Residual Data Driven Variational Multiscale Reduced Order Models for Convection-Dominated Problems in the Predictive Regime

Birgul Koc[1*], Samuele Rubino[2], Traian Iliescu[3], Tomás Chacón[2]

[1]Dpto. EDAN, Universidad de Sevilla, C/Tarfia, s/n. 41012 Sevilla, Spain; *bkoc@us.es

[2]Dpto. EDAN & IMUS, Universidad de Sevilla, C/Tarfia, s/n. 41012 Sevilla, Spain

[3]Virginia Tech, Blacksburg, VA 24061, United States

SNE 36(2), 2026, 95-98, DOI: 10.11128/sne.36.sn.10775
Selected MATHMOD 2025 Postconf. Publication: 2025-02-26
Rec. Revised Extended: 2026-05-13; Accepted: 2026-05-20
SNE - Simulation Notes Europe, ARGESIM Publisher Vienna
ISSN Print 2305-9974, Online 2306-0271, www.sne-journal.org

Abstract. In this work, we propose a novel residual based data driven closure strategy for reduced-order models (ROMs) of under-resolved, convection-dominated flows. The proposed closure model is developed within a variational multiscale (VMS) framework, leveraging available full order model (FOM) data and an ansatz that explicitly depends on the ROM residual. We emphasize that this approach is fundamentally different from existing data driven ROM closure models, which typically rely on the ROM coefficients as inputs. In contrast, the proposed residual based method utilizes the ROM residual to account for the effects of unresolved scales. We assess the performance of the proposed residual-based data-driven VMS-ROM in the numerical simulation of two-dimensional flow past a cylinder at Reynolds number Re $= 1000$. The results demonstrate that the proposed method yields significantly improved accuracy compared to standard coefficient based data driven VMS-ROM approaches.

Introduction

We consider the Navier-Stokes equations (NSE) (1)-(2) as the mathematical model:

$$\frac{\partial u}{\partial t} - Re^{-1}\Delta u + u\cdot\nabla u + \nabla p = 0, \tag{1}$$

$$\nabla\cdot u = 0, \tag{2}$$

where u denotes the velocity field, p the pressure, and Re the Reynolds number. Homogeneous Dirichlet boundary conditions are imposed.

To construct reduced order models (ROMs), we use the proper orthogonal decomposition (POD) to generate the ROM basis functions and associated operators. Owing to the orthogonality of the POD modes, the ROM space can be decomposed into large-scale and sub-scale components:

$$X^d = X^L \oplus X^S, \tag{3}$$

where

$$X^d := \text{span}\{\varphi_1,\ldots,\varphi_d\},$$
$$X^L := \text{span}\{\varphi_1,\ldots,\varphi_L\},$$
$$X^S := \text{span}\{\varphi_{L+1},\ldots,\varphi_d\}.$$

Using all d modes, the ROM approximation

$$u_d = \sum_{j=1}^{d} (a_d)_j \varphi_j \tag{4}$$

provides the most accurate representation of the full order model (FOM) solution in the POD sense.

For laminar flows, a low dimensional approximation u_L, with $L \ll d$, is typically sufficient to accurately represent the FOM solution. In this regime, the standard Galerkin ROM (G-ROM) is given by

$$\dot{a}_L = A_{LL}\, a_L + a_L^\top B_{LLL}\, a_L, \tag{5}$$

where $(A_{LL})_{ij} = -\nu(\nabla\varphi_i, \nabla\varphi_j)$, $(B_{LLL})_{ijk} = -(\varphi_i, \varphi_j \cdot \nabla\varphi_k)$, for all $i,j,k = 1,\ldots,L$.

The G-ROM system (5) is derived by substituting u_L into the NSE (1)-(2) and projecting the resulting equations onto the large-scale ROM space X^L.

However, for turbulent flows, the low dimensional G-ROM solution u_L is generally insufficient to accurately approximate the FOM solution. To alleviate this inaccurate behavior, standard ROMs are typically augmented with numerical stabilization techniques or ROM closure models, which aim to account for the effects of unresolved scales [2, 3, 4, 5, 6].

