Article

Non-intrusive inference reduced order model for fluids using deep multistep neural network

Xuping Xie [1,‡]*, Guannan Zhang [1,‡] and Clayton G. Webster [1,2,‡]

[1]   Computation and Applied Mathematics, Oak Ridge National Laboratory;
       xiex@ornlg.gov; zhangg@ornl.gov;
[2]   Department of Mathematics; University of Tennessee, Knoxville; cwebst13@utk.edu
*   Correspondence: xiex@ornl.gov
‡   These authors contributed equally to this work.

Abstract: In this effort we propose a data-driven learning framework for reduced order modeling of fluid dynamics. Designing accurate and efficient reduced order models for nonlinear fluid dynamic problems is challenging for many practical engineering applications. Classical projection-based model reduction methods generate reduced systems by projecting full-order differential operators into low-dimensional subspaces. However, these techniques usually lead to severe instabilities in the presence of highly nonlinear dynamics, which dramatically deteriorates the accuracy of the reduced-order models. In contrast, our new framework exploits linear multistep networks, based on implicit Adams-Moulton schemes, to construct the reduced system. The advantage is that the method optimally approximates the full order model in the low-dimensional space with a given supervised learning task. Moreover, our approach is non-intrusive, such that it can be applied to other complex nonlinear dynamical systems with sophisticated legacy codes. We demonstrate the performance of our method through the numerical simulation of a two-dimensional flow past a circular cylinder with Reynolds number Re = 100. The results reveal that the new data-driven model is significantly more accurate than standard projection-based approaches.

Keywords: Reduced-Order Model; Fluid Dynamics; Neural Network; Multistep Method; Optimization

---

1. Introduction

The full order model (FOM) of realistic engineering applications in fluid dynamics often represents a large scale dynamic system. High-fidelity CFD (computational fluid dynamics) simulations of the FOM are so computationally expensive that they put a heavy burden on the computational resources despite the available CFD software and supercomputers with thousands of cores. Consequently, the use of FOM for such simulations is often impractical and prohibitive for time-critical applications such as system identification, flow control, design optimization.

The reduced order modeling in fluid dynamics is to construct an accurate low-dimensional approximation to the full system with orders of magnitude reduction in computational cost. The first usage of reduced order model (ROM) in fluid dynamics was by [1] for studying the intensity of turbulence and coherent structures. Many recent successful applications of ROM in fluid problems can be found in [2–9]

The Galerkin projection based reduced order model (GP-ROM) is one of the most popular methods that has been widely used in practice. The GP-ROM, in an offline stage, first constructs a reduced space and then uses the Galerkin projection of the FOM operator to obtain a low-dimensional nonlinear dynamical system, i.e., ROM dynamics. The reduced space is often generated by proper orthogonal decomposition (POD), also known as principal component analysis. In an online stage, the obtained reduced dynamics can be used to approximate the full system efficiently for various applications, such as long-term prediction, flow control. However, the projection step requires that the full model

operators have to be available in order to obtain the ROM dynamics. This limits the applicability of projection based model reduction in situations where the full model is unknown [10,11]. More importantly, the computational cost of assembling the reduced operators – tensors from the projection of FOM operator – scales with the large dimension of the underlying high-dimensional of FOM. For this reason, the GP-ROM are efficient for problems where the reduced operators must be constructed only once.

On the other hand, the GP-ROM generates inaccurate approximations for highly non-stationary (nonlinear) fluids, e.g., turbulence. In the literature, a common explanation for this failure is that the Galerkin projection does not preserve the stability properties from the full model. A deeper reason is that the low-dimensional space used in the Galerkin projection cannot resolve the nonlinear interaction of the fluid system [12,13]. Consequently, resulting in a projection based stability error which makes GP-ROM fail in nonlinear fluid applications, e.g., see [14–17]

1.1. Related Work

Closure modeling. Numerous stabilization strategies have been devised to address the instability problem, known as closure modeling. The fundamental idea of closure modeling is to model the lost information from the low-dimensional space since it is generated through POD truncation. This truncation is to keep the first few POD modes that extract the most dominant structure of the full system and discard the rest modes. Closure models generally can be categorized into two common approaches. One is physically model the effect of the discarded POD mode by adding artificial viscosity to the reduced system [6,14,15,18]. Another approach is to mathematically model the ROM dynamics by solving a related optimization problem [16,17]. The most recent development for closure modeling is to apply the neural network to approximate the lost information [19].

