HAL open science

Verifiability of the Data-Driven Variational Multiscale Reduced Order Model

Birgul Koc, Changhong Mou, Honghu Liu, Zhu Wang, Gianluigi Rozza, Traian Iliescu

To cite this version:

Birgul Koc, Changhong Mou, Honghu Liu, Zhu Wang, Gianluigi Rozza, et al.. Verifiability of the Data-Driven Variational Multiscale Reduced Order Model. Journal of Scientific Computing, 2022, 93 (2), pp.54. 10.1007/s10915-022-02019-y . hal-03854187

HAL Id: hal-03854187

https://ifp.hal.science/hal-03854187v1

Submitted on 15 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

VERIFIABILITY OF THE DATA-DRIVEN VARIATIONAL MULTISCALE REDUCED ORDER MODEL

BIRGUL KOC, CHANGHONG MOU, HONGHU LIU, ZHU WANG, GIANLUIGI ROZZA, AND TRAIAN ILIESCU

Abstract. In this paper, we focus on the mathematical foundations of reduced order model (ROM) closures. First, we extend the verifiability concept from large eddy simulation to the ROM setting. Specifically, we call a ROM closure model verifiable if a small ROM closure model error (i.e., a small difference between the true ROM closure and the modeled ROM closure) implies a small ROM error. Second, we prove that the data-driven ROM closure studied here (i.e., the data-driven variational multiscale ROM) is verifiable. Finally, we investigate the verifiability of the data-driven variational multiscale ROM in the numerical simulation of the one-dimensional Burgers equation and a two-dimensional flow past a circular cylinder at Reynolds numbers $Re = 100$ and $Re = 1000$. Reduced order model and variational multiscale and data-driven model and verifiability

1. Introduction

Full order models (FOMs) are computational models obtained with classical numerical methods (e.g., finite element or finite difference methods). In the numerical simulation of fluid flows, FOMs often yield high-dimensional (e.g., $\mathcal{O}(10^6)$) systems of equations. Thus, the computational cost of using FOMs in important many-query fluid flow applications (e.g., uncertainty quantification, optimal control, and shape optimization) can be prohibitively high.

Reduced order models (ROMs) are computational models that yield systems of equations whose dimensions are dramatically lower than those corresponding to FOMs. For example, in the numerical simulation of fluid flows that are dominated by recurrent spatial structures (e.g., flow past bluff bodies), the dimensions of the resulting system of equations can be $\mathcal{O}(10)$ for ROMs and $\mathcal{O}(10^6)$ for FOMs, while the ROM and FOM accuracy is of the same order. Thus, ROMs have been used in many-query fluid flow applications to reduce the computational cost of FOMs. Probably the most popular type of ROM used in these applications is the Galerkin ROM (G-ROM), which is constructed by using the Galerkin method. The G-ROM is based on a simple yet powerful idea: Instead of using millions or even billions of general purpose basis functions (as in classical Galerkin methods, such as the tent functions in the finite element method), G-ROM uses a lower-dimensional data-driven basis. Specifically, the available numerical or experimental data is used to build a few ROM basis functions that model the spatial structures that dominate the flow dynamics.

The G-ROM has been successful in the efficient numerical simulation of relatively simple laminar flows, e.g., flow past a circular cylinder at low Reynolds numbers. However, the standard G-ROM generally fails in the numerical simulation of turbulent flows. The main reason is that, in order to ensure a relatively low computational cost, only a few ROM basis functions are used to build the standard G-ROM. These few ROM basis functions can represent the simple dynamics of laminar flows, but not the complex dynamics of turbulent flows. Thus, in the numerical simulation of turbulent flows, the standard G-ROM is equipped with a ROM closure model, i.e., a correction term that models the effect of the discarded ROM basis functions on the ROM dynamics.

(BK) Departamento EDAN & IMUS, Universidad de Sevilla, Campus de Reina Mercedes, 41012, Sevilla, Spain
(CM) Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA
(HL) Department of Mathematics, Virginia Tech, Blacksburg, VA 24061, USA
(ZW) Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA
(GR) mathLab, Mathematics Area, SISSA, I-34136 Trieste, Italy
(TI) Department of Mathematics, Virginia Tech, Blacksburg, VA 24061, USA
E-mail addresses: (BK) birkoc@alum.us.es , (CM) cmou3@wisc.edu , (HL) hhliu@vt.edu, (ZW) wangzhu@math.sc.edu, (GR) gianluigi.rozza@sissa.it, (TI) iliescu@vt.edu.

Over the last two decades, ROM closure modeling has witnessed a dynamic development. A survey of current ROM closure modeling strategies is presented in [2]. Three main types of ROM closure models have been proposed: (i) Functional ROM closures are constructed by using physical insight. Classical examples of functional ROM closures include eddy viscosity models [53], in which the main role of the ROM closure model is to dissipate energy. (ii) Structural ROM closures are a different class of models that are developed by using mathematical arguments. Examples of structural ROM closures include the approximate deconvolution ROM [56], the Mori-Zwanzig formalism [15, 34, 40], and the parameterizing manifolds [11, 12, 13]. (iii) The most active research area in ROM closure modeling is in the development of data-driven ROM closures in which available data is utilized to build the ROM closure model. An example of data-driven ROM closure is the data-driven variational multiscale ROM (DD-VMS-ROM) that was proposed in [37, 54]. The DD-VMS-ROM has been investigated numerically in [31, 36, 38, 37, 54, 55]. However, providing mathematical support for the DD-VMS-ROM is an open problem.

In classical CFD, there exists extensive mathematical support for closure modeling. For example, the monographs [9, 27, 45] present the mathematical analysis for many large eddy simulation (LES) models, as well as the numerical analysis of their discretization. In contrast, despite the recent increased interest in ROM closure modeling [2], the mathematical foundations of ROM closures are relatively scarce. Indeed, the ROM closure models are generally assessed heuristically: The proposed ROM closure model is used in numerical simulations and is shown to improve the numerical accuracy of the standard G-ROM and/or other ROM closure models. However, fundamental questions in ROM numerical analysis are still wide open for most of these ROM closure models: Is the proposed ROM closure model stable? Does the ROM closure model converge? If so, what does it converge to?

Only the first steps in the numerical analysis of ROM closures have been taken. To our knowledge, the first numerical analysis of a ROM closure model was performed in [10], where an eddy viscosity ROM closure model (i.e., the Smagorinsky model) was analyzed in a simplified setting. Next, the numerical analysis of eddy viscosity variational multiscale ROMs was carried out in [25, 26]. Finally, the numerical analysis of the Samagorinsky model in a reduced basis method (RBM) setting was performed in [7, 44]. We note that numerical analysis for regularized ROMs, which are related to but different from ROM closures, was performed in [19, 57]; see also [5] for related work.

In this paper, we take a next step in the development of numerical analysis for ROM closures and prove verifiability for a data-driven ROM closure model, i.e., the DD-VMS-ROM proposed in [37, 54]. Specifically, we show that the ROM closure model in the DD-VMS-ROM is accurate in a precise sense. More importantly, we prove that the DD-VMS-ROM is verifiable, i.e., we prove that since the DD-VMS-ROM closure model is accurate, the DD-VMS-ROM solution is accurate. We note that this is not a trivial task: The Navier-Stokes equations (and their filtered counterparts), which are the mathematical models that we use in this paper, are nonlinear and sensitive to perturbations, so adding to them a relatively small term (i.e., the ROM closure term) does not automatically imply that the resulting solution will be close to the original one. To prove that the DD-VMS-ROM closure model is verifiable, we use the following ingredients: (i) We use ROM spatial filtering to determine an explicit formula for the exact ROM closure term, which needs to be modeled. (ii) We use data-driven modeling to construct the DD-VMS-ROM closure model and show that this closure model is accurate, i.e., it is close to the exact ROM closure model. (iii) We use physical constraints to increase the accuracy of our data-driven ROM closure model. We note that the verifiability concept was defined in an LES context (see, e.g., [30] as well as [9] for a survey). However, to our knowledge, this is the first time the verifiability concept is defined and investigated in a ROM context.

The rest of the paper is organized as follows: In Section 2, we outline the construction of the standard G-ROM. In Sections 3 and 4, we use ROM spatial filtering to build LES-ROMs and utilize data-driven modeling to build the closure model in the DD-VMS-ROM, respectively. In Section 5, we prove the main theoretical result in this paper, i.e., we prove that the DD-VMS-ROM is verifiable. In Section 6, we illustrate the theoretical developments. Specifically, for the Burgers equation and the two-dimensional flow past a circular cylinder, we show the following: (i) the ROM closure error (i.e., the difference between the true ROM closure term and the DD-VMS-ROM closure term) is small and it becomes smaller and smaller as we increase the ROM dimension; and (ii) as the ROM closure error decreases, so does the ROM error (i.e., the DD-VMS-ROM is verifiable). Finally, in Section 7, we present the conclusions of our theoretical and numerical investigations and outline several directions for future research.

2. Galerkin ROM (G-ROM)

In this section, we outline the construction of the Galerkin ROM (G-ROM) for the Navier-Stokes equations (NSE):

$$\frac{\partial \boldsymbol{u}}{\partial t} - Re^{-1}\Delta \boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + \nabla p = \boldsymbol{f}, \tag{2.1}$$

$$\nabla\cdot\boldsymbol{u} = 0, \tag{2.2}$$

where $\boldsymbol{u}$ is the velocity, p the pressure, and Re the Reynolds number. The NSE (2.1)–(2.2) are equipped with an initial condition and, for simplicity, homogeneous Dirichlet boundary conditions. To build the ROM basis, we assume that we have access to the snapshots $\{\boldsymbol{u}_h^0, ..., \boldsymbol{u}_h^M\}$, which are the coefficient vectors of the FEM approximations of the NSE (2.1)–(2.2) at the time instances $t_0, t_1, \ldots, t_M$, respectively. The number of snapshots, M, is an arbitrary positive integer. In what follows, we assume that M is fixed. Next, we use these snapshots and the proper orthogonal decomposition (POD) [23, 52] to construct an orthonormal ROM basis $\{\boldsymbol{\varphi}_1, ..., \boldsymbol{\varphi}_d\}$, which generates the ROM space $\boldsymbol{X}^d$ defined as follows:

$$\boldsymbol{X}^d := \text{span}\{\boldsymbol{\varphi}_1, ..., \boldsymbol{\varphi}_d\}, \tag{2.3}$$

where d is the number of linearly independent snapshots $\{\boldsymbol{u}_h^0, ..., \boldsymbol{u}_h^M\}$. Thus, d is the maximal dimension of a basis that spans the same space as the space spanned by the given snapshots. By using the ROM basis functions in (2.3), we construct $\boldsymbol{u}_d$, which is the d-dimensional ROM approximation of NSE velocity, $\boldsymbol{u}$:

$$\boldsymbol{u}_d(\boldsymbol{x}, t) = \sum_{i=1}^{d} (\boldsymbol{a}_d)_i(t)\, \boldsymbol{\varphi}_i(\boldsymbol{x}). \tag{2.4}$$

To find the vector of ROM coefficients $\boldsymbol{a}_d$ in (2.4), we use the Galerkin projection, i.e., we replace $\boldsymbol{u}$ with $\boldsymbol{u}_d$ in the NSE (2.1)–(2.2), and then project the resulting equations onto the ROM space, $\boldsymbol{X}^d$. This yields the d-dimensional Galerkin ROM (G-ROM):

$$((\boldsymbol{u}_d)_t, \boldsymbol{v}_d) + Re^{-1}(\nabla\boldsymbol{u}_d, \nabla\boldsymbol{v}_d) + (\boldsymbol{u}_d\cdot\nabla\boldsymbol{u}_d, \boldsymbol{v}_d) = (\boldsymbol{f}, \boldsymbol{v}_d), \qquad \forall\, \boldsymbol{v}_d \in \boldsymbol{X}^d, \tag{2.5}$$

where $(\cdot,\cdot)$ denotes the L^2 inner product. We note that the G-ROM (2.5) does not include a pressure term, since the ROM basis functions are assumed to be discretely divergence-free. This is the case if, e.g., the snapshots are discretely divergence-free. Indeed, when POD is used to construct the ROM basis (as in our numerical investigation), the ROM basis functions are linear combinations of the snapshots. Since the snapshots are discretely divergence-free, so are the ROM basis functions. We also note that alternative formulations within the RBM framework are used in, e.g., [4, 6, 21, 22, 35, 42].

By using the backward Euler time discretization, we get the full discretization of the d-dimensional G-ROM (2.5) as follows: $\forall\, n = 1, ..., M$

$$\left(\frac{\boldsymbol{u}_d^n - \boldsymbol{u}_d^{n-1}}{\Delta t}, \boldsymbol{v}_d\right) + Re^{-1}(\nabla\boldsymbol{u}_d^n, \nabla\boldsymbol{v}_d) + (\boldsymbol{u}_d^n\cdot\nabla\boldsymbol{u}_d^n, \boldsymbol{v}_d) = (\boldsymbol{f}^n, \boldsymbol{v}_d), \quad \forall\, \boldsymbol{v}_d \in \boldsymbol{X}^d, \tag{2.6}$$

where the superscript n denotes the approximation at time step n. To obtain the finite-dimensional representation of the d-dimensional G-ROM (2.6), we choose $\boldsymbol{v}_d$ to be $\boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_d$, which yields the following system of equations:

$$\frac{\boldsymbol{a}_d^n - \boldsymbol{a}_d^{n-1}}{\Delta t} = \boldsymbol{b}^n + \boldsymbol{A}\,\boldsymbol{a}_d^n + (\boldsymbol{a}_d^n)^\top\,\boldsymbol{B}\,\boldsymbol{a}_d^n, \tag{2.7}$$

where $\boldsymbol{a}_d^n$ is the vector of unknown ROM coefficients, $\boldsymbol{b}$ is a $d\times 1$ vector, $\boldsymbol{A}$ is a $d\times d$ matrix, and $\boldsymbol{B}$ is a $d\times d\times d$ tensor. The system of equations in (2.7) can be written componentwise as follows:

$$\frac{(\boldsymbol{a}_d^n)_i - (\boldsymbol{a}_d^{n-1})_i}{\Delta t} = \boldsymbol{b}_i^n + \sum_{m=1}^{d} \boldsymbol{A}_{im}\boldsymbol{a}_m^n + \sum_{m=1}^{d}\sum_{k=1}^{d} \boldsymbol{B}_{imk}\boldsymbol{a}_m^n\boldsymbol{a}_k^n\,, \quad 1 \le i \le d\,, \tag{2.8}$$

where, for $1 \leq i, m, k \leq d$,

$$\boldsymbol{b}_i^n = (\boldsymbol{f}^n, \boldsymbol{\varphi}_i), \tag{2.9}$$

$$\boldsymbol{A}_{im} = -Re^{-1} \left(\nabla \boldsymbol{\varphi}_m, \nabla \boldsymbol{\varphi}_i\right), \tag{2.10}$$

$$\boldsymbol{B}_{imk} = -\big(\boldsymbol{\varphi}_m \cdot \nabla \boldsymbol{\varphi}_k, \boldsymbol{\varphi}_i\big). \tag{2.11}$$

3. Large Eddy Simulation ROM (LES-ROM)

The ROM closure that we investigate in this paper (i.e., the DD-VMS-ROM presented in Section 4) is a large eddy simulation ROM (LES-ROM). Thus, in this section, we briefly outline the construction of LES-ROMs.