1 ROM Closure Models

ROM closure modeling aims to approximate the closure term arising in a variational multiscale (VMS) framework [2, 3]. To construct this closure term, we first decompose the most accurate ROM solution u_d into large and sub-scale components:

$$u_L := \sum_{j=1}^{L} (a_L)_j \varphi_j, \qquad u_S := \sum_{j=L+1}^{d} (a_S)_j \varphi_j. \tag{6}$$

Next, we derive the governing equations for the large and sub-scales. To this end, we substitute $u = u_d = u_L + u_S$ into (1)–(2) and project the resulting system onto the ROM spaces X^L and X^S, respectively. This yields the following coupled system:

$$\begin{aligned} \dot{a}_L &= A_{LL}a_L + A_{LS}a_S + a_L^\top B_{LLL}a_L + a_L^\top B_{LLS}a_S \\ &\quad + a_S^\top B_{LSL}a_L + a_S^\top B_{LSS}a_S, \qquad (7a) \\ \dot{a}_S &= A_{SS}a_S + A_{SL}a_L + a_S^\top B_{SSS}a_S + a_S^\top B_{SSL}a_L \\ &\quad + a_L^\top B_{SLS}a_S + a_L^\top B_{SLL}a_L. \qquad (7b) \end{aligned}$$

Here, the matrices A_{IJ} and tensors B_{IJK}, with indices $I,J,K \in \{L,S\}$, are defined by

$$\begin{aligned} (A_{IJ})_{ij} &:= -Re^{-1}(\nabla\varphi_i^I, \nabla\varphi_j^J), \\ (B_{IJK})_{ijk} &:= -(\varphi_i^I, \varphi_j^J \cdot \nabla\varphi_k^K), \end{aligned}$$

where $\varphi^L := \{\varphi_1, \ldots, \varphi_L\}$ and $\varphi^S := \{\varphi_{L+1}, \ldots, \varphi_d\}$ denote the large-scale and sub-scale ROM basis functions, respectively. The subscript indices I,J,K indicate which ROM subspace each basis function belongs to, making the coupling structure between large and sub-scales explicit.

In this work, we consider two distinct ROM closure strategies, leading to two different models: the *coefficient based data driven variational multiscale* ROM (C-ROM) and the *residual based data driven variational multiscale* ROM (R-ROM).

2 Basis of ROM Closure Models

2.1 Coefficient based ROM (C-ROM)

The coefficient-based ROM, C-ROM, [2, 3] is derived from the large-scale equation (7a) by introducing a closure model and a corresponding ansatz. Since the **closure** term is not available in closed form, we approximate it using a quadratic, coefficient-based **ansatz** that depends on the large-scale coefficients a_L:

$$\begin{aligned} \textbf{closure} &= A_{LS}a_S + a_L^\top B_{LLS}a_S + a_S^\top B_{LSL}a_L + a_S^\top B_{LSS}a_S, \\ \textbf{ansatz} &= \widetilde{A}_{LL}\,a_L + a_L^\top \widetilde{B}_{LLL}\,a_L. \end{aligned} \tag{8}$$

2.2 Residual based ROM (R-ROM)

In the proposed residual-based ROM, R-ROM, the closure modeling is constructed using information from the sub-scale equation (7b). In this case, the **closure** term and the corresponding residual-based **ansatz** are defined as:

$$\begin{aligned} \textbf{closure} &= a_S, \\ \textbf{ansatz} &= \widetilde{A}_{SS}\,\textbf{Res}_S(a_L) + \textbf{Res}_S(a_L)^\top \widetilde{B}_{SSS}\,\textbf{Res}_S(a_L), \end{aligned} \tag{9}$$

where the sub-scale residual is given by

$$\textbf{Res}_S(a_L) := A_{SL}a_L + a_L^\top B_{SLL}a_L. \tag{10}$$

2.3 Minimization Problem: Optimal Data Driven Operators

To identify the unknown operators $\widetilde{A}_{LL}$, $\widetilde{A}_{SS}$, $\widetilde{B}_{LLL}$, and $\widetilde{B}_{SSS}$, we employ a *data-driven* (D2) approach [2, 3]. These operators are obtained by solving the following least-squares minimization problem:

$$\min_{\text{D2 operators}} \quad \sum_{j=1}^{M} \left\| \textbf{closure}(a_j^{\text{FOM}}) - \textbf{ansatz}(a_j^{\text{FOM}}) \right\|_{L^2}^2. \tag{11}$$