Differential equations learning. Bridging numerical differential equations and deep neural networks have gained enormous attention recently. Specifically, [20] proposed dynamical system viewing of residual networks (ResNet). [21] the first time introduced the linear multistep network architecture to analyze ResNet on classification task. Sparse regression, Gaussian process, multistep neural networks have been applied for dynamic system learning [22–25]. More recently, ordinary differential equation network (ODE-net) was introduced for supervised learning [26].

1.2. Our approach

In this paper, we propose a novel non-intrusive learning reduced order model framework for fluid dynamic system. The new framework provides a general concept of learning the optimal reduced dynamic system from the data. Inspired by the successful development in learning differential equations with deep networks, we apply the linear multistep neural network (LMNet) to learn the reduced order model (LMNet-ROM). Unlike closure modeling and the existed non-intrusive model reduction methods, in this work, we focus on a different perspective. First, we do not use Galerkin projection and model the closure problem. The optimal reduced dynamic system that can address the instability issue is learned for a given supervised learning task. Secondly, the new model does not approximate reduced operators whereas other non-intrusive models use interpolation or regression method to infer reduced operators [10,11]. Moreover, our viewpoint easily enables us to answer the common question – what is the best ROM dynamic system to approximate the full system for a given low-dimensional subspace? We demonstrate the performance of the new LMNet-ROM is better than GP-ROM in full order model approximation and long-term prediction. The main contributions can be summarized as follows:

- A novel non-intrusive learning reduced order model framework for fluid dynamics, which is applicable to general nonlinear dynamical systems with sophisticated legacy codes.
- Our framework overcomes the instability issue of the projection based model reduction, and provides accurate approximation and long-term prediction of the full system.

- The learning process of our approach is more computationally efficient than the construction of reduced operators in the classic projection based methods.

2. Reduced Order Modeling

In this section, we present the Galerkin projection based reduced order modeling framework in fluid dynamic system. The classical Navier-Stokes equations (NSE) are often used as a mathematical model in fluid dynamics,

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

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

where $\boldsymbol{u}$ is the velocity, p the pressure, and Re the Reynolds number. We use the initial condition $\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{u}_0(\boldsymbol{x})$ and (for simplicity) homogeneous Dirichlet boundary conditions: $\boldsymbol{u}(\boldsymbol{x}, t) = \mathbf{0}$. For convenience, we use $\frac{\partial \boldsymbol{u}}{\partial t} = \mathbf{f}(\boldsymbol{u}, Re)$ as the general notation for the NSE in the rest of the paper.

Reduced Space. Proper orthogonal decomposition (POD) is the dimension reduction method that we used to generate the reduced space. It starts with the data matrix, $\boldsymbol{U} = [\boldsymbol{u}^0, \boldsymbol{u}^1, \cdots, \boldsymbol{u}^s] \in \mathcal{R}^{D_h \times (s+1)}$, collected by the numerical solutions or experimental observations of the full system (1) at $s + 1$ different time instances. The POD method seeks a low-dimensional space X^r that approximates the data ($\boldsymbol{U}$) optimally with respect to the L^2-norm. It formulates the following eigenvalue problem:

$$\boldsymbol{U}\boldsymbol{U}^T \varphi_i = \lambda_i \varphi_i, \ \ i = 1, 2, ..., d \tag{3}$$

where d is the rank of the data matrix $\boldsymbol{U}\boldsymbol{U}^T$. λ_i and φ_i are the eigenvalues and POD basis, respectively. The reduced space is given after the truncation as $\Phi_r := \{\varphi_1, \cdots, \varphi_r\} \in \mathcal{R}^{D_h \times r}$.