LES-ROMs are ROM closures that have been developed over the last decade (see [53, 56] and the survey in Section V in [2], as well as related approaches in [17, 18]). LES-ROMs utilize mathematical principles used in classical LES [9, 47] to construct ROM closure models for ROMs in under-resolved regimes, i.e., when the number of ROM basis functions is insufficient to represent the complex dynamics of the underlying flows. Classical LES and LES-ROMs are similar in spirit: They both aim at approximating the large scales in the flow at the available coarse resolution (e.g., coarse mesh in classical LES and not enough ROM basis functions in LES-ROMs). Furthermore, they both use spatial filtering to define the large scales that need to be approximated. We emphasize, however, that there are also major differences between classical LES and LES-ROMs. One of the main differences is the type of spatial filtering used to define the large flow structures. In classical LES, continuous filters (e.g., the Gaussian filter) are used to define the filtered equations at a continuous level. In contrast, in LES-ROMs, due to the hierarchical structure of the ROM spaces, the ROM projection (which is a discrete spatial filter) is generally used instead. (For a notable exception, see the ROM differential filter, which is a continuous spatial ROM filter used in [56] to construct the approximate deconvolution ROM closure.) The ROM projection is used, in particular, to build variational multiscale (VMS) ROM closures (see, e.g., [8, 25, 26, 46, 49, 53] and the VMS-ROM survey in Section V.A in [2]), such as the closure that we investigate in this paper, which we describe next.

To construct the DD-VMS-ROM, we start by choosing the "truth" solution, i.e., the most accurate ROM solution that we can construct with the given snapshots.

Definition 3.1 (Truth Solution)**.** *For fixed M and d, we define the d-dimensional G-ROM solution of* (2.6) *as our "truth" solution.*

The goal of an LES-ROM is to construct an r-dimensional ROM whose solution, $\boldsymbol{u}_r$, approximates as accurately as possible the large scale component of the truth solution, $P_r(\boldsymbol{u}_d)$. We note that, since $r \ll d$, the LES-ROM development takes place in an under-resolved regime.

In what follows, our goal is to use data to construct an LES-ROM (specifically, the DD-VMS-ROM) whose solutions are as close as possible to $P_r(\boldsymbol{u}_d)$, i.e., the ROM projection of the truth solution. Thus, in the numerical analysis in Section 5, the DD-VMS-ROM solution will be compared to large scale component of the truth solution, which will be considered as *data*.

In what follows, we use the LES-ROM framework to achieve the following objectives: (i) Use the ROM projection to define the large ROM spatial scales; (ii) Use the ROM projection to filter the d-dimensional G-ROM (2.6) and obtain the LES-ROM, i.e., the set of equations for the filtered ROM variables; and (iii) Finally, use data-driven modeling to construct a ROM closure model for the filtered ROM equations developed in step (ii). In this section, we discuss steps (i) and (ii); in the next section, we discuss step (iii), i.e., we construct the DD-VMS-ROM.

To define the large ROM scales and build the VMS framework, we first decompose the d-dimensional ROM space $\boldsymbol{X}^d$ into two orthogonal subspaces

$$\boldsymbol{X}^r := \text{span}\{\boldsymbol{\varphi}_1, ..., \boldsymbol{\varphi}_r\}, \tag{3.1a}$$

$$(\boldsymbol{X}^r)^\perp := \text{span}\{\boldsymbol{\varphi}_{r+1}, ..., \boldsymbol{\varphi}_d\}, \tag{3.1b}$$

where $\boldsymbol{X}^r$ contains the first r dominant ROM basis functions, and $(\boldsymbol{X}^r)^\perp$, which is orthogonal to $\boldsymbol{X}^r$, contains the less energetic ROM basis functions. We also define the following orthogonal projections:

Definition 3.2 (Orthogonal Projections)**.** *Let $P_r : L^2 \to \boldsymbol{X}^r$ be the orthogonal projection onto $\boldsymbol{X}^r$, and $Q_r : L^2 \to (\boldsymbol{X}^r)^\perp$ be the orthogonal projection onto $(\boldsymbol{X}^r)^\perp$, which can be defined as*

$$P_r(\boldsymbol{u}) = \sum_{i=1}^{r} (\boldsymbol{u}, \boldsymbol{\varphi}_i)\boldsymbol{\varphi}_i, \quad \boldsymbol{u} \in L^2, \tag{3.2a}$$

$$Q_r(\boldsymbol{u}) = \sum_{i=r+1}^{d} (\boldsymbol{u}, \boldsymbol{\varphi}_i)\boldsymbol{\varphi}_i, \quad \boldsymbol{u} \in L^2, \tag{3.2b}$$

where L^2 denotes the space of square integrable functions on the spatial domain.

Next, in the LES spirit, we decompose the most accurate ROM solution at time step n, $\boldsymbol{u}_d^n$ (i.e., the d-dimensional G-ROM solution (2.6), which is the "truth" solution that is employed as a benchmark in our investigation) as

$$\boldsymbol{u}_d^n := \underbrace{P_r(\boldsymbol{u}_d^n)}_{\text{large scales}} + \underbrace{Q_r(\boldsymbol{u}_d^n)}_{\text{small scales}}, \tag{3.3}$$

where P_r and Q_r are the two orthogonal projections in Definition 3.2. Equation (3.3) represents the LES-ROM decomposition of the "truth" solution, $\boldsymbol{u}_d^n$, into its large scale component, $P_r(\boldsymbol{u}_d^n)$, and its small scale component, $Q_r(\boldsymbol{u}_d^n)$.

The ROM spatial filter that we use to construct the LES-ROM is the ROM projection filter [39, 53], i.e., the orthogonal projection P_r defined in Definition 3.2, which satisfies the following equation: For given $\boldsymbol{u} \in L^2$,

$$\big(P_r(\boldsymbol{u}), \boldsymbol{\varphi}_i\big) = \big(\boldsymbol{u}, \boldsymbol{\varphi}_i\big), \ \forall\, i = 1, ..., r. \tag{3.4}$$

To construct the LES-ROM, we need to construct the equation satisfied by the large scales, $P_r(\boldsymbol{u}_d^n)$, defined in (3.3). We note that, by using Definition 3.2 and the ROM orthogonality property, we obtain the following formula for the large scale component $P_r(\boldsymbol{u}_d^n)$:

$$P_r(\boldsymbol{u}_d^n) = \sum_{i=1}^{r} (\boldsymbol{a}_d^n)_i\, \boldsymbol{\varphi}_i. \tag{3.5}$$

To construct the LES-ROM satisfied by $P_r(\boldsymbol{u}_d^n)$, we apply the ROM spatial filter, P_r, to the equation satisfied by the "truth" solution, $\boldsymbol{u}_d^n$ (i.e., to the full discretization of the d-dimensional G-ROM (2.6)), we restrict the test functions in (2.6) to the r-dimensional ROM subspace $\boldsymbol{X}^r$ defined in (3.1a), and we use the decomposition (3.3). This yields the equations satisfied by the large scales, $P_r(\boldsymbol{u}_d^n)$, i.e., the LES-ROM equations:

$$\begin{aligned} \left(\frac{P_r(\boldsymbol{u}_d^n) - P_r(\boldsymbol{u}_d^{n-1})}{\Delta t}, \boldsymbol{v}_r\right) &+ Re^{-1}(\nabla P_r(\boldsymbol{u}_d^n), \nabla \boldsymbol{v}_r) + (P_r(\boldsymbol{u}_d^n) \cdot \nabla P_r(\boldsymbol{u}_d^n), \boldsymbol{v}_r) \\ &+ \mathcal{E}^n + (\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n), \boldsymbol{v}_r) = (\boldsymbol{f}^n, \boldsymbol{v}_r), \qquad \forall\, \boldsymbol{v}_r \in \boldsymbol{X}^r, \end{aligned} \tag{3.6}$$

where we used that, by (3.4), $(P_r(\boldsymbol{f}^n), \boldsymbol{v}_r) = (\boldsymbol{f}^n, \boldsymbol{v}_r)$. In the LES-ROM equations (3.6), the Reynolds stress tensor $\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n)$ and commutation error $\mathcal{E}$ are defined as follows:

$$\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n) := \boldsymbol{u}_d^n \cdot \nabla \boldsymbol{u}_d^n - P_r(\boldsymbol{u}_d^n) \cdot \nabla P_r(\boldsymbol{u}_d^n), \tag{3.7}$$

$$\mathcal{E}^n := Re^{-1}(\nabla Q_r(\boldsymbol{u}_d^n), \nabla \boldsymbol{v}_r), \tag{3.8}$$

respectively. We note that, to obtain the LES-ROM equations (3.6), we used the fact that the term $(Q_r(\boldsymbol{u}_d^n), \boldsymbol{v}_r)$ vanishes since $Q_r(\boldsymbol{u}_d^n)$ is orthogonal to any vector in $\boldsymbol{X}^r$. We also note that the term $(\nabla Q_r(\boldsymbol{u}_d^n), \nabla \boldsymbol{v}_r)$ in the commutation error term (3.8) does not vanish since the ROM basis functions are only L^2-orthogonal, not H_0^1-orthogonal.

Remark 3.1 (Commutation Error)**.** *In* [31]*, we investigated the effect of the commutation error* (3.8) *on ROMs. We showed that the commutation error is generally nonzero, but becomes negligible for large Re. Since our current investigation centers around LES-ROMs for turbulent flows, for simplicity, we do not consider the commutation error.*

Definition 3.3 (Closure Model)**.** *A closure model consists of replacing the Reynolds stress tensor* $\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n)$ *in* (3.6) *with another tensor* $\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n))$ *depending only on* $P_r(\boldsymbol{u}_d^n)$*.*

Thus, the role of the closure model $\boldsymbol{\tau}^{ROM}$ is to replace the true closure model $\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n)$ (which cannot be computed in $\boldsymbol{X}^r$) with a term that can actually be computed in $\boldsymbol{X}^r$. Since a closure model cannot in general be exact (i.e., $\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n) \neq \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n))$), when $\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n))$ is inserted for $\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n)$ in (3.6) the solution of the resulting system is just an approximation to $P_r(\boldsymbol{u}_d^n)$. We denote this LES-ROM approximation to $P_r(\boldsymbol{u}_d^n)$ as $\boldsymbol{u}_r^n$, which can be written as

$$\boldsymbol{u}_r^n = \sum_{i=1}^{r} (\boldsymbol{a}_r^n)_i \, \boldsymbol{\varphi}_i. \tag{3.9}$$

Thus, the LES-ROM equations for $\boldsymbol{u}_r^n$ are

$$\left(\frac{\boldsymbol{u}_r^n - \boldsymbol{u}_r^{n-1}}{\Delta t}, \boldsymbol{v}_r\right) + Re^{-1}(\nabla \boldsymbol{u}_r^n, \nabla \boldsymbol{v}_r) + (\boldsymbol{u}_r^n \cdot \nabla \boldsymbol{u}_r^n, \boldsymbol{v}_r) + (\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{v}_r) = (\boldsymbol{f}^n, \boldsymbol{v}_r), \ \forall \boldsymbol{v}_r \in \boldsymbol{X}^r. \tag{3.10}$$

Inserting (3.9) into (3.10) yields the following matrix form of the LES-ROM:

$$\frac{\boldsymbol{a}_r^n - \boldsymbol{a}_r^{n-1}}{\Delta t} = \boldsymbol{b}^n + \boldsymbol{A}\boldsymbol{a}_r^n + (\boldsymbol{a}_r^n)^T \boldsymbol{B} \boldsymbol{a}_r^n + [-(\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{\varphi}_i)_{i=1,\ldots,r}], \tag{3.11}$$

where the vector $\boldsymbol{b}^n$, the matrix $\boldsymbol{A}$, and the tensor $\boldsymbol{B}$ are defined in (2.9)-(2.11), but here are truncated to the first r components, i.e., the indices i, k, m in (2.9)-(2.11) are restricted between 1 and r. We opt for this slight abuse of notation in order to avoid introducing new variables that would overload the presentation. We also note that $[-(\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{\varphi}_i)_{i=1,\ldots,r}]$ in (3.11) denotes the $r \times 1$ vector whose i^{th} component is given by $-(\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{\varphi}_i)$.

4. Data Driven Variational Multiscale ROM (DD-VMS-ROM)

In this section, we outline the construction of the data-driven variational multiscale ROM (DD-VMS-ROM) closure model proposed in [37, 54]. We also describe the physical constraints that we add to the DD-VMS-ROM in order to increase its stability and accuracy. The construction of the DD-VMS-ROM is carried out within the LES-ROM framework described in Section 3.

