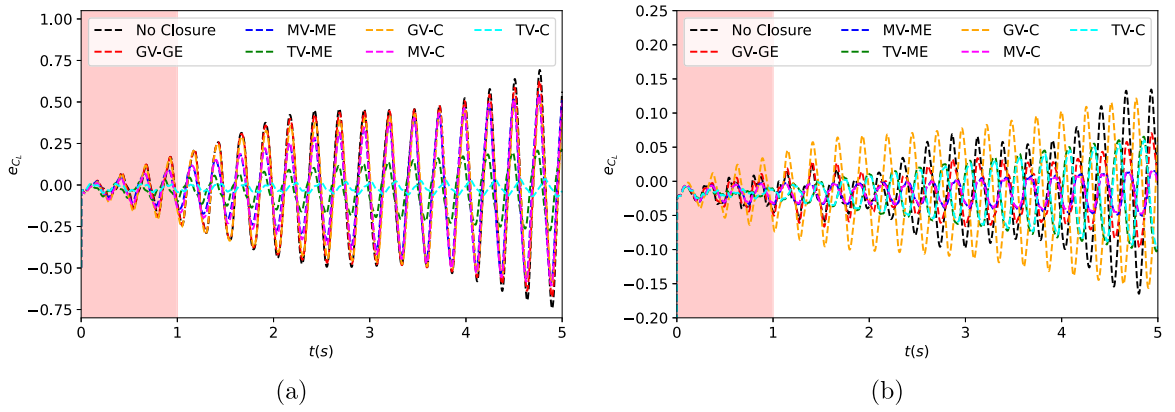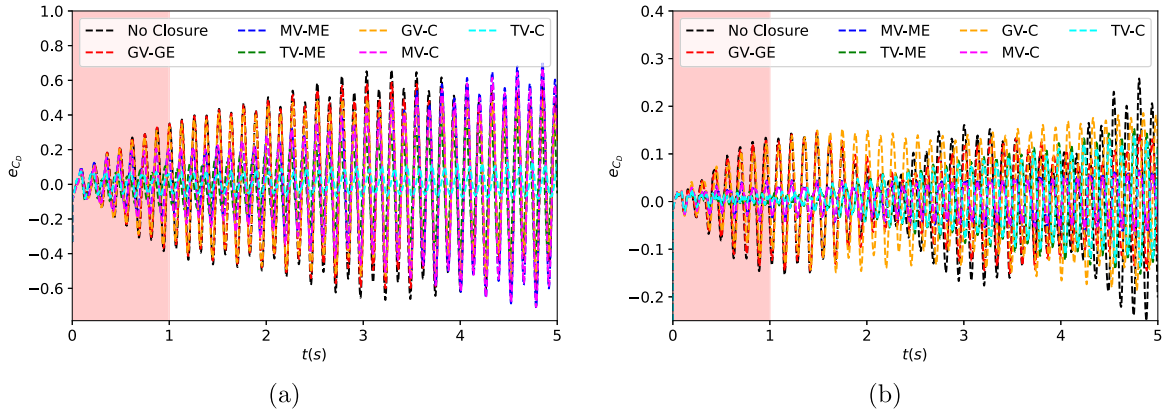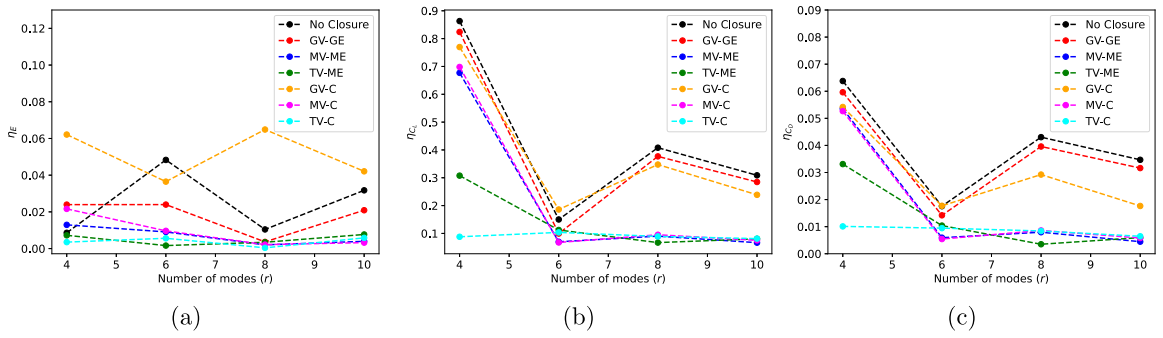