Using the respective closure terms and ansatz for the C-ROM and R-ROM, we solve (11) to obtain the corresponding D2 operators. Substituting the resulting ansatz into the large-scale equation (7a), the C-ROM reads

$$\dot{a}_L = \left(A_{LL} + \widetilde{A}_{LL}\right) a_L + a_L^\top \left(B_{LLL} + \widetilde{B}_{LLL}\right) a_L, \tag{12}$$

while the R-ROM is given by

$$\dot{a}_L = A_{LL}a_L + a_L^\top B_{LLL}a_L + A_{LS}\widetilde{a}_S + a_L^\top B_{LLS}\widetilde{a}_S + \widetilde{a}_S^\top B_{LSL}a_L + \widetilde{a}_S^\top B_{LSS}\widetilde{a}_S, \quad (13)$$

where the approximated sub-scale coefficients are computed as

$$\widetilde{a}_S := \widetilde{A}_{SS}\mathbf{Res}_S(a_L) + \mathbf{Res}_S(a_L)^\top \widetilde{B}_{SSS}\mathbf{Res}_S(a_L). \quad (14)$$

3 Numerical Results

We investigate the numerical accuracy of the G-ROM, C-ROM, and R-ROM for 2D channel flow past a circular cylinder at $\mathrm{Re} = 1000$ in the **predictive regime**: the ROM basis and operators are built from FOM snapshots over $t \in [13, 16]$, the data-driven operators are trained over $t \in [13, 13.134]$, and all ROMs are tested over $t \in [16, 23]$, with $d = 22$.

In addition to the ROM accuracy, we investigate the consistency between the closure term and the ansatz for both the C-ROM and R-ROM. To assess ROM performance, we examine the average L^2 projection errors, the evolution of the kinetic energy, the average L^2 kinetic energy errors, Pareto plots, and the vortex shedding frequency.

3.1 Consistency and ROM Projection Errors

We define the following metrics:

$$\mathscr{E}_{\mathrm{cons}} := \frac{1}{M}\sum_{k=1}^{M}\left\|\mathbf{closure}(a_{L,S}^{\mathrm{FOM},k}) - \mathbf{ansatz}(a_L^{\mathrm{FOM},k})\right\|_{L^2}, \quad (15)$$

$$\mathscr{E}_{\mathrm{proj}} := \frac{1}{M}\sum_{k=1}^{M}\left\|u_L(t_k) - \sum_{i=1}^{L}\left(u^{\mathrm{FOM}}(t_k), \varphi_i\right)_{L^2}\varphi_i\right\|_{L^2}, \quad (16)$$

where $\mathscr{E}_{\mathrm{cons}}$ measures the discrepancy between the closure term and its approximation, and $\mathscr{E}_{\mathrm{proj}}$ denotes the ROM projection error.

In Table 1, we list the average L^2 consistency error (15). The R-ROM consistently yields lower consistency errors than C-ROM, with an overall decaying trend despite non-monotone behavior due to the sensitivity of (11). In Table 2, we list the average L^2 ROM projection errors. Both C-ROM and R-ROM significantly outperform G-ROM, with R-ROM consistently achieving better accuracy overall.

L	C-ROM	R-ROM
2	2.28e-01	2.59e-02
5	1.49e+00	1.72e-02
8	4.01e-01	1.22e-04
11	1.35e+00	1.27e-08
14	2.46e-01	1.29e-02
20	1.43e-01	2.34e-03
22	0	0

Table 1: Average L^2 FOM consistency error (15) for C-ROM and R-ROM across different values of L.

L	G-ROM	C-ROM	R-ROM
2	1.15e+00	4.11e-01	3.59e-01
3	9.22e-01	5.51e-01	6.16e-02
4	7.21e-01	1.98e-01	1.18e-01
5	7.28e-01	5.81e-01	3.11e-01
6	3.54e-01	1.48e-01	4.36e-02
7	3.02e-01	2.81e-01	5.29e-02
8	1.59e-01	9.44e-02	2.53e-02

Table 2: Average L^2 ROM projection errors (16) for different values of L.

In Figure 1, we present a Pareto plot comparing C-ROM and R-ROM in terms of average L^2 error and offline ansatz computational cost, averaged over low-dimensional ($L = 2, 3, 4, 5$) and higher-dimensional ($L = 6, 7, 8$) ROMs.

For low-dimensional ROMs, R-ROM yields higher accuracy than C-ROM at a higher computational cost. For higher-dimensional ROMs, R-ROM is both more accurate and more efficient than C-ROM.

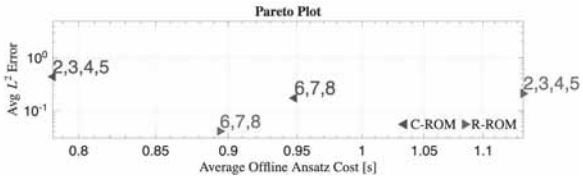

Figure 1: Pareto plot of average L^2 error of C-ROM and R-ROM.