To construct the DD-VMS-ROM, we start from the LES-ROM equations (3.11). First, we notice that since we used the ROM projection as a spatial filter, the LES-ROM (3.11) is in fact a variational multiscale ROM (VMS-ROM). However, the VMS-ROM (3.11) is not closed since the closure term $\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)$ still needs to be determined. To construct a VMS-ROM closure model, we use data-driven modeling. Specifically, we first postulate a linear ansatz for the VMS-ROM closure term, and then we determine the parameters in the linear ansatz that best match the FOM data. The linear ansatz for the VMS-ROM closure term can be written as follows:

$$[-(\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{\varphi}_i)_{i=1,\ldots,r}] \approx \widetilde{\boldsymbol{A}} \, \boldsymbol{a}_r^n, \tag{4.1}$$

where $\boldsymbol{a}_r^n$ is the vector of ROM coefficients of the solution $\boldsymbol{u}_r^n$; cf. (3.9). To determine the $r \times r$ matrix $\widetilde{\boldsymbol{A}}$ in (4.1), in the offline stage, we solve the following low-dimensional *least squares problem*:

$$\min_{\widetilde{\boldsymbol{A}}} \sum_{n=1}^{M} \left\| -\left[\left(\boldsymbol{u}_d^n \cdot \nabla \boldsymbol{u}_d^n - P_r(\boldsymbol{u}_d^n) \cdot \nabla P_r(\boldsymbol{u}_d^n) \, , \, \boldsymbol{\varphi}_i\right)_{i=1,\ldots,r}\right] - \underbrace{[(\widetilde{\boldsymbol{A}} \, \boldsymbol{b}_r^n)_{i=1,..,r}]}_{:=[(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n)), \boldsymbol{\varphi}_i)_{i=1,\ldots,r}]} \right\|^2, \tag{4.2}$$

where $\boldsymbol{u}_d^n$ and $P_r(\boldsymbol{u}_d^n)$ are obtained from the available FOM data and are defined in (2.4) and (3.5), respectively, and $\boldsymbol{b}_r^n$ is the r-dimensional vector that contains the first r entries of the vector $\boldsymbol{a}_d^n$.

Physical Constraint. In the numerical investigation in [16], it was shown that, in the mean, the LES-ROM closure model dissipates energy. Thus, to mimic this behavior, in [36] we equipped the DD-VMS-ROM with a similar physical constraint. Specifically, in the least squares problem (4.2), we added the constraint that $\widetilde{\boldsymbol{A}}$ be negative semidefinite:

$$(\boldsymbol{a}_r^n)^T \widetilde{\boldsymbol{A}} \boldsymbol{a}_r^n \leq 0 \qquad \forall\, \boldsymbol{a}_r^n \in \mathbb{R}^r. \tag{4.3}$$

For the numerical results presented in Section 6, this condition (4.3) is guaranteed by enforcing a particular structure on $\widetilde{\boldsymbol{A}}$. Specifically, we require the entries of $\widetilde{\boldsymbol{A}}$ to satisfy the following relations:

$$\widetilde{A}_{ij} = -\widetilde{A}_{ji},\ \forall\, i \neq j, \text{ and } \widetilde{A}_{ii} \leq 0,\ \forall\, i. \tag{4.4}$$

Solving the least squares problem (4.2) with the physical constraint (4.3), using the resulting matrix $\widetilde{\boldsymbol{A}}$ in the linear ansatz (4.1), and plugging this in the VMS-ROM (3.11) yields the data-driven variational multiscale ROM (DD-VMS-ROM):

$$\frac{\boldsymbol{a}_r^n - \boldsymbol{a}_r^{n-1}}{\Delta t} = \boldsymbol{b}^n + (\boldsymbol{A} + \widetilde{\boldsymbol{A}})\boldsymbol{a}_r^n + (\boldsymbol{a}_r^n)^T \boldsymbol{B} \boldsymbol{a}_r^n. \tag{4.5}$$

5. Verifiability of the DD-VMS-ROM

In this section, we prove the verifiability of the DD-VMS-ROM described in Section 4. In Section 5.1, we introduce the verifiability and mean dissipativity concepts in the ROM setting. In Section 5.2, we prove that the DD-VMS-ROM is verifiable.

5.1. Definition of Verifiability and Mean Dissipativity.

The goal of this subsection is to define the verfiability of ROM closure models. Verifiability of closure models has been investigated for decades in classical CFD (see, e.g., [30] as well as [9] for a survey of verifiability methods in LES). We emphasize, however, that, to our knowledge, the verifiability concept has not been defined in a ROM context. In this section, we take a first step in this direction and define verifiability of ROM closure models. We also define the mean dissipativity of ROM closures, which will be used in Section 5.2 to prove the verifiability of the DD-VMS-ROM.

In the remainder of this paper, we also use the following notation:

Definition 5.1 (Generic Constant C)**.** *We denote with C a generic constant that can depend on the fixed data (e.g., the solution, $\boldsymbol{u}$, the number of snapshots, M, the number of linearly independent snapshots, d, and the "truth" solution, $\boldsymbol{u}_d$), but not on the ROM parameters (e.g., the ROM dimension, r, and the ROM solution, $\boldsymbol{u}_r$).*

Definition 5.2 (Verifiability)**.** *Let the number of snapshots, M, (and, thus, the number of linearly independent snapshots, d) be fixed. A ROM closure model is verifiable in the L^2 norm, $||\cdot||$, if there is a constant C such that, for all $r \leq d$ and for all $n = 1, \ldots, M$, the following a priori error bound holds:*

$$\boxed{||\, P_r(\boldsymbol{u}_d^n) - \boldsymbol{u}_r^n||^2 \leq C \frac{1}{n} \sum_{j=1}^{n} ||P_r(\, \boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)\,)||^2,} \tag{5.1}$$

where $\boldsymbol{u}_d^j$ represents the "truth" solution (i.e., the d-dimensional G-ROM solution of (2.6)) at $t = t_j,\ j = 1, \ldots, M$, and $\boldsymbol{u}_r^n$ solves the ROM equipped with the given ROM closure model at $t = t_n,\ n = 1, \ldots, M$.

Definition 5.2 says that a ROM closure model is verifiable if a small average error in the ROM closure term implies a small error in the LES-ROM approximation.

Remark 5.1 (*A Priori* Error Bound)**.** *We emphasize that inequality (5.1) in the verifiability definition is an a priori error bound. This ROM error bound is similar to the a priori error bounds for classical FOMs, e.g., the FE method, which are often of the following form (see, e.g., Theorem 1.5 in [51]):*

$$\textit{error} \leq C\,(h^{p_1} + \Delta t^{p_2})\,, \tag{5.2}$$

where h is the spatial mesh size, Δt is the time step, p_1 and p_2 are exponents that depend on the particular finite element and time discretization used, and C is a generic constant that can depend on the problem data

(including the solution of the continuous problem), but not on the discretization parameters. As explained in Section 2.4 of [33]*, the a priori error bound* (5.2) *shows asymptotic convergence as* $h \to 0$ *and* $\Delta t \to 0$*, and can give the asymptotic rate of convergence with respect to the spatial and temporal discretizations. We emphasize that one essential feature of the FE a priori error bound* (5.2) *is that it can be proven before actually running the FE model (which explains the error bound's a priori qualifier). We note, however, that since the constant* C *on the right-hand side of* (5.2) *can depend on the unknown solution of the continuous problem, the a priori error bound* (5.2) *cannot be used to decide where the spatial mesh should be refined or coarsened. For that purpose, one could instead use a posteriori error bounds, in which the right-hand side depends entirely on computable quantities, e.g., the FE solution* [3]*.*

The ROM error bound (5.1) *in the verifiability definition is similar to the a priori FE error bound* (5.2)*. Indeed, the right-hand side of* (5.1) *does not depend on the ROM solution and can be evaluated before actually running the ROM. Thus, the ROM error bound* (5.1) *is an a priori error bound, just like the FE error bound* (5.2)*. Furthermore, the right-hand side of* (5.1) *is the product of a generic constant that does not depend on the ROM discretization parameters, and a term that can be tuned by the user (i.e., the average ROM closure error term). Thus, as the average ROM closure error in* (5.1) *decreases, we expect the ROM error to decrease at the same rate. Our numerical investigation in Section 6 shows that this is indeed the case. There is, however, a difference between the a priori ROM error bound* (5.1) *and the a priori FE error bound* (5.2)*: The latter depends on two FE parameters that can be easily adjusted (i.e., the spatial mesh size,* h*, and the time step,* Δt*). The former, however, depends on the average ROM closure error, which can be tuned by varying the parameters in the numerical discretization of the least squares problem* (4.2)*. This process is explained in Sections 6.1 and 6.2.*

Remark 5.2. *We note that the terms on the right-hand side of* (5.1) *in the verifiability definition are the same as those used in the least squares problem* (4.2)*. Furthermore, the* L^2 *norm is used in both* (5.1) *and* (4.2)*. Thus, solving the least squares problem* (4.2) *to construct the DD-VMS-ROM and proving that the DD-VMS-ROM is verifiable (as we will do in Theorem* (5.2)*) should yield accurate DD-VMS-ROM approximations. The numerical investigation in Section 6 will show that, as expected, the DD-VMS-ROM approximations are accurate.*

Definition 5.3 (Mean Dissipativity)**.** *A ROM closure model satisfies the mean dissipativity condition if for the* $\boldsymbol{u}_d^n, \boldsymbol{u}_r^n$*, and* n *given in Definition 5.2, the following inequalities are satisfied:*

$$\boxed{0 \leq (\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)\,,\, P_r(\boldsymbol{u}_d^n) - \boldsymbol{u}_r^n) < \infty.} \tag{5.3}$$

5.2. **Proof of DD-VMS-ROM's Verifiability.** In this section, we first prove that the DD-VMS-ROM is mean dissipative. Then, we use this result to prove that the DD-VMS-ROM is verifiable.

Theorem 5.1. *The DD-VMS-ROM with linear ansatz* (4.5) *and physical constraint* (4.3) *satisfies mean dissipativity according to Definition 5.3.*

Proof. The least squares problem (4.2) yields the ROM operator $\widetilde{\boldsymbol{A}}$ for $-(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n), \boldsymbol{\varphi}_i)$, which is the VMS-ROM closure term. We note that the same ROM operator $\widetilde{\boldsymbol{A}}$ is used to construct the VMS-ROM closure term $-(\tau^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{\varphi}_i)$. Specifically, the ROM operator $\widetilde{\boldsymbol{A}}$ that is created by solving the least squares problem (4.2) for the VMS-ROM closure term $-(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n), \boldsymbol{\varphi}_i)$ is used in the linear ansatz $-(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n), \boldsymbol{\varphi}_i)_{i=1,\ldots,r} \approx \widetilde{\boldsymbol{A}}\,\boldsymbol{b}_r$, where $\boldsymbol{b}_r^n$ is the r-dimensional vector defined in (4.2), i.e., the r-dimensional vector that contains the first r entries of the vector $\boldsymbol{a}_d^n$. The same ROM operator $\widetilde{\boldsymbol{A}}$ is also used in the linear ansatz (4.1) for the VMS-ROM closure term $-(\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{\varphi}_i)$: $-(\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{\varphi}_i)_{i=1,\ldots,r} \approx \widetilde{\boldsymbol{A}}\,\boldsymbol{a}_r$. We approximate the VMS-ROM closure terms with these ansatzes and obtain the following equalities:

$$\begin{aligned}
(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)\,,\, \boldsymbol{\varphi}_i) &= \Big(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n))\,,\, \boldsymbol{\varphi}_i\Big) - \Big(\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)\,,\, \boldsymbol{\varphi}_i\Big) \\
&= (-\widetilde{\boldsymbol{A}}\,\boldsymbol{b}_r^n)_i - (-\widetilde{\boldsymbol{A}}\,\boldsymbol{a}_r^n)_i \\
&= \big(-\widetilde{\boldsymbol{A}}\,(\boldsymbol{b}_r^n - \boldsymbol{a}_r^n)\big)_i \qquad \forall i = 1, .., r.
\end{aligned} \tag{5.4}$$

To prove that the inner product $(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)\,,\, P_r(\boldsymbol{u}_d^n) - \boldsymbol{u}_r^n)$ is non-negative, we use the definitions of $P_r(\boldsymbol{u}_d^n)$ in (3.5) and $\boldsymbol{u}_r^n$ in (3.9), and rewrite it as follows:

$$\begin{aligned}\Big(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)\,,\, P_r(\boldsymbol{u}_d^n) - \boldsymbol{u}_r^n\Big) &= \Big(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)\,,\, \sum_{i=1}^{r}(\boldsymbol{a}_d^n - \boldsymbol{a}_r^n)_i\,\boldsymbol{\varphi}_i\Big) \\ &= \sum_{i=1}^{r}(\boldsymbol{a}_d^n - \boldsymbol{a}_r^n)_i\,\Big(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)\,,\ \boldsymbol{\varphi}_i\Big).\end{aligned} \tag{5.5}$$

By applying (5.4) to (5.5) and using the physical constraint (4.3), we get

$$\begin{aligned}(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n)\,,\, P_r(\boldsymbol{u}_d^n) - \boldsymbol{u}_r^n) &= \sum_{i=1}^{r}(\boldsymbol{a}_d^n - \boldsymbol{a}_r^n)_i\,\big(-\widetilde{\boldsymbol{A}}\,(\boldsymbol{b}_r^n - \boldsymbol{a}_r^n)\big)_i \\ &= -(\boldsymbol{b}_r^n - \boldsymbol{a}_r^n)^T\,\widetilde{\boldsymbol{A}}\,(\boldsymbol{b}_r^n - \boldsymbol{a}_r^n) \geq 0,\end{aligned} \tag{5.6}$$

since $\widetilde{\boldsymbol{A}}$ is negative semi-definite. In (5.6), we have used that $\boldsymbol{b}_r^n$ is an r-dimensional vector that contains the first r entries of the $\boldsymbol{a}_d^n$. The inequality in (5.6) concludes the proof. □