Fig. 10. Cylinder flow at Re = 500: Relative errors in C_L for (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.

with the TV-C closure model gives lower errors in energy than ROMs with other models and without a closure model. For the case of six mode ROMs, the ROM without a closure model exhibits a high error that rapidly rises outside the data generation window. ROMs with GV-GE and GV-C closure models display lower errors than those without a closure model. However, this error increases rapidly outside the data generation window. ROMs with MV-ME and MV-C closure models exhibit very low and similar errors. Lastly, ROMs with TV-ME and TV-C closure models exhibit even lower errors, with the former ROM displaying the lowest error among all ROMs.

The temporal evolution of relative errors in C_L and C_D for 4 and 6 mode ROMs with different closure models is shown in Fig. 10 and Fig. 11 respectively. When GV-GE, GV-C, MV-ME and MV-C closure models are used with ROMs consisting of 4 modes, errors in C_L and C_D are very similar to ROMs without closure models, especially outside the initial data generation window. On the other hand, ROMs with TV-ME and TV-C closure models exhibit a much lower error for these aerodynamic coefficients, with the latter model providing the lowest errors. The results for six mode ROMs indicate that ROMs with MV-ME and MV-C closure models give the lowest error in C_L and C_D for $t > 2s$. ROMs with GV-GE, TV-ME, and TV-C closure models give much lower errors than those without closure models, but this error is slightly higher than those for ROMs with MV-ME and MV-C closure models.
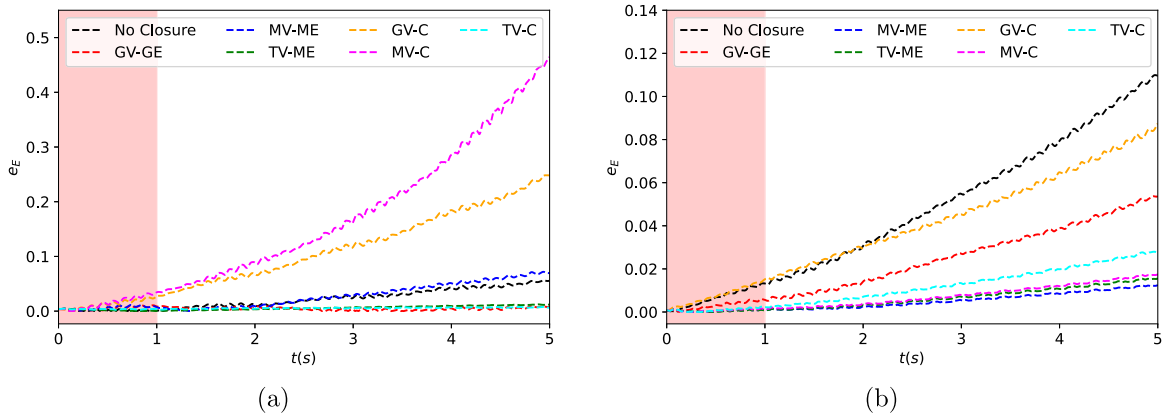
The L_2-integrated error in energy, C_L and C_D with different closure models is shown in Fig. 12. The error in energy for ROMs with the GV-C closure model is mostly higher than the ROMs without closure models, indicating this closure model downgrades the performance of ROM compared to the baseline ROMs without any closure model. ROMs with the GV-C closure model exhibit a slightly lower error in C_L and C_D than ROMs without a closure model. ROMs with the GV-GE closure model give lower errors in C_L, C_D and energy than ROMs without the closure model for all modes except when six mode ROMs are used. As highlighted earlier, the lower error for four mode ROM without a closure model is not expected and is peculiar for this test case. ROMs with MV-ME and MV-C closure models yield high errors for four mode ROMs. However, these errors decrease significantly for six to 10 mode ROMs. ROMs with TV-ME and TV-C closure models consistently give the lowest errors for all mode ROMs considered in this study.

Fig. 11. Cylinder flow at Re = 500: Relative errors in C_D for (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.


**Fig. 12.** Cylinder flow at Re = 500: Relative L_2-integrated errors in (a) energy, (b) C_L and (c) C_D.

Fig. 13. Cylinder flow at Re = 1000: Relative errors in energy for (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.