Galerkin Projection-ROM (GP-ROM). For a given space Φ_r, the GP-ROM finds the approximation of the velocity field spanned by the low-dimensional space,

$$\boldsymbol{u} \approx \boldsymbol{u}_r(\mathbf{x}, t) \equiv \sum_{j=1}^{r} a_j(t) \boldsymbol{\varphi}_j(\mathbf{x}), \tag{4}$$

where $\{a_j(t)\}_{j=1}^r$ are the sought time-varying coefficients. The GP-ROM can be obtained by projecting the FOM onto the POD space: $\forall\, i = 1, \ldots, r$,

$$\left(\frac{\partial \boldsymbol{u}_r}{\partial t}, \boldsymbol{\varphi}_i \right) = (\mathbf{f}(\boldsymbol{u}_r, Re), \boldsymbol{\varphi}_i). \tag{5}$$

Here, $(\cdot, \cdot)$ is the L^2 inner product. The solution of GP-ROM can be determined by the following nonlinear dynamic system:

$$\dot{\boldsymbol{a}} = L\,\boldsymbol{a} + \boldsymbol{a}^\top N\,\boldsymbol{a}, \tag{6}$$

where L and N are ROM operators that can be obtained by projection

$$L_{im} = \frac{1}{Re}\,(\Delta \boldsymbol{\varphi}_m, \boldsymbol{\varphi}_i)\,, N_{imn} = -\left(\boldsymbol{\varphi}_m \cdot \nabla \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_i\right) \tag{7}$$

Note that the reduced system (6) efficiently approximates the full model of NSE as the dimension r is generally very small ($\sim \mathcal{O}(10)$) compared to the high dimension of $\boldsymbol{u}$ ($D_h \sim \mathcal{O}(\geq 10^5)$).

ROM closure models. The ROM closure models can be generally written as the following dynamic system:

$$\dot{\boldsymbol{a}} = L\,\boldsymbol{a} + \boldsymbol{a}^\top N\,\boldsymbol{a} + \boldsymbol{\tau}, \tag{8}$$

108 τ is an artificial term that model the effect of the discarded POD modes using various approaches
109 e.g., see [17,19,27–29]. We note that the dynamics of closure model (8) is more accurate than GP-ROM
110 dynamics (6) in approximation of the model. The challenge of closure modeling is that the τ is
111 unknown, i.e., no explicit formula. Therefore, closure models are empirical modification of the
112 GP-ROM from physical or mathematical perspective.

3. Learning Reduced Order Model

114 In this section, we present the architecture of learning the reduced order model from deep neural
115 networks. In contrast to the standard GP-ROM framework, we learn the reduced dynamical system
116 from the data without intrusively using ROM operators (e.g., L, N)

3.1. ROM dynamics

118 We consider the low dimensional ROM dynamic system as a general function,

$$\dot{a} = \mathbf{f}_r(a). \tag{9}$$

119 We claim that this function $\mathbf{f}_r$ has a general representation of the ROM dynamics including system (6)
120 and (8). Our goal is to learn the ROM dynamics (9) in a given set of temporal data and return a closed
121 form model that can be used to accurately approximate and predict the full system.

122 For a given data-set of snapshot solutions of NSE (1), $U = u^0, ..., u^s \in \mathcal{R}^{D_h \times s+1}$, at time steps
123 $t_0, ..., t_s$. The best approximation of snapshots data by the POD space is given by, $u^j = \sum_{i=1}^{r} b_i^j \varphi_i$,
124 $j = t_0, ..., t_s$. The reduced dynamics (9) is to find coefficients a such that,

$$u_r = \sum_{i=1}^{r} a_i^j \varphi_i \approx u = \sum_{i=1}^{r} b_i^j \varphi_i \tag{10}$$

125 It indicates that the optimal solution from the r-dimensional ROM dynamic system is given by the full
126 model data, b^j, such that

$$\dot{b} = \mathbf{f}_r(b) \tag{11}$$

127 This provides a framework that the ROM dynamics can be learned from the data-set $B = [b^0, ..., b^s] \in$
128 $\mathcal{R}^{r \times (s+1)}$, i.e., time-varying coefficients of FOM data. The training data-set can be computed by the
129 following,

$$B = \Phi_r^T W U \tag{12}$$

130 The formula is derived by the L^2 projection of data U to the low-dimensional space Φ_r. $W \in \mathcal{R}^{D_h \times D_h}$
131 is the weight matrix of the L^2 inner product, we use finite element weights in this work.