Remark 5.3. *We note that in Theorem 5.1 we proved the ROM mean dissipativity property only for $P_r(\boldsymbol{u}_d^n)$ and $\boldsymbol{u}_r^n$. This is contrast with the FEM context, where mean dissipativity is proven for general FEM functions (see, e.g.,* [30]*). However, the result presented in Theorem 5.1 is sufficient for proving the verifibility property given in Theorem 5.2 below.*

Next, we prove that the DD-VMS-ROM is verifiable. We note that, as explained in Section 3, the goal for the DD-VMS-ROM solution is to approximate as accurately as possible $P_r(\boldsymbol{u}_d^n)$, which is the large scale component of the d-dimensional G-ROM solution (2.6), i.e., the "truth" solution that is employed as a benchmark in our investigation. Furthermore, as explained in the second paragraph following Definition 3.1, the "truth" solution, $\boldsymbol{u}_d$, will be considered as given data. We also note that $P_r(\boldsymbol{u}_d^n)$ satisfies the LES-ROM equations (3.6), which, for clarity, we rewrite below:

$$\begin{aligned}(\frac{P_r(\boldsymbol{u}_d^n) - P_r(\boldsymbol{u}_d^{n-1})}{\Delta t}, \boldsymbol{v}_r) + Re^{-1}(\nabla P_r(\boldsymbol{u}_d^n), \nabla \boldsymbol{v}_r) + (P_r(\boldsymbol{u}_d^n)\cdot\nabla P_r(\boldsymbol{u}_d^n), \boldsymbol{v}_r) \\ +(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n), \boldsymbol{v}_r) = (\boldsymbol{f}^n, \boldsymbol{v}_r),\end{aligned} \tag{5.7}$$

where we used the fact that $(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n), \boldsymbol{v}_r)$ is equal to $(P_r(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n)), \boldsymbol{v}_r)$. We also rewrite the full discretization of the DD-VMS-ROM (3.10):

$$\begin{aligned}(\frac{\boldsymbol{u}_r^n - \boldsymbol{u}_r^{n-1}}{\Delta t}, \boldsymbol{v}_r) + Re^{-1}(\nabla \boldsymbol{u}_r^n, \nabla \boldsymbol{v}_r) + (\boldsymbol{u}_r^n\cdot\nabla \boldsymbol{u}_r^n, \boldsymbol{v}_r) \\ +(\boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^n), \boldsymbol{v}_r) = (\boldsymbol{f}^n, \boldsymbol{v}_r).\end{aligned} \tag{5.8}$$

Furthermore, we use the linear ansatz (4.1) and the physical constraint (4.3) for the ROM closure model in the DD-VMS-ROM (5.8). We also choose the initial condition $\boldsymbol{u}_r^0 = P_r(\boldsymbol{u}_d^0)$.

The DD-VMS-ROM error at time step n, which we denote with $\boldsymbol{e}^n$, is defined as the difference between the large scale component of the "truth" solution, $P_r(\boldsymbol{u}_d^n)$ (which is the solution of (5.7)), and the DD-VMS-ROM solution of (5.8), $\boldsymbol{u}_r^n$: $\boldsymbol{e}^n = P_r(\boldsymbol{u}_d^n) - \boldsymbol{u}_r^n$.

To prove the DD-VMS-ROM's verifiability, we use the following bound on the nonlinear term, which is given in Lemma 22 in [33] (see also Lemma 61.1 in [48]):

Lemma 5.1. *Let $\Omega \subset \mathbb{R}^q$ be an open, bounded set of class C^2, with $q = 2$ or 3. For all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in [\boldsymbol{H}_0^1(\Omega)]^q$,*

$$b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \leq C(\Omega)\sqrt{||\boldsymbol{u}||\,||\nabla \boldsymbol{u}||}\,||\nabla \boldsymbol{v}||\,||\nabla \boldsymbol{w}||, \tag{5.9}$$

where the trilinear form $b(\cdot,\cdot,\cdot)$ [33, 50] *is defined as*

$$b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{u}\cdot\nabla \boldsymbol{v}, \boldsymbol{w}). \tag{5.10}$$

Theorem 5.2. *The DD-VMS-ROM* (5.8) *with linear ansatz* (4.1)*, physical constraint* (4.3)*, and the initial condition* $\boldsymbol{u}_r^0 = P_r(\boldsymbol{u}_d^0)$ *is verifiable: For a small enough time step,* $\Delta t\, d_j < 1,\ \forall\, j = 1, ..., M$*, where* $d_j = \Big(\frac{27(Re)^3 C(\Omega)^4}{16}||\nabla P_r(\boldsymbol{u}_d^j)||^4 + Re\Big)$ *and* $C(\Omega)$ *is the constant in Lemma 5.1, the following inequality holds for all* $n = 1, \ldots, M$*:*

$$
\begin{aligned}
&||\boldsymbol{e}^n||^2 + \Delta t \sum_{j=1}^{n} Re^{-1}||\nabla \boldsymbol{e}^j||^2 \leq \\
&\exp\Big(\Delta t \sum_{j=1}^{n} \frac{d_j}{1-\Delta t d_j}\Big)\Big(\Delta t \sum_{j=1}^{n} Re^{-1}||P_r(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)))\,||^2\Big),
\end{aligned}
\tag{5.11}
$$

where $\boldsymbol{e}^n = P_r(\boldsymbol{u}_d^n) - \boldsymbol{u}_r^n$*.*

Proof. We subtract (5.8) from (5.7), and replace n with j to get the error equation:

$$
\begin{aligned}
&(\frac{\boldsymbol{e}^j - \boldsymbol{e}^{j-1}}{\Delta t}, \boldsymbol{v}_r) + Re^{-1}(\nabla \boldsymbol{e}^j, \nabla \boldsymbol{v}_r) + b(P_r(\boldsymbol{u}_d^j), P_r(\boldsymbol{u}_d^j), \boldsymbol{v}_r) - b(\boldsymbol{u}_r^j, \boldsymbol{u}_r^j, \boldsymbol{v}_r) \\
&\qquad + \big(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^j), \boldsymbol{v}_r\big) = -\big(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)), \boldsymbol{v}_r\big).
\end{aligned}
\tag{5.12}
$$

We set $\boldsymbol{v}_r = \boldsymbol{e}^j$ in (5.12), add and subtract $b(\boldsymbol{u}_r^j, P_r(\boldsymbol{u}_d^j), \boldsymbol{e}^j)$, and use the fact that $b(\boldsymbol{u}_r^j, \boldsymbol{e}^j, \boldsymbol{e}^j) = 0$ to get the following equation:

$$
\begin{aligned}
&\Delta t^{-1}(\boldsymbol{e}^j - \boldsymbol{e}^{j-1}, \boldsymbol{e}^j) + Re^{-1}||\nabla \boldsymbol{e}^j||^2 + b(\boldsymbol{e}^j, P_r(\boldsymbol{u}_d^j), \boldsymbol{e}^j) \\
&\qquad + (\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^j), \boldsymbol{e}^j) = -(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)), \boldsymbol{e}^j).
\end{aligned}
\tag{5.13}
$$

From Theorem 5.1, we have the following inequality:

$$
(\boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)) - \boldsymbol{\tau}^{ROM}(\boldsymbol{u}_r^j), \boldsymbol{e}^j) \geq 0.
\tag{5.14}
$$

By applying (5.14) to (5.13), we get the following inequality:

$$
\Delta t^{-1}\big(\boldsymbol{e}^j - \boldsymbol{e}^{j-1}, \boldsymbol{e}^j\big) + Re^{-1}||\nabla \boldsymbol{e}^j||^2 \leq -b(\boldsymbol{e}^j, P_r(\boldsymbol{u}_d^j), \boldsymbol{e}^j) - \big(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)), \boldsymbol{e}^j\big).
\tag{5.15}
$$

Applying Hölder's and Young's inequalities to the terms $(\boldsymbol{e}^j - \boldsymbol{e}^{j-1}, \boldsymbol{e}^j)$ and $-(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)), \boldsymbol{e}^j)$ in (5.15) we obtain that, for any $C_1, C_2 > 0$, the following inequalities hold:

$$
\begin{aligned}
(\boldsymbol{e}^j - \boldsymbol{e}^{j-1}, \boldsymbol{e}^j) &= ||\boldsymbol{e}^j||^2 - (\boldsymbol{e}^j, \boldsymbol{e}^{j-1}) \\
&\geq ||\boldsymbol{e}^j||^2 - ||\boldsymbol{e}^j||\,||\boldsymbol{e}^{j-1}|| \\
&\geq ||\boldsymbol{e}^j||^2 - \frac{C_1}{2}||\boldsymbol{e}^j||^2 - \frac{1}{2C_1}||\boldsymbol{e}^{j-1}||^2
\end{aligned}
\tag{5.16}
$$

and

$$
\begin{aligned}
|-(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)), \boldsymbol{e}^j)| &= |-(P_r(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j))), \boldsymbol{e}^j)| \\
&\leq \frac{1}{2C_2}||P_r(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)))\,||^2 + \frac{C_2}{2}||\boldsymbol{e}^j||^2.
\end{aligned}
\tag{5.17}
$$

Applying Lemma 5.1 to the term $-b(\boldsymbol{e}^j, P_r(\boldsymbol{u}_d^j), \boldsymbol{e}^j)$, we obtain the following inequality for any $C_3 > 0$:

$$
\begin{aligned}
|-b(\boldsymbol{e}^j, P_r(\boldsymbol{u}_d^j), \boldsymbol{e}^j)| &\leq C(\Omega)\,||\nabla \boldsymbol{e}^j||^{3/2}\,||\nabla P_r(\boldsymbol{u}_d^j)||\,||\boldsymbol{e}^j||^{1/2} \\
&\leq \frac{3C_3 C(\Omega)}{4}||\nabla \boldsymbol{e}^j||^2 + \frac{C(\Omega)}{4(C_3)^3}||\nabla P_r(\boldsymbol{u}_d^j)||^4||\boldsymbol{e}^j||^2,
\end{aligned}
\tag{5.18}
$$

where $C(\Omega)$ is the constant in Lemma 5.1.

By choosing $C_1 = 1$, $C_2 = Re$, and $C_3 = 2Re^{-1}/3C(\Omega)$, we get the following inequality:

$$
\begin{aligned}
&\frac{1}{2\Delta t}(||\boldsymbol{e}^j||^2 - ||\boldsymbol{e}^{j-1}||^2) + \frac{Re^{-1}}{2}||\nabla \boldsymbol{e}^j||^2 \\
&\leq \Big(\frac{27(Re)^3 C(\Omega)^4}{32}||\nabla P_r(\boldsymbol{u}_d^j)||^4 + \frac{Re}{2}\Big)||\boldsymbol{e}^j||^2 + \frac{Re^{-1}}{2}||P_r(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)))\,||^2.
\end{aligned}
\tag{5.19}
$$

By multiplying (5.19) by $2\Delta t$ and summing the resulting inequalities from $j = 1$ to n, we obtain the following inequality:

$$
\begin{aligned}
||\boldsymbol{e}^n||^2 + \Delta t \sum_{j=1}^{n} Re^{-1}||\nabla \boldsymbol{e}^j||^2 \leq ||\boldsymbol{e}^0||^2 + \Delta t \sum_{j=1}^{n} \Big(\frac{27(Re)^3 C(\Omega)^4}{16} ||\nabla P_r(\boldsymbol{u}_d^j)||^4 + Re \Big) ||\boldsymbol{e}^j||^2 \\
+ \Delta t \sum_{j=1}^{n} Re^{-1} ||P_r(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)))\,||^2.
\end{aligned} \tag{5.20}
$$

To apply the discrete Gronwall's lemma, we first make the following notation:

$$
\begin{aligned}
a_j &:= ||\boldsymbol{e}^j||^2 \geq 0, \\
b_j &:= Re^{-1}||\nabla \boldsymbol{e}^j||^2 \geq 0, \\
d_j &:= \Big(\frac{27(Re)^3 C(\Omega)^4}{16} ||\nabla P_r(\boldsymbol{u}_d^j)||^4 + Re \Big) \geq 0, \\
c_j &:= Re^{-1}||P_r(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^j) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^j)))\,||^2 \geq 0, \\
H &:= ||\boldsymbol{e}^0||^2 \geq 0.
\end{aligned} \tag{5.21}
$$

We also recall that, by the small time step assumption, the following inequality holds: $\Delta t\, d_j < 1, \forall j$. By using the notation in (5.21), we rewrite (5.20) as follows:

$$
a_n + \Delta t \sum_{j=1}^{n} b_j \leq \Delta t \sum_{j=1}^{n} d_j\, a_j + \Delta t \sum_{j=1}^{n} c_j + H. \tag{5.22}
$$

By using the discrete Gronwall's lemma (see Lemma 27 in [33]) in (5.22), we obtain the following inequality:

$$
a_n + \Delta t \sum_{j=1}^{n} b_j \leq \exp\Big(\Delta t \sum_{j=1}^{n} \frac{d_j}{1 - \Delta t d_j} \Big) \Big(\Delta t \sum_{j=1}^{n} c_j + H \Big). \tag{5.23}
$$

We note that choosing the initial condition $\boldsymbol{u}_r^0 = P_r(\boldsymbol{u}_d^0)$, implies that $\boldsymbol{e}^0 = \boldsymbol{u}_r^0 - P_r(\boldsymbol{u}_d^0) = 0$, and thus $H = 0$. As a result, (5.23) implies that (5.11) holds.