5.5. Cylinder flow at Re = 1000

The temporal evolution of relative errors in energy for four and six mode ROMs with different closure models is shown in Fig. 13. For four mode ROMs, the error in energy for ROMs with GV-C, MV-ME and MV-C closure models is higher than the ROM without any closure model. On the other hand, ROMs with GV-GE, TV-ME and TV-C closure models yield much lower errors in energy. For six mode ROMs, ROMs with any closure model deliver lower errors compared to the ROM without a closure model. ROMs with MV-ME, MV-C and TV-ME yield the lowest errors in energy that do not rise considerably even outside the initial time used for computing ROM operators. The errors for the ROM with the TV-C closure model increase after $t > 1$s. However, these errors are

18
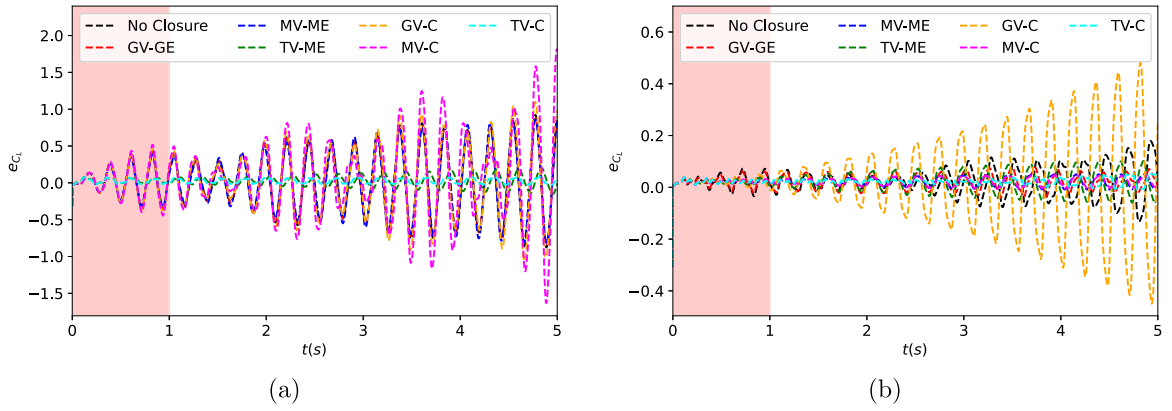
Fig. 14. Cylinder flow at Re = 1000: Relative errors in C_L for (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.
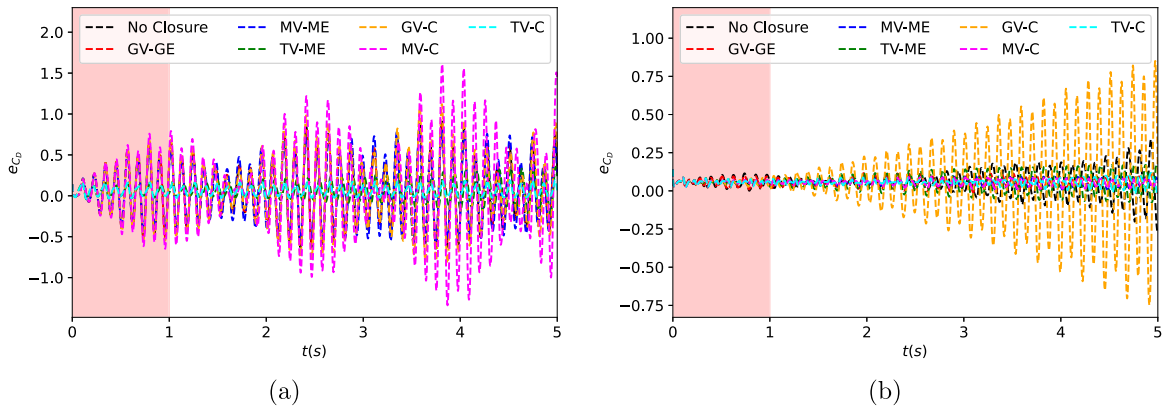
Fig. 15. Cylinder flow at Re = 1000: Relative errors in C_D for (a) four modes and (b) six modes. The shaded area is the time interval from which data is extracted to determine the POD basis, equation operators and closure models.

still very low, even for the prediction time window. On the other hand, the errors for ROMs with GV-GE and GV-C closure models rise considerably, with the latter giving much higher errors.