3.2. Linear Multistep Network (LMNet)

133 Motivated by differential equation learning, we adopt the linear multistep network architecture
134 [21,30] to construct a structured nonlinear regression model that can learn the reduced dynamics. In
135 this work, we only consider the implicit multistep method, Adams-Moulton (AM) scheme [31] as it
136 has better stability property. The K-step AM method is defined as follows:

$$a^n = \sum_{i=1}^{K} (\alpha_i a^{n-i} + \beta_i \Delta t \mathbf{f}_r(a^{n-i})) + \beta_0 \Delta t \mathbf{f}_r(a^n) \tag{13}$$

We discretize the ROM dynamic system (9) by using AM scheme (13) with a neural network. The parameters of this neural network can be learned by minimizing the mean squared error loss function:

$$MSE := \frac{1}{N - K + 1} \sum_{n=K}^{N} |\mathcal{L}_n|^2 \tag{14}$$

Where S is the total number of time instance in the ROM dynamic system. $\mathcal{L}_n$ is the local truncation error from the Taylor expansion of K-step AM method (13),

$$\mathcal{L}_n = \sum_{i=0}^{K} \alpha_i a^{n-i} + \Delta t \beta_i \mathbf{f}_r(a^{n-1}). \tag{15}$$

One advantage of this nonlinear regression is that we do not have to approximate the temporal gradients [11,23] since the time derivatives are discretized by the AM method.

3.3. LMNet-ROM

We use the trained neural network as the ROM dynamic system (9) to approximate the full system of NSE. We emphasize that the novelty of our approach is the non-intrusively learning of the reduced system, whereas GP-ROM and closure models require the use of FOM operators. Fig.1 is the flowchart of the LMNet-ROM and GP-ROM framework.
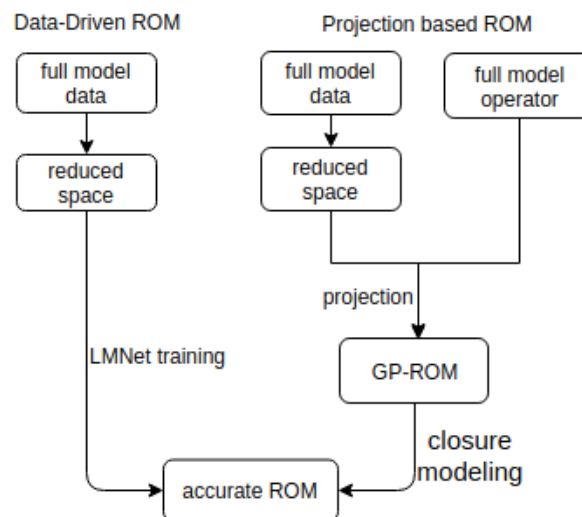

Figure 1. Flowchart of projection based model reduction and the new non-intrusive learning reduced order modeling framework

We outline the algorithm of the framework in the following: We claim that the learned reduced

---

Algorithm 1 Linear Multistep Network Reduced Order Model Learning (LMNet-ROM)

  Compute the reduced POD space from the data of NSE by (3)
  Compute the training dataset B via (12)
  Train the neural network using loss function (14)
  The LMNet-ROM for NSE is obatined from the trained low-dimensional dynamic system:

$$\boxed{\dot{a} = \mathbf{f}_r^{Net}(a)} \tag{16}$$

---

dynamics (16) has a better approximation to the full model than system (7) and (8) since it is learned optimally from the FOM data. The new framework, see Fig.1, only requires input data from a system

and does not use any FOM operator, which can be generally applied to reduced order modeling of any fluid dynamical system. The main offline computational cost of the LMNet-ROM is training the neural network whereas the GP-ROM is the construction of ROM operators in (7). In the numerical experiment, we show that the offline computation of our model is faster than the GP-ROM.

4. Numerical Experiment

In this section, we present the preliminary numerical results to demonstrate the advantages of our new model. The test case is a 2D channel flow past a circular cylinder at a $Re = 100$. It is a benchmark problem that has been used as a numerical test in fluid dynamics, see [8,17,32–34].