□

Remark 5.4. *We note that the small time step assumption that we made in the theorem, i.e., that $\Delta t\, d_j < 1\ \forall j = 1, ..., M$, is also made in a FE context (see Lemma 27 and the proof of Theorem 24 in* [33]*).*

Remark 5.5. *In this paper, we used backward Euler time discretization to obtain the full discretizations of the ROMs. However, other time discretization schemes could be applied as well.*

6. Numerical Results

In Theorem 5.2, we proved that the DD-VMS-ROM presented in Section 4 is verifiable. In this section, we present numerical support for the theoretical results in Theorem 5.2. In Section 6.1, we provide details on the numerical implementation of the DD-VMS-ROM. We numerically show that the DD-VMS-ROM is verifiable for the Burgers equation in Section 6.3 and for the flow past a cylinder in Section 6.4.

6.1. **Numerical Implementation.**

"Truth" Solution. For computational efficiency, instead of solving the very large-dimensional G-ROM (2.5) to get the "truth" solution, $\boldsymbol{u}_d$, we simply project the FOM data on the ROM space, i.e., $\boldsymbol{u}_d = P_r(\boldsymbol{u}_h)$, $r = d$. In our numerical investigation, the two approaches yield similar results (i.e., the difference between the two approaches is on the order of the time discretization error). Thus, using the projection of the FOM data as "truth" solution does not affect our numerical investigation of the DD-VMS-ROM's verifiability.

Algorithm 1: Least Squares Regularization: Truncated SVD

1: Formulate the standard linear least squares problem for the unknown vector $\boldsymbol{x}_u$:

$$\min_{\boldsymbol{x}_u} \left\| E\boldsymbol{x}_u - \boldsymbol{f} \right\|^2, \tag{6.1}$$

where $E \in \mathbb{R}^{Mr \times r^2}$ is a matrix whose entries are determined by $\boldsymbol{a}_d(t_j), j = 1, \cdots, M$, $\boldsymbol{f} \in \mathbb{R}^{Mr \times 1}$ is a vector whose entries are determined by $P_r(\boldsymbol{\tau}^{FOM}(t_j))$, and $\boldsymbol{x}_u \in \mathbb{R}^{r^2 \times 1}, j = 1, \cdots, M$, is a vector whose entries are determined by $\widetilde{\boldsymbol{A}}$.

2: Calculate the SVD of E:

$$E = U \Sigma V^\top. \tag{6.2}$$

3: Specify a tolerence *tol*.

4: Keep the entries in Σ that are larger than *tol*; the resulting matrix is $\widetilde{\Sigma}$ ($\widetilde{\sigma} = \sigma$ if $\sigma > tol$; the singular values of E can be chosen as *tol* values).

5: Construct $\widetilde{E}$, the truncated SVD of E:

$$\widetilde{E} = \widetilde{U} \widetilde{\Sigma} \widetilde{V}^\top, \tag{6.3}$$

where $\widetilde{U}$ and $\widetilde{V}$ are the entries of U and V that correspond to $\widetilde{\Sigma}$, respectively.

6: The solution is given by

$$\boldsymbol{x}_u = \left(\widetilde{V} \widetilde{\Sigma}^{-1} \widetilde{U}^\top \right) \boldsymbol{f}. \tag{6.4}$$

Least Squares Regularization: Truncated SVD. As is often the case in data-driven modeling [41], the least squares problem (4.2) that we need to solve in order to determine the entries in the ROM closure operator $\widetilde{A}$ used to construct the DD-VMS-ROM (4.5) can be ill conditioned. To alleviate the ill conditioning of the least squares problem, we proposed the use of the truncated SVD [54, 37] as a regularization method [20, Chapter 4] (see also [58] for a related approach). For completeness, in Algorithm 1, we outline the construction of the DD-VMS-ROM with the truncated SVD procedure.

The tolerance *tol* specified in step 3 of Algorithm 1 (which yields the truncation parameter k, i.e., the index of the lowest singular value retained in the matrix $\widetilde{\Sigma}$ constructed in step 4 of Algorithm 1; see equations (4.2) and (4.3) in [20, Chapter 4]) plays an important role in the numerical implementation of the DD-VMS-ROM. Specifying a large *tol* value yields a well conditioned least squares problem in step 1 and, as a result, minimizes the numerical errors in the least squares problem. However, a large *tol* value also decreases the accuracy of the least squares problem, i.e., yields a DD-VMS-ROM closure operator $\tilde{A}$ that does not accurately match the FOM data. On the other hand, choosing a small *tol* value does not significantly decrease the accuracy of the DD-VMS-ROM closure operator $\tilde{A}$, but does not significantly alleviate the ill conditioning of the least squares problem either. In our numerical investigation, a careful choice of the tolerance *tol* yields optimal DD-VMS-ROM results.

If physical constraints such as that given by (4.4) are added when solving the minimization problem (6.1), then the optimal $\widetilde{A}$ given by (6.4) associated with a specified *tol* should be replaced by the solution of a constrained linear least squares solver with $\widetilde{E}$ given by (6.3) as the data matrix. For all the numerical results presented in Section 6.3 and Section 6.4, we use the Matlab built-in solver `lsqlin` for this purpose. Specifically, we use the interior-point algorithm option for `lsqlin` with `ConstraintTolerance` = 1E-10, `OptimalityTolerance` = 1E-9, `StepTolerance` = 1E-12, and `MaxIter` = 1000.

Time Discretization. Although the DD-VMS-ROM's verifiability was proven in Theorem 5.2 for the backward Euler time discretization, in the numerical investigation of the flow past a cylinder (Section 6.4), we use the linearized BDF2 time discretization. We use this higher-order time discretizations in order to decrease the impact of the time discretization error onto the LES-ROM error, which is the main focus of the numerical investigation in this section. Furthermore, we believe that the mathematical arguments used to prove the

DD-VMS-ROM's verifiability in Theorem 5.2 can be extended to higher-order time discretizations such as that considered in Section 6.4.

Criteria. To illustrate numerically the DD-VMS-ROM verifiability proven in Theorem 5.2, we use the following approach, which was outlined in Section 3 (see, e.g., the discussion after Definition 3.1) and Section 5 (see, e.g., Definitions 5.1 and 5.2): First, we fix the number of snapshots, M. Therefore, the maximal dimension of the ROM space, d, is also fixed. Furthermore, the "truth" solution $\boldsymbol{u}_d$ (i.e., the solution of the d-dimensional G-ROM (2.5)) is also fixed. The goal of our numerical investigation is to show that, for fixed M, d, and $\boldsymbol{u}_d$, there exists a constant C (see Definition 5.1) such that for varying r values and for varying *tol* values, the inequality (5.11) is satisfied. Thus, the goal is to bound the error between the DD-VMS-ROM solution, $\boldsymbol{u}_r$, and the large scale component of the "truth" solution, $P_r(\boldsymbol{u}_d)$.

To this end, we use the following metrics: To quantify the LES-ROM error, i.e., the averaged error associated with the first term on the LHS of inequality (5.11) (see also the LHS of (5.1)), we use the following average L^2 norm:

$$\mathcal{E}(L^2) = \frac{1}{M}\sum_{n=1}^{M} \|P_r(\boldsymbol{u}_d^n) - \boldsymbol{u}_r^n\|^2 = \frac{1}{M}\sum_{n=1}^{M} \|\boldsymbol{e}^n\|^2 \,. \tag{6.5}$$

To quantify the LES-ROM closure error, i.e., the term on the RHS of inequality (5.11), we use the following metric:

$$\eta(L^2) = \frac{1}{M}\sum_{n=1}^{M} \left\|P_r(\boldsymbol{\tau}^{FOM}(\boldsymbol{u}_d^n) - \boldsymbol{\tau}^{ROM}(P_r(\boldsymbol{u}_d^n))\right\|_{L^2}^2 \,. \tag{6.6}$$

Note that the ROM error $\mathcal{E}(L^2)$ and the closure error $\eta(L^2)$ depend on both the dimension r of the LES-ROM and the aforementioned tolerance index (i.e., truncation parameter) k associated with the tolerance of the truncated SVD used for constructing $\widetilde{A}$ for the given r. We suppressed these dependencies to simplify the notation. It should be clear from the context which parameter is varied for each of the numerical results presented below.

6.2. **Assessment of Results.** To illustrate numerically the DD-VMS-ROM verifiability proven in Theorem 5.2, we need to show that as $\eta(L^2)$ in (6.6) decreases, so does $\mathcal{E}(L^2)$ in (6.5). Specifically, according to (5.11) (see also Definition 5.2), we should see $\log(\eta(L^2))$ and $\log(\mathcal{E}(L^2))$ obey the following relation:

$$\log(\mathcal{E}(L^2)) \leq \alpha \log(\eta(L^2)) + \beta, \tag{6.7}$$

with $\alpha = 1$ and some $\beta > 0$. As pointed out above, both $\mathcal{E}(L^2)$ and $\eta(L^2)$ depend on two parameters: the ROM dimension r and the tolerance index k in the truncated SVD. In the numerical investigation, we perform two types of experiments:

(i) For a fixed r, we aim to show that (6.7) holds with $\alpha \geq 1$ as k is varied;
(ii) For each r, we pick the corresponding k that minimizes $\mathcal{E}(L^2)$, and aim to show that (6.7) holds with $\alpha \geq 1$ as r is varied.

Since in practice one is interested in the settings for which $\eta(L^2)$ is relatively small, a rate $\alpha > 1$ indicates a better rate than the rate predicted by Theorem 5.2.

We would like to note that our numerical investigation is somewhat different from the standard investigations used in the numerical analysis literature. While increasing the ROM dimension r is analogous to reducing the mesh size h in numerical analysis, the tolerance index k for the truncated SVD (which is tied specifically to the data-driven aspect of the LES-ROM closure examined here) has no analogue in classical numerical analysis.

6.3. **Burgers Equation.** In this section, we investigate the DD-VMS-ROM verifiability in the numerical simulation of the one-dimensional viscous Burgers equation:

$$\begin{cases} u_t - \nu u_{xx} + u u_x = 0 \ , \ x \in (0,1), \ t \in (0,1], \\ u(0,t) = u(1,t) = 0 \ , \ t \in (0,1], \\ u(x,0) = u_0(x) \ , \ x \in [0,1], \end{cases} \tag{6.8}$$

with non-smooth initial condition (6.9):

$$u_0(x) = \begin{cases} 1, & x \in (0, 1/2], \\ 0, & x \in (1/2, 1]. \end{cases} \tag{6.9}$$

This test problem has been used in, e.g., [1, 32, 54].

Snapshot Generation. We generate the FOM results by using a linear finite element (FE) spatial discretization with mesh size $h = 1/2048$, a backward Euler time discretization with timestep size $\Delta t = 5 \times 10^{-4}$, and a viscosity coefficient $\nu = 10^{-2}$. Due to the parabolic nature of the Burgers equation (6.8), the discontinuity in the initial data (6.9) is smoothed out as soon as $t > 0$. It becomes a (smooth) viscous shock with relatively steep gradient due to the small viscosity used, and persists for the whole duration of the time integration, i.e., for t in $[0, 1]$. See also [14, 24], where a stochastic version of this type of viscous shocks is considered within a reduced order modeling context.

ROM Construction. We run the FOM from $t = 0$ to $t = 1$, which yields a total of 2001 solution snapshots. Since the spatial derivatives of the FOM solution are involved in the $\boldsymbol{\tau}^{FOM}$ part of the closure error $\eta(L^2)$ (see (3.7)) and the initial condition given by (6.9) is discontinuous, we remove the FOM solution in the time interval $[0, 0.01)$, and thus collect a total of 1981 equally spaced snapshots in the time interval $[0.01, 1]$ to generate the ROM basis functions. To train the DD-VMS-ROM closure operator $\widetilde{\boldsymbol{A}}$, we use FOM data on the same time interval $[0.01, 1]$. We also test the DD-VMS-ROM on the time interval $[0.01, 1]$. That is, each ROM is initialized at $t = 0.01$ using the projected FOM data and run up to $t = 1$, and the ROM error $\mathcal{E}(L^2)$ in (6.5) and the closure error $\eta(L^2)$ in (6.6) are both computed over the time interval $[0.01, 1]$. Thus, we consider the reconstructive regime. The ROMs are integrated with the backward Euler time discretization and the same timestep size as that used for the FOM.

Numerical Results. We begin by presenting the results obtained for the first type of experiments outlined in Section 6.2. That is, we fix the ROM dimension r, and examine how $\mathcal{E}(L^2)$ in (6.5), which measures the DD-VMS-ROM error, and $\eta(L^2)$ in (6.6), which measures the DD-VMS-ROM closure error, vary as the tolerance index k in the truncated SVD used in the data-driven modeling part is varied. Specifically, we monitor the decaying rate of $\mathcal{E}(L^2)$ with respect to $\eta(L^2)$ as k is varied. The results in Figure 1, for $r = 8, 14$, and 20, generally show that, as $\eta(L^2)$ decreases (red curves), so does $\mathcal{E}(L^2)$ (blue curves). We note, however, that as shown for $r = 8$ and $r = 14$ in Figure 1, the global minimum of the ROM error $\mathcal{E}(L^2)$ may not be achieved at $k = r^2$, which corresponds to the case when the full SVD is used for constructing the data-driven closure term $\widetilde{A}$; see the caption of Figure 1. We also note that larger local fluctuations in both curves are displayed for $r = 14$ and $r = 20$, which is due to the fact that the condition number of the data matrix $E^\top E$ increases significantly for these two values. Indeed, the condition number of $E^\top E$ is 6.35×10^6 for $r = 20$, 1.2×10^5 for $r = 14$, and 1.6×10^3 for $r = 8$ [1].

With the k-dependence data available, we turn now to examining the relation (6.7) for fixed r values while k is varied. For this purpose, in Figure 2, we plot the corresponding linear regression (LR) slope. We note that the LR slopes shown in Figure 2 are computed based on those ($\mathcal{E}(L^2)$, $\eta(L^2)$) data pairs for which $\eta(L^2) \leq 100$ since most of the data pairs are aggregated below that threshold and, more importantly, the cases with small $\eta(L^2)$ are those of practical interest. The results in Figure 2 show that (6.7) holds with α either greater than 1 or just slightly below 1.