The temporal evolution of relative errors in C_L and C_D for four and six mode ROMs with different closure models is shown in Fig. 14 and Fig. 15 respectively. For four mode ROMs, the error in C_L and C_D is much lower when TV-ME and TV-C closure models are used than ROMs with other closure models and without a closure model. Similarly, for six mode ROMs, ROMs with MV-ME, MV-C and TV-C closure models yield the lowest errors, whereas ROMs with GV-GE and TV-ME closure models exhibit slightly higher errors. The errors for ROMs with these closure models are still much lower than those without a closure model, indicating the feasibility of most closure models, except for the ROM with the GV-C closure model.

The L_2-integrated error in energy, C_L and C_D with different closure models is shown in Fig. 16. Similar to the results for the other two Reynolds numbers, the ROMs with the GV-C closure model exhibit higher errors than ROMs without closure models for all mode ROMs considered in this article. ROMs with the GV-GE closure model show a significantly lower energy error than ROMs without a closure model. However, the differences in errors in aerodynamic coefficients between ROMs with these two closure models are minor. ROMs with TV-ME and TV-C closure models yield the lowest errors in all quantities for all modes, where the latter gives slightly better C_L and C_D for four modes and the former yields more accurate energy for six and eight mode ROMs. ROMs with MV-ME and MV-C closure models also yield similar low error at a higher number of modes, that is, for six to 10 mode ROMs. However, this error rises sharply for four mode ROMs. These results indicate that ROMs with TV-ME and TV-C closure models yield consistently low errors for all modes; therefore, these are the preferred closure models.

5.6. Discussion

The following conclusions can be drawn from the results obtained from the flow over a 2-D cylinder at three Reynolds numbers:
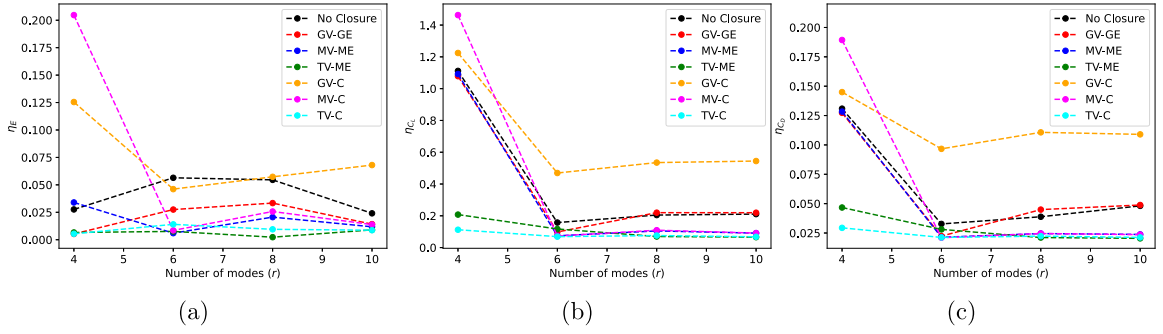
Fig. 16. Cylinder flow at Re = 1000: Relative L_2-integrated errors in (a) energy, (b) C_L and (c) C_D.

- The incremental pressure correction scheme provides accurate estimates of the several quantities of interest for different mode selections and gives significantly lower errors when coupled with an appropriate closure model.
- A constant artificial viscosity model form may not be expressive enough to resolve the closure problem and provide significantly lower errors than Galerkin ROMs without a closure model.
- The parameters of modal and tensor artificial viscosity closure models obtained using the two calibration techniques, based on energy and closure term approximation, do not lead to an appreciable difference in the closure model performance.
- Both modal and tensor artificial viscosity model forms are sufficient for a higher number of modes to account for the interaction of unresolved states on resolved states. However, at a lower number of modes, ROMs with tensor artificial viscosity models consistently yield better results than ROMs with modal artificial viscosity closure models.

These observations inform decision-making for selecting the appropriate solution method and closure models for ROMs for incompressible flows. The inapplicability of global artificial viscosity closure models is demonstrated through results indicating that GV-GE and GV-C closure models do not consistently yield lower errors than ROMs without a closure model. This conclusion resonates with the observation in [18], where a modal artificial viscosity was used to overcome this drawback. A comparison of the two calibration techniques indicated that energy transfer between unresolved and resolved states is sufficient to account for the non-linear dynamics, even for the highest Reynolds numbers under consideration. This result echoes the observation in [16] where the mean modal artificial viscosities computed using two calibration techniques are similar. Lastly, the modal artificial viscosity closure form is insufficient to capture the dynamics between the more energetic resolved and unresolved states, which appear to be better handled by a tensor artificial viscosity closure form. These conclusions guide the selection of appropriate model forms for even higher Reynolds number flows where the closure problem becomes even more prominent and ROMs with more modes become less favorable due to higher evaluation costs during the online stage.