4.1. Implementation Details

The domain is a 2.2×0.41 rectangular channel with a radius=0.05 cylinder, centered at $(0.2, 0.2)$, see Figure 2. No slip boundary conditions are prescribed for the walls and on the cylinder, and the inflow and outflow profiles are given by [17,34] $u_1(0, y, t) = u_1(2.2, y, t) = \frac{6}{0.41^2} y(0.41 - y)$, $u_2(0, y, t) = u_2(2.2, y, t) = 0$. The kinematic viscosity is $\nu = 10^{-3}$, there is no forcing, and the flow starts from rest.
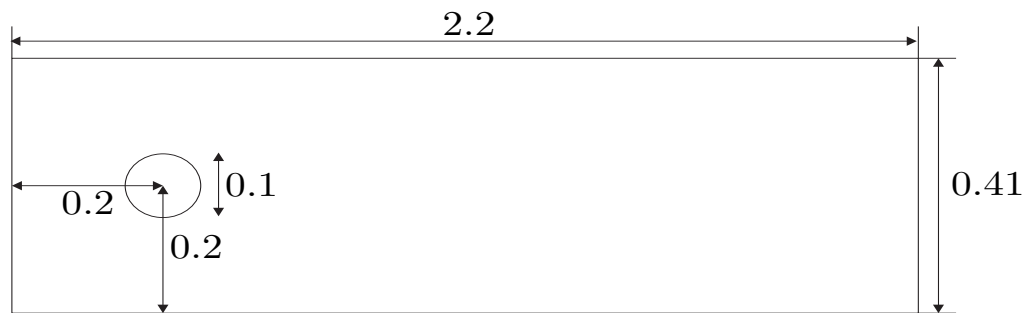

Figure 2. Channel flow around a cylinder domain.

The velocity snapshots of NSE (1) are generated by finite element method with meshsize approximately 103K which gives a fully resolved solution. The lift and drag computation agree well with results from references in [35]: $c_{d,max} = 3.2261$, $c_{l,max} = 1.0040$.

A total number of 2500 snapshots U are collected from $T = [0, 5]$ at every time step $\Delta t = 0.002$. The LMNet-ROM is built and trained in Tensorflow.

4.2. Full Order Model Approximation

After obtaining the trained neural network, we use the ROM dynamic system (16) to approximate the full model of NSE (1). The following average L^2 error formula has been used to quantify the accuracy of the model,

$$\mathcal{E} = \frac{1}{s+1} \sqrt{\sum_{j=0}^{s} \int_{\Omega} (\boldsymbol{u}^j - \boldsymbol{u}_r^j) d\Omega}. \tag{17}$$

We first evaluate the model with different layers and neurons. The dimension of the ROM dynamics is fixed to be $r = 8$ with the one step Adams Moulton method. Table 1 provides a crude estimate that increasing network width (256 neurons) might have the potential over-fitting issue whereas decreasing units (64) may not be enough to reach a good accuracy. Also, the network depth (number of hidden layers) have a positive effect on the performance of the model. To reduce the numerical efforts, we use one hidden layer and 128 neurons for the neural network in the rest of our

evaluations. We emphasize that to fully understand the model sensitivity with respect to network architecture, a systematic study involving regularization, batch normalization and drop out is needed in future research.

Table 1. Average L^2 error between trajectories of the learned reduced order model with dimension $r = 8$ and the exact data for the different number of hidden layers and neurons

layers \ neurons	64	128	256
1	5.62e-02	7.22e-03	4.31e-01
2	4.01e-03	5.75e-03	2.95e-02
3	1.64e-02	4.04e-03	3.11e-03

We also test the LMNet-ROM with different steps in the training. Table 2 lists the error between the new model and the exact data. As a comparison, we add GP-ROM result in the last entry. Table 2 shows that the LMNet-ROM consistently provide more accurate results as increasing steps of Adams Moulton (AM) method. An intuitive explanation is that the stability property of AM method helps to regularize the network and eventually achieve a good calibration. Large steps (K) requires high computational cost of training the network. To balance the output accuracy and training cost, we use step $K = 1$ for the rest of our numerical tests.