[1] For the case $r = 20$, for about 5% of the total 400 possible k values, the constrained linear least squares solver `lsqlin` fails to converge. These k values are scattered around $k = 300$. We did not include the corresponding $\mathcal{E}(L^2)$ and $\eta(L^2)$ data in Figure 1 and we also excluded them when computing the corresponding linear regression slope presented in Figure 2.

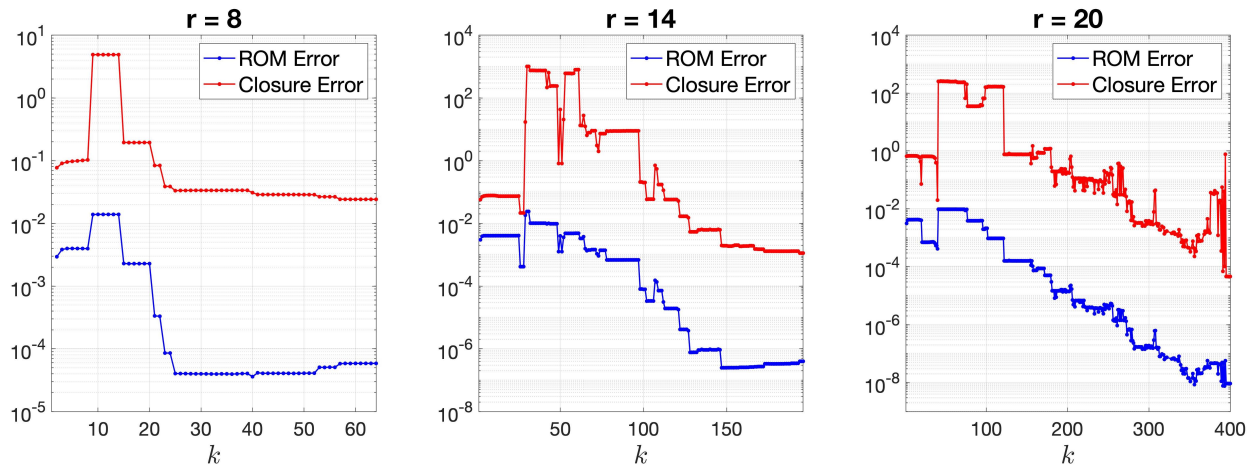

FIGURE 1. Burgers equation (6.8), reconstructive regime: $\mathcal{E}(L^2)$ and $\eta(L^2)$ for three fixed r values and different tolerance index k values in the truncated SVD. Recall that $\mathcal{E}(L^2)$ and $\eta(L^2)$ are defined by (6.5) and (6.6), respectively. As mentioned in Section 6.1, the tolerance values in the truncated SVD take the form of the truncation index k, which is the index of the lowest singular value retained in the matrix $\widetilde{\Sigma}$ constructed in step 4 of Algorithm 1. For an r-dimensional ROM, the matrix E in Algorithm 1 is of dimension $Mr \times r^2$; cf. (6.1). Thus, the tolerance index k can take values between 1 and r^2. As a result, there are r^2 data points in each of the three panels for both $\mathcal{E}(L^2)$ (blue curve) and $\eta(L^2)$ (red curve).

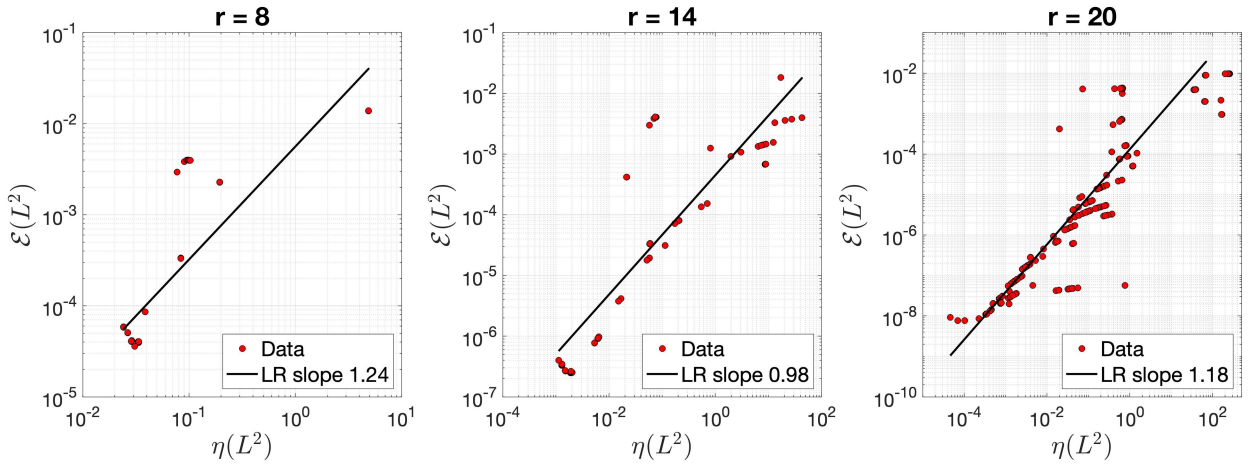

FIGURE 2. Burgers equation (6.8), reconstructive regime: linear regression for $\mathcal{E}(L^2)$ and $\eta(L^2)$ for fixed r values and different tolerance values in the truncated SVD. The red dots in each panel correspond to the data points $(\mathcal{E}(L^2), \eta(L^2))$ shown in the corresponding panel in Figure 1. The linear regression for $\mathcal{E}(L^2)$ in terms of $\eta(L^2)$ in each panel is indicated by the solid black line.

Next, we consider the other type of experiments, in which we vary r, and for each r we pick the corresponding k that minimizes $\mathcal{E}(L^2)$. These results are plotted in Figure 3, which shows again that (6.7) holds with $\alpha \geq 1$, this time when r is varied.

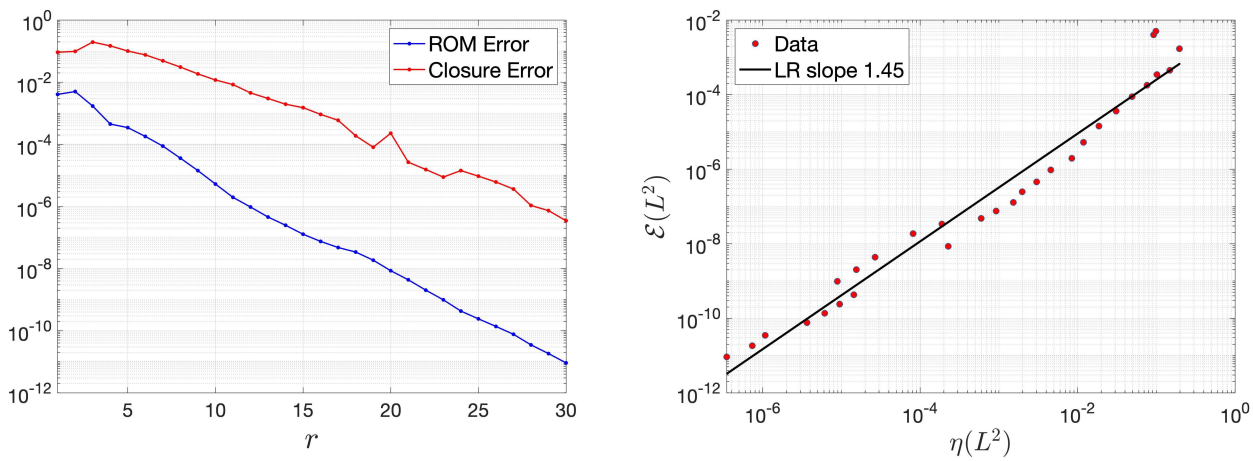

FIGURE 3. Burgers equation (6.8), reconstructive regime: $\mathcal{E}(L^2)$ and $\eta(L^2)$ as r increases. For each r, the tolerance index k in the truncated SVD is chosen to minimize the corresponding ROM error $\mathcal{E}(L^2)$.

Overall, the results in this section provide strong numerical support to the theoretical understanding put forth in Theorem 5.2 in the Burgers equation setting.

6.4. **Flow Past A Cylinder.** In this section, we investigate the DD-VMS-ROM verifiability in the numerical simulation of a 2D channel flow past a circular cylinder at Reynolds numbers $Re = 100$ and $Re = 1000$. This test problem has been used in, e.g., [36, 37, 54].
Computational Setting. As a mathematical model, we use the NSE (2.1)–(2.2). The computational domain is a 2.2×0.41 rectangular channel with a cylinder of radius 0.05, centered at $(0.2, 0.2)$, see Figure 4.

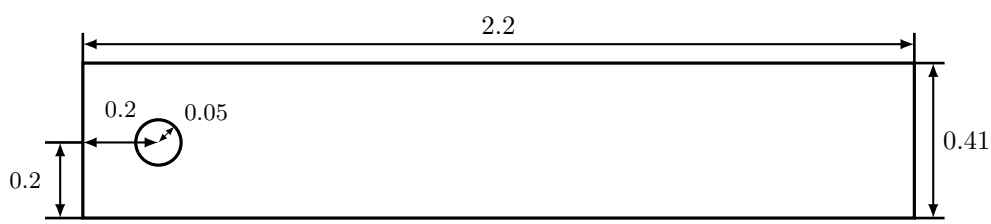

FIGURE 4. Geometry of the flow past a circular cylinder numerical experiment.

We prescribe no-slip boundary conditions on the walls and cylinder, and the following inflow and outflow profiles [28, 36, 43]:

$$u_1(0, y, t) = u_1(2.2, y, t) = \frac{6}{0.41^2} y(0.41 - y), \tag{6.10}$$

$$u_2(0, y, t) = u_2(2.2, y, t) = 0, \tag{6.11}$$

where $\boldsymbol{u} = \langle u_1, u_2 \rangle$. There is no forcing and the flow starts from rest.
Snapshot Generation. For the spatial discretization, we use the pointwise divergence-free, LBB stable (P_2, P_1^{disc}) Scott-Vogelius finite element pair on a barycenter refined regular triangular mesh [29]. The mesh yields $103K$ (102962) velocity and $76K$ (76725) pressure degrees of freedom. We use the linearized BDF2 temporal discretization and a time step size $\Delta t = 0.002$ for both FOM and ROM time discretizations. On the first time step, we use a backward Euler scheme so that we have the two initial time step solutions required for the BDF2 scheme.

ROM Construction. The FOM simulations settle down to periodic dynamics at different time instances for the two Reynolds numbers used in the numerical investigation: For $Re = 100$ after $t = 5$, and for $Re = 1000$ after $t = 13$. To construct the ROM basis functions, we use 10 time units of FOM data. Thus, to ensure a fair comparison of the numerical results at different Reynolds numbers, we collect FOM snapshots on the following time intervals: For $Re = 100$ from $t = 7$ to $t = 17$, and for $Re = 1000$ from $t = 13$ to $t = 23$.

To train the DD-VMS-ROM closure operator $\tilde{A}$, we use FOM data for one period. The period length of the FOM dynamics is different for the two different Reynolds numbers: From $t = 7$ to $t = 7.332$ for $Re = 100$, and from $t = 13$ to $t = 13.268$ for $Re = 1000$. Thus, we collect 167 snapshots for $Re = 100$, and 135 snapshots for $Re = 1000$.

6.4.1. *Numerical Results for $Re = 100$.* In Figure 5, for three different r values, we plot $\mathcal{E}(L^2)$ in (6.5), which measures the DD-VMS-ROM error, and $\eta(L^2)$ in (6.6), which measures the DD-VMS-ROM closure error. To compute $\mathcal{E}(L^2)$ and $\eta(L^2)$, we fix the r value and decrease the tolerance index k in the truncated SVD, which is used in the data-driven modeling part. As the tolerance decreases, we monitor the decaying rate of $\mathcal{E}(L^2)$ with respect to $\eta(L^2)$. The results in Figure 5, for $r = 4, 6$, and 8, generally show that, as $\eta(L^2)$ decreases, so does $\mathcal{E}(L^2)$. We note that, in each panel, the minimal $\mathcal{E}(L^2)$ value is actually achieved at $k = r^2$, i.e., when the full SVD is used in constructing the closure term $\widetilde{A}$. This is due to the fact that for all the r values considered, the condition number of the corresponding data matrix $E^\top E$ is always below 10^3. The same observation is true for the $Re = 1000$ test case presented in Section 6.4.2.

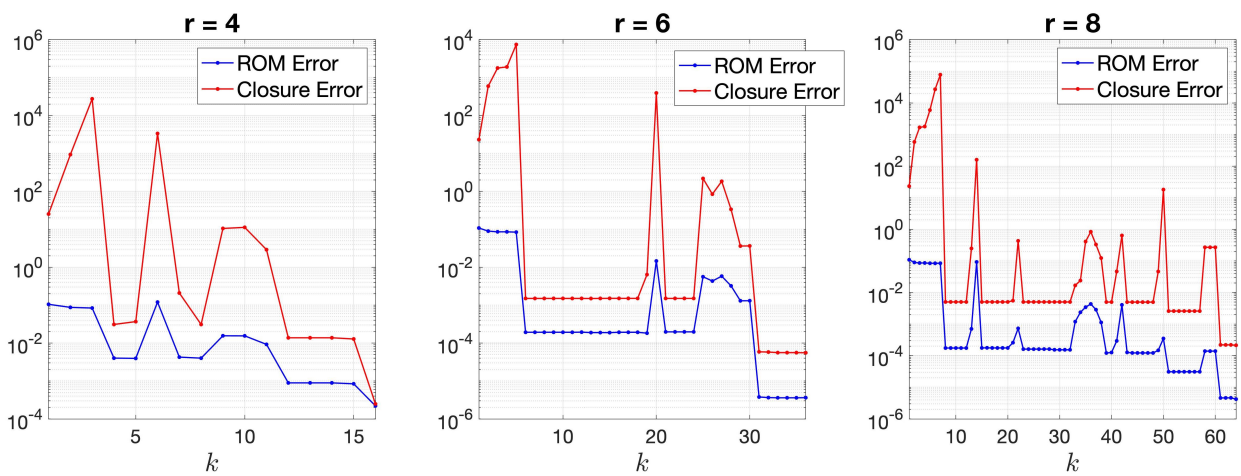

FIGURE 5. Flow past a cylinder, $Re = 100$, reconstructive regime: $\mathcal{E}(L^2)$ and $\eta(L^2)$ values for fixed r values and different tolerance index k values in the truncated SVD.