6. Conclusions

ROMs are essential for multi-query applications, real-time control and dynamics forecasting. These models must achieve high accuracy while exhibiting a lower online computation cost. The cost of the online stage scales as r^3 where r is the number of resolved states for incompressible fluid flows. Therefore, it is desirable to use a lower number of resolved states. However, with a decrease in the number of resolved states, the need for accounting for the interaction between unresolved and resolved states becomes more important. The first contribution of this article is to propose closure models for the interaction between the unresolved and resolved states. We consider six different closure models determined using three model forms, based on global, modal and tensor artificial viscosity, where the unknown coefficients are determined by two calibration techniques: least squares minimization of error in energy approximation and closure term approximation. These closure models can be considered hybrid data-physics models as physical arguments inspire the model form, but the model parameters are directly learned through raw data. The proposed models did not require regularization for the cases considered in this study, unlike other common data-driven closure models [25,31]. The flow over a 2-D cylinder is considered at three Reynolds numbers and errors in energy, C_L and C_D are compared to validate these closure models. The results indicate that modal and tensor artificial viscosity model forms yield more accurate ROMs than those without any closure model. ROMs with tensor artificial viscosity model form deliver even more consistent results as the results did not deteriorate even for four mode ROMs, indicating that this model form is the most consistent amongst those considered in this article. Lastly, the two calibration techniques used for determining the unknown parameters in the closure model forms, least squares minimization for energy error or closure term error, provide closure models with slight differences in results. This observation indicates that both calibration techniques are equally well suited for determining modal and tensor artificial viscosity closure model parameters.

ROMs for incompressible fluid flows exhibit coupling between the reduced pressure states and reduced velocity states, where the primary role of the reduced pressure states is to enforce the incompressibility constraint. The unique structure of these equations results in a saddle-point problem that requires special solution techniques to obtain the dynamic evolution of velocity and pressure states. The second contribution of this article is to demonstrate the applicability of a solution technique: incremental pressure

correction for projection-based ROMs. This method is adapted from the incremental pressure correction scheme readily used in full order modeling of incompressible fluid flows [51]. The ROMs obtained through this solution method give low errors for incompressible flows under consideration, thereby demonstrating the applicability of this method for ROM of incompressible fluid flows.

The solution method and the closure models proposed in this article require further validation on complex three-dimensional turbulent flows to ensure the accuracy and efficiency remains. The advancements proposed in this article are directly applicable to such flows but will require additional efforts to ensure a scalable and efficient offline stage for ROMs of such large-scale flows. Furthermore, a detailed numerical analysis of the stability and convergence of the incremental pressure correction scheme in the context of ROMs will be performed. Alternative solution methods for saddle-point problems, such as velocity correction schemes, exist in the full order modeling literature. The applicability of schemes for ROMs must be validated. The closure models considered in this article are linear as they depend linearly on reduced states. These models could not adequately account for the interaction between resolved and unresolved states when less than four modes were used for ROMs. Non-linear model forms have the potential to be more accurate for such a lower number of modes and will be explored in future work. Despite the potential to be more accurate than linear models, these models would involve a higher online stage cost. Therefore, a comprehensive study is required to investigate the trade-off between cost and accuracy. As part of future work, we also plan to explore novel applications of ROMs in material transport regulation of neurite networks [64–67] and heat exchanger design for additive manufacturing [68,69] where complex and challenging fluid flows need to be addressed.

CRediT authorship contribution statement

Aviral Prakash: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yongjie Jessica Zhang:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

References

[1] L. Sirovich, Turbulence and the dynamics of coherent structures part III: Dynamics and scaling, Quart. Appl. Math. 45 (3) (1987) 583–590.

[2] N. Aubry, P. Holmes, J.L. Lumley, E. Stone, The dynamics of coherent structures in the wall region of a turbulent boundary layer, J. Fluid Mech. 192 (1988) 115–173.

[3] J.L. Lumley, The structure of inhomogeneous turbulent flows, in: Atmospheric Turbulence and Radio Wave Propagation, Nauka, Moscow, 1967, pp. 166–178.

[4] J.L. Lumley, Coherent structures in turbulence, in: Richard E. Meyer (Ed.), Transition and Turbulence, Academic Press, 1981, pp. 215–242.