Table 2. Average L^2 error between the new model and exact data for the different number of steps and dimensions

K/model \ dimension	r=4	r=6	r=8
1	1.88e-03	3.05e-03e	7.22e-03
2	2.83e-04	6.04e-04	1.45e-03
3	2.67e-04	6.22e-04	7.14e-03
4	3.79e-04	6.20e-04	6.37e-04
GP-ROM	1.66e-01	7.30e-02	1.92e-02

Table 3. Average L^2 error for different noise magnitudes.

noise \ model	GP-ROM	LMNet-ROM
0.0%	1.92e-02	7.22e-03
0.5%	1.93e-02	2.44e-03
1%	2.05e-02	2.20e-02
5%	2.92e-02	8.76e+01

Noisy data. The above numerical tests are carried out on the deterministic data. In some situations, however, problems may contain noise measurements. We study the robustness of the new method with respect to noise data by adding Gaussian noise to the data-set for both models. Table 3 lists the error comparison between the LMNet-ROM and GP-ROM for different level of noise data. The results show that the LMNet-ROM cannot preserve good performance when the noise magnitude ($\geq 1\%$) is high while GP-ROM does. The argument is that the learned dynamical system fully depends on the data which makes it vulnerable to noise interference. The GP-ROM, however, requires the use of FOM operators making it less sensitive than LMNet-ROM to noise. Further approaches

should be investigated for the potential improvement for this problem. As for deterministic data, the LMNet-ROM is better than GP-ROM in full system approximation.

4.3. Long-Term Prediction

In this section, we make a thorough study of long-term predictability of the new LMNet-ROM. The solution to our new model and the GP-ROM are computed by the dynamic system (16) and (6), respectively. We use direct numerical simulation (DNS) to denote the exact solution (data) of NSE. We use snapshots data that are collected from the time interval $T = [0,3]$ to generate the POD space and train the neural network. We then run the reduced systems (16) and (6) for $T = [0,5]$ to make the prediction. Fig. 3 plots the phase portraits of the first few coefficients, a_2, a_3, a_4, from both models and the DNS data. The red line that depicts the result from LMNet-ROM has a closer mimic of DNS data, whereas the portraits from the GP-ROM have small deviation. This behavior tells that the dynamics of LMNet-ROM predict future states better than GP-ROM with the given information. Note that a_1 is a constant and not meaningful to discuss since the first POD mode φ_1 represents the mean flow.
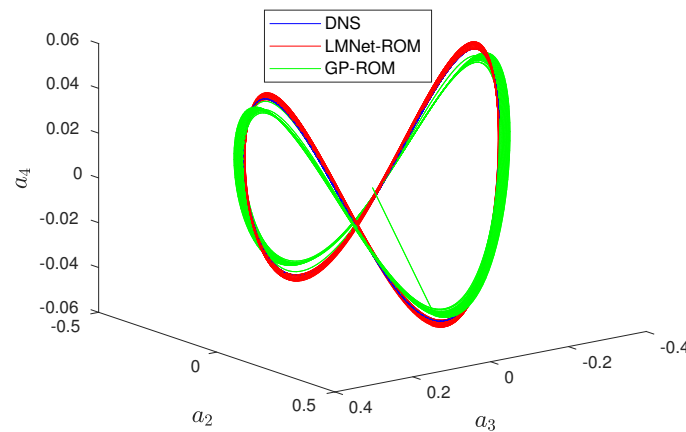

Figure 3. phase portraits of the coefficients a_2, a_3, a_4 from LMNet-ROM (red), GP-ROM (green) and DNS data (blue) with dimension $r = 8$

We also looked at the prediction of the time evolution of energy ($E(t_j) = \frac{1}{2}\|\boldsymbol{u}^j\|_{L^2}$), vorticity construction, and drag coefficients. Fig. 4 shows the long-term prediction of energy travel and drag. The main observation is that the LMNet-ROM performs much better than GP-ROM. The energy and drag generated from GP-ROM have a huge deterioration when time evolves, which means the prediction is not accurate. Clearly, the prediction from LMNet-ROM is impressively good as it is stable and close to DNS, see Fig. 4. Vorticity describes the local spinning motion of a fluid system. Fig. 5 plots the vorticity construction from the velocity field around the cylinder at the end time $T = 5$. As depicted in Fig. 5, the LMNet-ROM correctly predicts the vortex street behind the cylinder while GP-ROM not. The above results are presented for dimension $r = 8$, but similar results can be found for $r = 4, 6$. Overall, the long-term predictability of LMNet-ROM is much better than GP-ROM.