3.2 Kinetic Energy

In this section, we use the kinetic energy (KE) criterion to compare the numerical accuracy of G-ROM, C-ROM, and R-ROM, where the KE is defined as follows:

$$E_{kin} := \frac{1}{2}\|u\|_{L^2}^2 = \frac{1}{2}\int_\Omega |u|^2 \, d\Omega. \tag{17}$$

In Figure 2, we compare the KE evolution of C-ROM and R-ROM for $L = 6$. R-ROM is significantly more accurate than C-ROM.

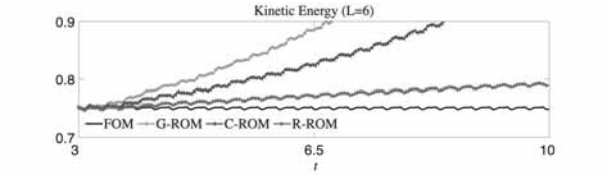

Figure 2: Time evolution of the kinetic energy for ROMs.

3.3 Vortex Shedding Frequency

In this section, we compute the average vortex shedding frequency f_s of C-ROM, R-ROM, and FOM based on the vortex shedding period $\mathscr{T}_s$, which is defined as follows:

$$\mathscr{T}_s = \frac{1}{N_s}\sum_{k=1}^{N_s}(t_s(k+1) - t_s(k)), \tag{18}$$

where $t_s(k)$ denotes successive KE peaks within $t \in [18, 23]$. In Figure 3, R-ROM more accurately recovers the FOM vortex shedding periods than C-ROM, and the KE peak amplitudes in C-ROM are noticeably higher than in R-ROM.

We observe that R-ROM more accurately recovers the vortex shedding periods of the FOM than C-ROM. Furthermore, the amplitudes of the KE peaks in C-ROM are noticeably higher than those in R-ROM.

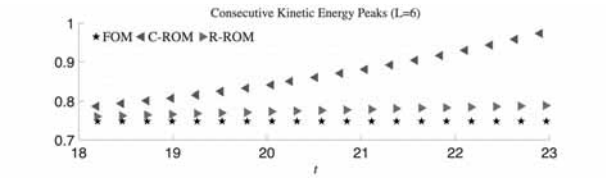

Figure 3: Comparison of vortex shedding frequency for FOM, C-ROM, and R-ROM.

Acknowledgement

The first and second authors are partially supported by Project PID2021-123153OB-C21 funded by MCIN / AEI / 10.13039 / 501100011033 / FEDER, UE, the third author is partially supported by Junta de Andalucía - Spanish Government - European Union grant SOL2024-31708, and the fourth author is partially supported by National Science Foundation grant DMS-2012253.

References

[1] Koc B, Rubino S, Chacón T, Iliescu T. Residual Data-Driven Variational Multiscale Reduced Order Models for Convection-Dominated Problems. In A. Körner & al. *MATHMOD 2025 Short Contribution Volume*, reposiTUm TU Wien, p. 11-12 DOI 10.34726/9005.

[2] Koc B, Rubino S, Chacón Rebollo T, Iliescu, T. Residual-based data-driven variational multiscale reduced order models for parameter-dependent problems. *Computational and Applied Mathematics*, 2025, 44(6), 308.

[3] Mou C, Koc B, San O, Rebholz LG, Iliescu T. Data-driven variational multiscale reduced order models. *Computer Methods in Applied Mechanics and Engineering*, 2021, 373, 113470.

[4] Azaïez M, Rebollo TC, Rubino S. A cure for instabilities due to advection-dominance in POD solution to advection-diffusion-reaction equations. *Journal of Computational Physics*, 2021, 425, 109916.

[5] Rebollo TC, Rubino S, Oulghelou M, Allery C. Error analysis of a residual-based stabilization-motivated POD-ROM for incompressible flows. *Computer Methods in Applied Mechanics and Engineering*, 2022, 401, 115627.

[6] Reyes R, Codina R. Projection-based reduced order models for flow problems: A variational multiscale approach. *Computer Methods in Applied Mechanics and Engineering*, 2020, 363, 112844.

Publication Remark. This work is an improved version of our MATHMOD 2025 conference contribution [1], closely related to [2]. Compared to [1], we include three additional results: (i) a consistency analysis of the ansatz operators; (ii) a Pareto plot comparing accuracy and computational cost; and (iii) a vortex shedding frequency analysis. Note that, unlike [2], absolute errors are reported throughout. For a comprehensive treatment including the R2-ROM variant and additional test cases, we refeer to [2].