[5] J.S. Hesthaven, G. Rozza, B. Stamm, Certified Reduced Basis Methods for Parametrized Partial Differential Equations, Springer, 2016.

[6] S. Sirisup, G.E. Karniadakis, A spectral viscosity method for correcting the long-term behavior of POD models, J. Comput. Phys. 194 (1) (2004) 92–116.

[7] Z. Wang, I. Akhtar, J. Borggaard, T. Iliescu, Proper orthogonal decomposition closure models for turbulent flows: A numerical comparison, Computer Methods in Applied Mechanics and Engineering 237-240 (2012) 10–26.

[8] M. Bergmann, C.-H. Bruneau, A. Iollo, Enablers for robust POD models, J. Comput. Phys. 228 (2) (2009) 516–538.

[9] K. Carlberg, C. Bou-Mosleh, C. Farhat, Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations, Internat. J. Numer. Methods Engrg. 86 (2) (2011) 155–181.

[10] S.E. Ahmed, S. Pawar, O. San, A. Rasheed, T. Iliescu, B.R. Noack, On closures for reduced order models - A spectrum of first-principle to machine-learned avenues, Phys. Fluids 33 (9) (2021).

[11] K. Carlberg, M. Barone, H. Antil, Galerkin v. least-squares Petrov–Galerkin projection in nonlinear model reduction, J. Comput. Phys. 330 (2017) 693–734.

[12] E.J. Parish, C.R. Wentland, K. Duraisamy, The adjoint Petrov–Galerkin method for non-linear model reduction, Comput. Methods Appl. Mech. Engrg. 365 (2020) 112991.

[13] E. Parish, M. Yano, I. Tezaur, T. Iliescu, Residual-based stabilized reduced-order models of the transient convection–diffusion-reaction equation obtained through discrete and continuous projection, 2023, arXiv:2302.09355.

[14] T. Iliescu, Z. Wang, Variational multiscale proper orthogonal decomposition: Navier–Stokes equations, Numer. Methods Partial Differential Equations 30 (2) (2014) 641–663.

[15] D. Rempfer, H.F. Fasel, Dynamics of three-dimensional coherent structures in a flat-plate boundary layer, J. Fluid Mech. 275 (1994) 257–283.

[16] M. Couplet, P. Sagaut, C. Basdevant, Intermodal energy transfers in a proper orthogonal decomposition – Galerkin representation of a turbulent separated flow, J. Fluid Mech. 491 (2003) 275–284.

[17] J. Borggaard, T. Iliescu, Z. Wang, Artificial viscosity proper orthogonal decomposition, Math. Comput. Modelling 53 (1) (2011) 269–279.

[18] J. Östh, B.R. Noack, S. Krajnović, D. Barros, J. Borée, On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high-Reynolds-number flow over an Ahmed body, J. Fluid Mech. 747 (2014) 518–544.

[19] X. Xie, D. Wells, Z. Wang, T. Iliescu, Approximate deconvolution reduced order modeling, Comput. Methods Appl. Mech. Engrg. 313 (2017) 512–534.

[20] R. Reyes, R. Codina, Projection-based reduced order models for flow problems: A variational multiscale approach, Comput. Methods Appl. Mech. Engrg. 363 (2020) 112844.

[21] X. Xie, D. Wells, Z. Wang, T. Iliescu, Numerical analysis of the leray reduced order model, J. Comput. Appl. Math. 328 (2018) 12–29.

[22] L. Cordier, B.R. Noack, G. Tissot, G. Lehnasch, J. Delville, M. Balajewicz, G. Daviller, R.K. Niven, Identification strategies for model-based control topics in flow control, Exp. Fluids 54 (8) (2013).

[23] B. Protas, B.R. Noack, J. Östh, Optimal nonlinear eddy viscosity in Galerkin models of turbulent flows, J. Fluid Mech. 766 (2015) 337–367.

[24] S.E. Ahmed, K. Bhar, O. San, A. Rasheed, Forward sensitivity approach for estimating eddy viscosity closures in nonlinear model reduction, Phys. Rev. E 102 (2020).

[25] X. Xie, M. Mohebujjaman, L.G. Rebholz, T. Iliescu, Data-driven filtered reduced order modeling of fluid flows, SIAM J. Sci. Comput. 40 (3) (2018) B834–B857.

[26] K. Duraisamy, G. Iaccarino, H. Xiao, Turbulence modeling in the age of data, Annu. Rev. Fluid Mech. 51 (1) (2019) 357–377.