In Figure 6, for $r = 4, 6$, and 8, we plot the LR slope for $\mathcal{E}(L^2)$ with respect to $\eta(L^2)$. For $r = 4$, the LR slope is 0.56, for $r = 6$ the LR slope is 0.99, and for $r = 8$ the LR slope is 1.03. These results indicate an almost linear correlation between $\mathcal{E}(L^2)$ and $\eta(L^2)$, again in agreement with (6.7) with $\alpha = 1$, except for $r = 4$. One possible explanation for the $r = 4$ case is that, due to the low-dimensionality of the ROM, we do not have sufficient data points to accurately estimate the LR slope.

When we vary r and choose the tolerance index k in the truncated SVD to minimize the corresponding $\mathcal{E}(L^2)$, the results are shown in Figure 7. This figure shows again that (6.7) holds with $\alpha = 1$ in the varying r setting.

Overall, the results in Figures 5–7 support the theoretical results in Theorem 5.2.

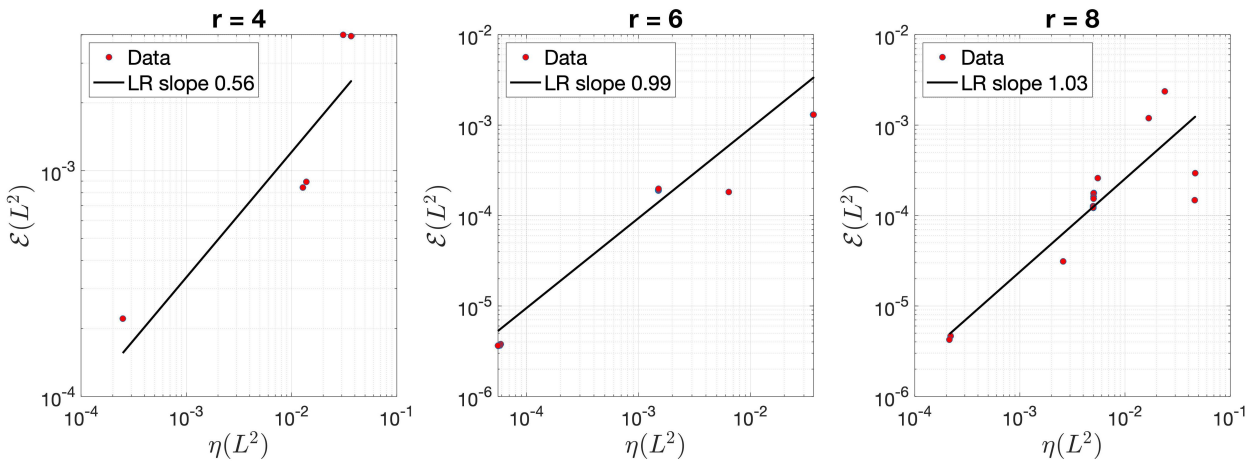

FIGURE 6. Flow past a cylinder, $Re = 100$, reconstructive regime: linear regression for $\mathcal{E}(L^2)$ and $\eta(L^2)$ for fixed r values and different tolerance values in the truncated SVD.

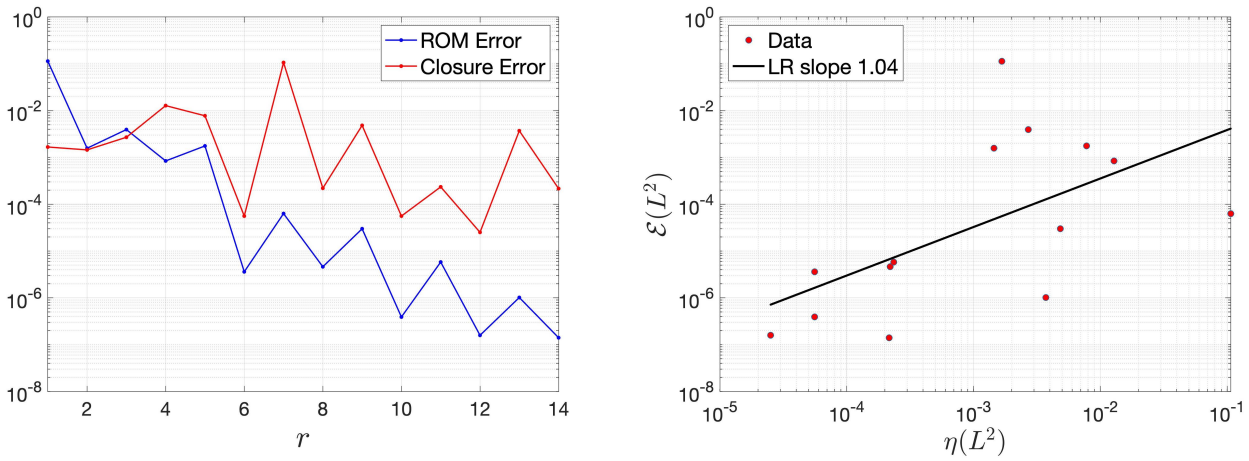

FIGURE 7. Flow past a cylinder, $Re = 100$, reconstructive regime: $\mathcal{E}(L^2)$ and $\eta(L^2)$ values as r increases. For each r, the tolerance index k in the truncated SVD is chosen to minimize the corresponding ROM error $\mathcal{E}(L^2)$.

6.4.2. *Numerical Results for $Re = 1000$.* In Figure 8, for three different r values, we plot $\mathcal{E}(L^2)$ in (6.5), which measures the DD-VMS-ROM error, and $\eta(L^2)$ in (6.6), which measures the DD-VMS-ROM closure error. To compute $\mathcal{E}(L^2)$ and $\eta(L^2)$, we fix the r value and decrease the tolerance in the truncated SVD, which is used in the data-driven modeling part. As the tolerance decreases, we monitor the decaying rate of $\mathcal{E}(L^2)$ with respect to $\eta(L^2)$. The results in Figure 8, for $r = 4, 6$, and 8, generally show that, as $\eta(L^2)$ decreases, so does $\mathcal{E}(L^2)$.

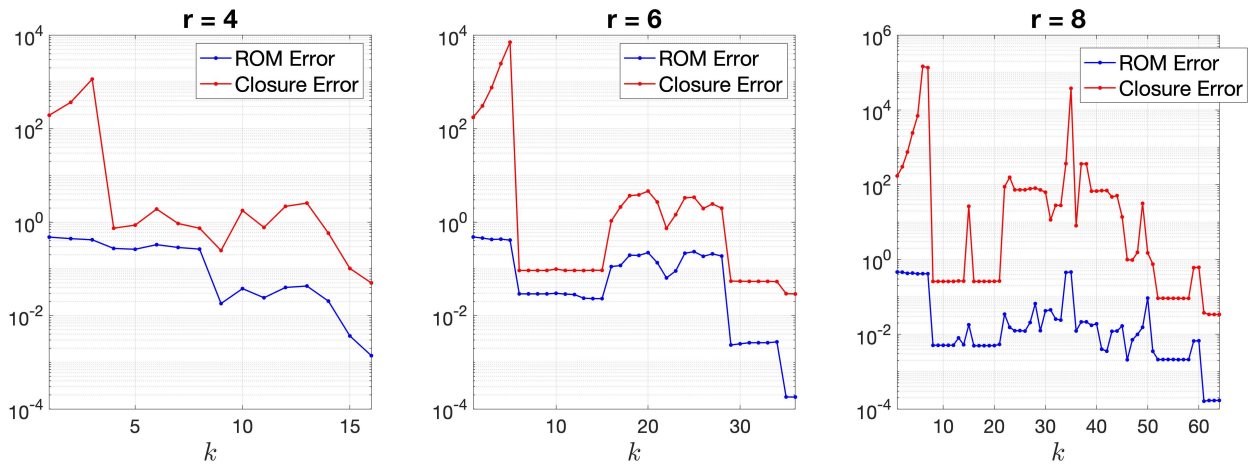

FIGURE 8. Flow past a cylinder, $Re = 1000$, reconstructive regime: $\mathcal{E}(L^2)$ and $\eta(L^2)$ values for fixed r values and different tolerance index k values in the truncated SVD.

In Figure 9, for $r = 4, 6$, and 8, we plot the LR slope for $\mathcal{E}(L^2)$ with respect to $\eta(L^2)$. For $r = 4$, the LR slope is 1.71, for $r = 6$ the LR slope is 2.07, and for $r = 8$ the LR slope is 1.00. These results indicate that $\mathcal{E}(L^2)$ decays at least linearly as $\eta(L^2)$ is reduced, again in agreement with (6.7) with $\alpha \geq 1$.

When we vary r and choose the tolerance index k in the truncated SVD to minimize the corresponding $\mathcal{E}(L^2)$, the results are shown in Figure 10. This figure shows again that (6.7) holds with $\alpha \geq 1$ in this varying r setting.

Overall, the results in Figures 8–10 support the theoretical results in Theorem 5.2, yielding the same conclusion as that in Section 6.4.1.

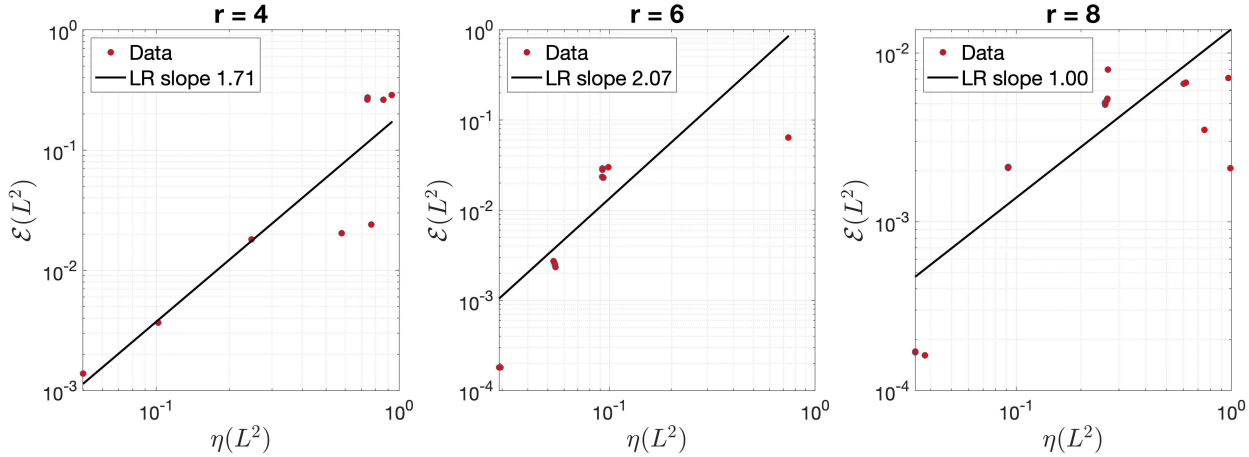

FIGURE 9. Flow past a cylinder, $Re = 1000$, reconstructive regime: linear regression for $\mathcal{E}(L^2)$ and $\eta(L^2)$ for fixed r values and different tolerance values in the truncated SVD.

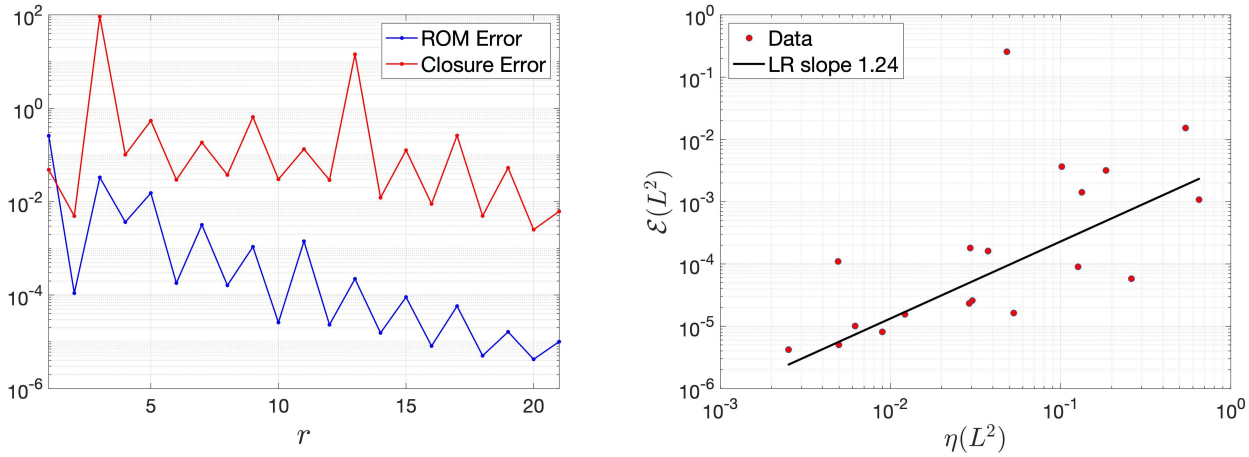

FIGURE 10. Flow past a cylinder, $Re = 1000$, reconstructive regime: $\mathcal{E}(L^2)$ and $\eta(L^2)$ values as r increases. For each r, the tolerance index k in the truncated SVD is chosen to minimize the corresponding ROM error $\mathcal{E}(L^2)$.