[27] A. Prakash, K.E. Jansen, J.A. Evans, Invariant data-driven subgrid stress modeling in the strain-rate eigenframe for large eddy simulation, Comput. Methods Appl. Mech. Engrg. 399 (2022) 115457.

[28] A. Prakash, K.E. Jansen, J.A. Evans, Invariant data-driven subgrid stress modeling on anisotropic grids for large eddy simulation, Comput. Methods Appl. Mech. Engrg. 422 (2024) 116807.

[29] A. Prakash, Scale-Resolving Simulations and Data-Driven Subgrid Modeling for Complex Turbulent Boundary Layer Flows (Ph.D. thesis), University of Colorado Boulder, 2023.

[30] B. Koc, C. Mou, H. Liu, Z. Wang, G. Rozza, T. Iliescu, Verifiability of the data-driven variational multiscale reduced order model, J. Sci. Comput. 93 (2) (2022) 54.

[31] M. Mohebujjaman, L.G. Rebholz, T. Iliescu, Physically constrained data-driven correction for reduced-order modeling of fluid flows, Internat. J. Numer. Methods Fluids 89 (3) (2019) 103–122.

[32] O. San, R. Maulik, Neural network closures for nonlinear model order reduction, Adv. Comput. Math. 44 (6) (2018) 1717–1750.

[33] L. Cordier, B. Abou El Majd, J. Favier, Calibration of POD reduced-order models using Tikhonov regularization, Internat. J. Numer. Methods Fluids 63 (2) (2010) 269–296.

[34] B.R. Noack, P. Papas, P.A. Monkewitz, The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows, J. Fluid Mech. 523 (2005) 339–365.

[35] A. Caiazzo, T. Iliescu, V. John, S. Schyschlowa, A numerical investigation of velocity–pressure reduced order models for incompressible flows, J. Comput. Phys. 259 (2014) 598–616.

[36] F. Ballarin, A. Manzoni, A. Quarteroni, G. Rozza, Supremizer stabilization of POD–Galerkin approximation of parametrized steady incompressible Navier–Stokes equations, Internat. J. Numer. Methods Engrg. 102 (5) (2015) 1136–1161.

[37] V. DeCaria, T. Iliescu, W. Layton, M. McLaughlin, M. Schneier, An artificial compression reduced order model, SIAM J. Numer. Anal. 58 (1) (2020) 565–589.

[38] S. Rubino, Numerical analysis of a projection-based stabilized POD-ROM for incompressible flows, SIAM J. Numer. Anal. 58 (4) (2020) 2019–2058.

[39] J.L. Guermond, P. Minev, J. Shen, An overview of projection methods for incompressible flows, Comput. Methods Appl. Mech. Engrg. 195 (44) (2006) 6011–6045.

[40] A. Prakash, K.E. Jansen, J.A. Evans, Extension of the Smagorinsky subgrid stress model to anisotropic filters, in: AIAA SCITECH 2023 Forum, 2023.

[41] A. Abbà, L. Valdettaro, Tensorial turbulent viscosity model for LES: Properties and applications, in: From Kinetic Theory to Turbulence Modeling, Springer Nature Singapore, Singapore, 2023, pp. 9–20.

[42] I. Babuška, The finite element method with Lagrangian multipliers, Numer. Math. 20 (3) (1973) 179–192.

[43] F. Brezzi, On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers, R.A.I.R.O. Anal. Numér. 8 (1974) 129–151.

[44] C.W. Rowley, T. Colonius, R.M. Murray, Model reduction for compressible flows using POD and Galerkin projection, Physica D 189 (1) (2004) 115–129.

[45] I. Kalashnikova, M.F. Barone, On the stability and convergence of a Galerkin reduced order model (ROM) of compressible flow with solid wall and far-field boundary treatment, Internat. J. Numer. Methods Engrg. 83 (10) (2010) 1345–1375.

[46] C. Greif, K. Urban, Decay of the Kolmogorov N-width for wave problems, Appl. Math. Lett. 96 (2019) 216–222.

[47] Giovanni Stabile, Bojana Rosic, Bayesian identification of a projection-based reduced order model for computational fluid dynamics, Comput. Fluids 201 (2020) 104477.

[48] T.J.R. Hughes, L.P. Franca, M. Balestra, A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška-Brezzi condition: A stable Petrov–Galerkin formulation of the Stokes problem accommodating equal-order interpolations, Comput. Methods Appl. Mech. Eng. 59 (1) (1986) 85–99.