7. Conclusions and Future Work

Over the last two decades, a plethora of ROM closure models have been developed for reduced order modeling of convection-dominated flows. Various ROM closure models have been constructed by using physical insight, mathematical arguments, or data. Although these ROM closure models are built by using different arguments, they are constructed by using the same *heuristic* algorithm: (i) In the offline stage, the ROM closure model is built so that it is as close as possible (in some norm) to the "true" ROM closure term. (ii) In the online stage, one needs to check whether the ROM closure model yields a ROM solution that is as close as possible to the filtered FOM solution. If the ROM solution is an accurate approximation of the filtered FOM solution, the ROM closure model is deemed accurate. This heuristic algorithm is the most popular approach used in assessing the success of the current ROM closure models. However, a natural question is whether one can actually *prove* anything about these ROM closure models. For example, can one prove that an accurate ROM closure model (constructed in the offline phase) yields an accurate ROM solution (in the online phase)?

In this paper, we took a step in this direction and answered the above question by extending the verifiability concept from classical LES to a ROM setting. Specifically, we defined a ROM closure model as verifiable if the ROM error is bounded (in some norm) by the ROM closure model error. Furthermore, we proved that a recently introduced data-driven ROM closure model (i.e., the DD-VMS-ROM [37, 54]) is verifiable. Finally, we showed numerically that the DD-VMS-ROM closure is verifiable. Specifically, in the numerical simulation of the one-dimensional Burgers equation and the two-dimensional flow past a circular cylinder at Reynolds numbers $Re = 100$ and $Re = 1000$, we showed that by reducing the error in the ROM closure term, we can achieve a decrease in the ROM error, as predicted by the theoretical results.

There are several natural research directions that can be pursued in the quest to lay mathematical foundations for ROM closure models. For example, one could investigate the verifiability of (functional, structural, or data-driven) ROM closure models that are different from the DD-VMS-ROM investigated in this paper. One could also extend the verifiability concept to ROM closures that are built from experimental data. In that case, one could replace the high-dimensional "truth" solution used in this paper with the experimental solution interpolated onto a discrete mesh. Another potential research direction is the investigation of different norms (e.g., the H^1 norm) in the least squares problem (4.2), verifiability definition (i.e., Definition 5.2), and verfiability theorem (i.e., Theorem 5.2). Finally, one could consider other mathematical concepts that are used in classical LES (see, e.g., [9]) and extend them to a ROM setting.

Acknowledgements

We thank the reviewers for the insightful comments and suggestions, which have significantly improved the paper. The work of the first, second, and sixth authors was supported by NSF through grants DMS-2012253 and CDS&E-MSS-1953113. The third author acknowledges the support by NSF through grant DMS-2108856. The fifth author acknowledges the support by European Union Funding for Research and Innovation – Horizon 2020 Program – in the framework of European Research Council Executive Agency: Consolidator Grant H2020 ERC CoG 2015 AROMA-CFD project 681447 "Advanced Reduced Order Methods with Applications in Computational Fluid Dynamics," the PRIN 2017 "Numerical Analysis for Full and Reduced Order Methods for the efficient and accurate solution of complex systems governed by Partial Differential Equations" (NA-FROM-PDEs), and the INDAM-GNCS project "Tecniche Numeriche Avanzate per Applicazioni Industriali."

Data Availability

The datasets generated during the current study are available from the corresponding author on reasonable request.

References

[1] Ahmed, M., San, O.: Stabilized principal interval decomposition method for model reduction of nonlinear convective systems with moving shocks. Comp. Appl. Math. **37**(5), 6870–6902 (2018)

[2] Ahmed, S.E., Pawar, S., San, O., Rasheed, A., Iliescu, T., Noack, B.R.: On closures for reduced order models − a spectrum of first-principle to machine-learned avenues. arXiv preprint `http://arxiv.org/abs/2106.14954` (2021)

[3] Ainsworth, M., Oden, J.T.: A Posteriori Error Estimation in Finite Element Analysis, vol. 37. John Wiley & Sons (2000)

[4] Ali, S., Ballarin, F., Rozza, G.: Stabilized reduced basis methods for parametrized steady Stokes and Navier–Stokes equations. Comput. Math. Appl. **80**(11), 2399–2416 (2020)

[5] Azaïez, M., Rebollo, T.C., Rubino, S.: A cure for instabilities due to advection-dominance in POD solution to advection-diffusion-reaction equations. J. Comput. Phys. **425**, 109916 (2021)

[6] Ballarin, F., Manzoni, A., Quarteroni, A., Rozza, G.: Supremizer stabilization of POD–Galerkin approximation of parametrized steady incompressible Navier–Stokes equations. Int. J. Numer. Meth. Engng. **102**, 1136–1161 (2015)

[7] Ballarin, F., Rebollo, T.C., Ávila, E.D., Mármol, M.G., Rozza, G.: Certified reduced basis vms-smagorinsky model for natural convection flow in a cavity with variable height. Computers & Mathematics with Applications **80**(5), 973–989 (2020)

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

[9] Berselli, L.C., Iliescu, T., Layton, W.J.: Mathematics of Large Eddy Simulation of Turbulent Flows. Scientific Computation. Springer-Verlag, Berlin (2006)

[10] Borggaard, J., Iliescu, T., Wang, Z.: Artificial viscosity proper orthogonal decomposition. Math. Comput. Modelling **53**(1-2), 269–279 (2011)

[11] Chekroun, M.D., Liu, H., McWilliams, J.C.: Variational approach to closure of nonlinear dynamical systems: Autonomous case. J. Stat. Phys. **179**, 1073–1160 (2020)

[12] Chekroun, M.D., Liu, H., McWilliams, J.C.: Stochastic rectification of fast oscillations on slow manifold closures. Proc. Natl. Acad. Sci. USA **118**, e2113650118 (2021)

[13] Chekroun, M.D., Liu, H., Wang, S.: Stochastic parameterizing manifolds and non-Markovian reduced equations: stochastic manifolds for nonlinear SPDEs II. Springer (2015)

[14] Chen, N., Liu, H., Lu, F.: Shock trace prediction by reduced models for a viscous stochastic Burgers equation. Chaos **32**(4), 043109 (2022)

[15] Chorin, A.J., Lu, F.: Discrete approach to stochastic parametrization and dimension reduction in nonlinear dynamics. Proc. Natl. Acad. Sci. USA **112**(32), 9804–9809 (2015)

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

[17] Girfoglio, M., Quaini, A., Rozza, G.: A POD-Galerkin reduced order model for a LES filtering approach. J. Comput. Phys. **436**, 110260 (2021)

[18] Girfoglio, M., Quaini, A., Rozza, G.: Pressure stabilization strategies for a LES filtering reduced order model. arXiv preprint,`http://arxiv.org/abs/2106.15887` (2021)

[19] Gunzburger, M., Iliescu, T., Schneier, M.: A Leray regularized ensemble-proper orthogonal decomposition method for parameterized convection-dominated flows. IMA J. Numer. Anal. **40**(2), 886–913 (2020)

[20] Hansen, P.C.: Discrete Inverse Problems: Insight and Algorithms, vol. 7. Society for Industrial and Applied Mathematics (2010)

[21] Hess, M.W., Quaini, A., Rozza, G.: Reduced basis model order reduction for Navier–Stokes equations in domains with walls of varying curvature. Int. J. Comput. Fluid Dyn. **34**(2), 119–126 (2020)

[22] Hesthaven, J.S., Rozza, G., Stamm, B.: Certified Reduced Basis Methods for Parametrized Partial Differential Equations. Springer (2015)
[23] Holmes, P., Lumley, J.L., Berkooz, G.: Turbulence, Coherent Structures, Dynamical Systems and Symmetry. Cambridge (1996)
[24] Iliescu, T., Liu, H., Xie, X.: Regularized reduced order models for a stochastic burgers equation. International Journal of Numerical Analysis & Modeling **15**, 594–607 (2018)
[25] Iliescu, T., Wang, Z.: Variational multiscale proper orthogonal decomposition: Convection-dominated convection-diffusion-reaction equations. Math. Comput. **82**(283), 1357–1378 (2013)
[26] Iliescu, T., Wang, Z.: Variational multiscale proper orthogonal decomposition: Navier-Stokes equations. Num. Meth. P.D.E.s **30**(2), 641–663 (2014)
[27] John, V.: Large Eddy Simulation of Turbulent Incompressible Flows, *Lecture Notes in Computational Science and Engineering*, vol. 34. Springer-Verlag, Berlin (2004). Analytical and Numerical Results for a Class of LES Models
[28] John, V.: Reference values for drag and lift of a two dimensional time-dependent flow around a cylinder. Int. J. Num. Meth. Fluids **44**, 777–788 (2004)
[29] John, V., Linke, A., Merdon, C., Neilan, M., Rebholz, L.G.: On the divergence constraint in mixed finite element methods for incompressible flows. SIAM Rev. (2016)
[30] Kaya, M., Layton, W., et al.: On" verifiability" of models of the motion of large eddies in turbulent flows. Differential and Integral Equations **15**(11), 1395–1407 (2002)
[31] Koc, B., Mohebujjaman, M., Mou, C., Iliescu, T.: Commutation error in reduced order modeling of fluid flows. Adv. Comput. Math. **45**(5-6), 2587–2621 (2019)
[32] Kunisch, K., Volkwein, S.: Galerkin proper orthogonal decomposition methods for parabolic problems. Numer. Math. **90**(1), 117–148 (2001)
[33] Layton, W.J.: Introduction to the numerical analysis of incompressible viscous flows, vol. 6. Society for Industrial and Applied Mathematics (SIAM) (2008)
[34] Lu, F.: Data-driven model reduction for stochastic Burgers equations. Entropy **22**(12), 1360 (2020)
[35] Martini, I., Haasdonk, B., Rozza, G.: Certified reduced basis approximation for the coupling of viscous and inviscid parametrized flow models. J. Sci. Comput. **74**(1), 197–219 (2018)
[36] Mohebujjaman, M., Rebholz, L.G., Iliescu, T.: Physically-constrained data-driven correction for reduced order modeling of fluid flows. Int. J. Num. Meth. Fluids **89**(3), 103–122 (2019)
[37] Mou, C., Koc, B., San, O., Rebholz, L.G., Iliescu, T.: Data-driven variational multiscale reduced order models. Comput. Methods Appl. Mech. Engrg. **373**, 113470 (2021)
[38] Mou, C., Liu, H., Wells, D.R., Iliescu, T.: Data-driven correction reduced order models for the quasi-geostrophic equations: A numerical investigation. Int. J. Comput. Fluid Dyn. **34**, 147–159 (2020)
[39] Oberai, A.A., Jagalur-Mohan, J.: Approximate optimal projection for reduced-order models. Int. J. Num. Meth. Engng. **105**(1), 63–80 (2016)
[40] Parish, E.J., Duraisamy, K.: A unified framework for multiscale modeling using the Mori-Zwanzig formalism and the variational multiscale method. arXiv preprint, `http://arxiv.org/abs/1712.09669` (2017)
[41] Peherstorfer, B., Willcox, K.: Data-driven operator inference for nonintrusive projection-based model reduction. Comput. Methods Appl. Mech. Engrg. **306**, 196–215 (2016)
[42] Quarteroni, A., Manzoni, A., Negri, F.: Reduced Basis Methods for Partial Differential Equations: An Introduction, vol. 92. Springer (2015)
[43] Rebholz, L., Xiao, M.: Improved accuracy in algebraic splitting methods for Navier-Stokes equations. SIAM J. Sci. Comput. **39**(4), A1489–A1513 (2017)
[44] Rebollo, T.C., Ávila, E.D., Mármol, M.G., Ballarin, F., Rozza, G.: On a certified Smagorinsky reduced basis turbulence model. SIAM J. Numer. Anal. **55**(6), 3047–3067 (2017)
[45] Rebollo, T.C., Lewandowski, R.: Mathematical and Numerical Foundations of Turbulence Models and Applications. Springer (2014)
[46] Reyes, R., Codina, R.: Projection-based reduced order models for flow problems: A variational multiscale approach. Comput. Methods Appl. Mech. Engrg. **363**, 112844 (2020)
[47] Sagaut, P.: Large Eddy Simulation for Incompressible Flows, third edn. Scientific Computation. Springer-Verlag, Berlin (2006)
[48] Sell, G.R., You, Y.: Dynamics of evolutionary equations, vol. 143. Springer Science & Business Media (2013)
[49] Stabile, G., Ballarin, F., Zuccarino, G., Rozza, G.: A reduced order variational multiscale approach for turbulent flows. Adv. Comput. Math. pp. 1–20 (2019)
[50] Temam, R.: Navier-Stokes equations: Theory and numerical analysis, vol. 2. American Mathematical Society (2001)
[51] Thomée, V.: Galerkin finite element methods for parabolic problems. Springer Verlag (2006)
[52] Volkwein, S.: Proper orthogonal decomposition: Theory and reduced-order modelling. Lecture Notes, University of Konstanz (2013). `http://www.math.uni-konstanz.de/numerik/personen/volkwein/teaching/POD-Book.pdf`
[53] Wang, Z., Akhtar, I., Borggaard, J., Iliescu, T.: Proper orthogonal decomposition closure models for turbulent flows: A numerical comparison. Comput. Meth. Appl. Mech. Eng. **237-240**, 10–26 (2012)
[54] Xie, X., Mohebujjaman, M., Rebholz, L.G., Iliescu, T.: Data-driven filtered reduced order modeling of fluid flows. SIAM J. Sci. Comput. **40**(3), B834–B857 (2018)

[55] Xie, X., Webster, C., Iliescu, T.: Closure learning for nonlinear model reduction using deep residual neural network. Fluids **5**(1), 39 (2020)
[56] Xie, X., Wells, D., Wang, Z., Iliescu, T.: Approximate deconvolution reduced order modeling. Comput. Methods Appl. Mech. Engrg. **313**, 512–534 (2017)
[57] Xie, X., Wells, D., Wang, Z., Iliescu, T.: Numerical analysis of the Leray reduced order model. J. Comput. Appl. Math. **328**, 12–29 (2018)
[58] Yıldız, S., Goyal, P., Benner, P., Karasozen, B.: Data-driven learning of reduced-order dynamics for a parametrized shallow water equation. arXiv preprint, `http://arxiv.org/abs/arXiv:2007.14079` (2020)