[49] A.J. Chorin, Numerical solution of the Navier–Stokes equations, Math. Comp. 22 (104) (1968) 745–762.

[50] R. Témam, Sur l'approximation de la solution des équations de Navier–Stokes par la méthode des pas fractionnaires (I), Arch. Ration. Mech. Anal. 32 (2) (1969) 135–153.

[51] J.V. Kan, A second-order accurate pressure-correction scheme for viscous incompressible flow, SIAM J. Sci. Stat. Comput. 7 (3) (1986) 870–891.

[52] C. Taylor, P. Hood, A numerical solution of the Navier–Stokes equations using the finite element technique, Comput. & Fluids 1 (1) (1973) 73–100.

[53] G. Rozza, K. Veroy, On the stability of the reduced basis method for Stokes equations in parametrized domains, Comput. Methods Appl. Mech. Engrg. 196 (7) (2007) 1244–1260.

[54] J.L. Guermond, L. Quartapelle, On stability and convergence of projection methods based on pressure Poisson equation, Internat. J. Numer. Methods Fluids 26 (9) (1998) 1039–1053.

[55] R. Rannacher, On Chorin's projection method for the incompressible Navier–Stokes equations, in: The Navier–Stokes Equations II — Theory and Numerical Methods, Springer Berlin Heidelberg, 1992, pp. 167–183.

[56] I. Akhtar, A.H. Nayfeh, C.J. Ribbens, On the stability and extension of reduced-order Galerkin models in incompressible flows, Theor. Comput. Fluid Dyn. 23 (3) (2009) 213–237.

[57] M. Schäfer, S. Turek, F. Durst, E. Krause, R. Rannacher, Benchmark computations of laminar flow around a cylinder, in: Flow Simulation with High-Performance Computers II: DFG Priority Research Programme Results 1993–1995, Vieweg+Teubner Verlag, 1996, pp. 547–566.

[58] I.A. Baratta, J.P. Dean, J.S. Dokken, M. Habera, J.S. Hale, C.N. Richardson, M.E. Rognes, M.W. Scroggs, N. Sime, G.N. Wells, DOLFINx: The next generation FEniCS problem solving environment, 2023.

[59] H.P. Langtangen, A. Logg, Solving PDEs in Python: The FEniCS Tutorial I, Springer, 2016.

[60] J.S. Dokken, The FEniCSx tutorial, 2024, https://jsdokken.com/dolfinx-tutorial/. [Accessed: 2024-01-07].

[61] I. Kalashnikova, M.F. Barone, S. Arunajatesan, B.G.V.B. Waanders, Construction of energy-stable projection-based reduced order models, Appl. Math. Comput. 249 (2014) 569–596.

[62] S. Ingimarson, L.G. Rebholz, T. Iliescu, Full and reduced order model consistency of the nonlinearity discretization in incompressible flows, Comput. Methods Appl. Mech. Engrg. 401 (2022) 115620.

[63] L.D. Dalcin, R.R. Paz, P.A. Kler, A. Cosimo, Parallel distributed computing using Python, Adv. Water Resour. 34 (9) (2011) 1124–1139.

[64] A. Li, A.B. Farimani, Y.J. Zhang, Deep learning of material transport in complex neurite networks, Sci. Rep. 11 (1) (2021) 11280.

[65] A. Li, Y.J. Zhang, Modeling material transport regulation and traffic jam in neurons using PDE-constrained optimization, Sci. Rep. 12 (1) (2022) 3902.

[66] A. Li, Y.J. Zhang, Modeling intracellular transport and traffic jam in 3D neurons using PDE-constrained optimization, J. Mech. 38 (2022) 44–59.

[67] A. Li, Y.J. Zhang, Isogeometric analysis-based physics-informed graph neural network for studying traffic jam in neurons, Comput. Methods Appl. Mech. Engrg. 403 (2023) 115757.

[68] X. Liang, A. Li, A.D. Rollett, Y.J. Zhang, An isogeometric analysis-based topology optimization framework for 2D cross-flow heat exchangers with manufacturability constraints, Eng. Comput. 38 (6) (2022) 4829–4852.

[69] X. Liang, L. White, J. Cagan, A.D. Rollett, Y.J. Zhang, Unit-based design of cross-flow heat exchangers for LPBF additive manufacturing, J. Mech. Des. 145 (1) (2022) 012002.