4.4. LMNet-ROM vs. Closure Models

In this section, we present a preliminary numerical comparison between the LMNet-ROM and the closure models. Due to the wide class of stabilization closure models and the availability of their open source implementation, a thorough comparison between our new model and all other models is not practical. Therefore, we use the two most recent models (open source implemenation), Data-Driven-Filtered ROM (DDF-ROM) and Evolve-then-Filter ROM (EF-ROM), to illustrate the numerical comparison. The EF-ROM uses a two steps regularization strategy to improve the GP-ROM [29]. The DDF-ROM solves an optimization problem to approximate the τ in system (8) [17]. Table 4
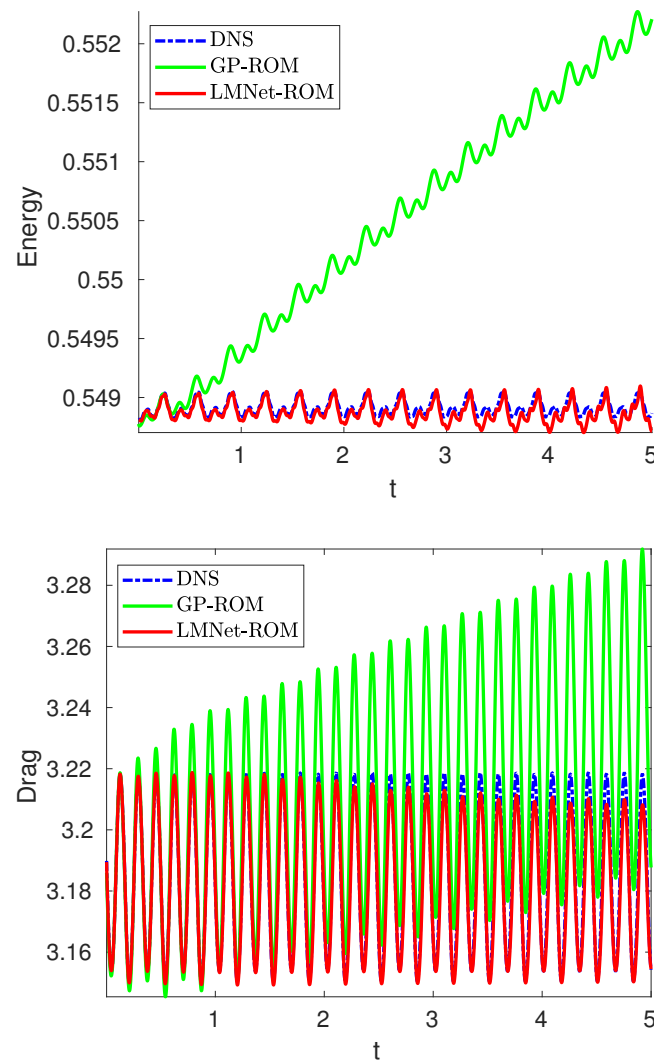
Figure 4. Plots of the time evolution of energy $E(t_j)$ (top) and drag (bottom). The solutions are generated from LMNet-ROM and GP-ROM with dimension $r = 8$

lists the average L^2 error from the three models with different dimensions. It surprisingly indicates that the closure models with mathematical methods cannot outperform the LMNet-ROM. This result is impressive, given that the LMNet-ROM is learned through the neural network from the data without acquiring any FOM information.

We note that the goal of both the stabilization closure models and our new model is to provide an accurate reduced dynamics to approximate the full system from different perspective. The former generally model the artificial term (τ) in the system (8) physically or mathematically, whereas the latter use data to learn the dynamics (16). We claim that the non-intrusive learning framework has the generality that can be applied to any nonlinear fluid system since it only requires the training data.

4.5. Computational Cost

In this section, we discuss the computational efficiency of the proposed LMNet-ROM. The main computational cost of reduced order models is the offline computation since the cost of solving a small ODE system is negligible at the online stage. The FOM simulation (DNS) time is used as a benchmark to evaluate the performance of each model. The computation is carried out on a 64-bit Linux system with a single 2.70 GHz CPU. The DNS CPU time is 36828.53(s). Table 5 lists the CPU time
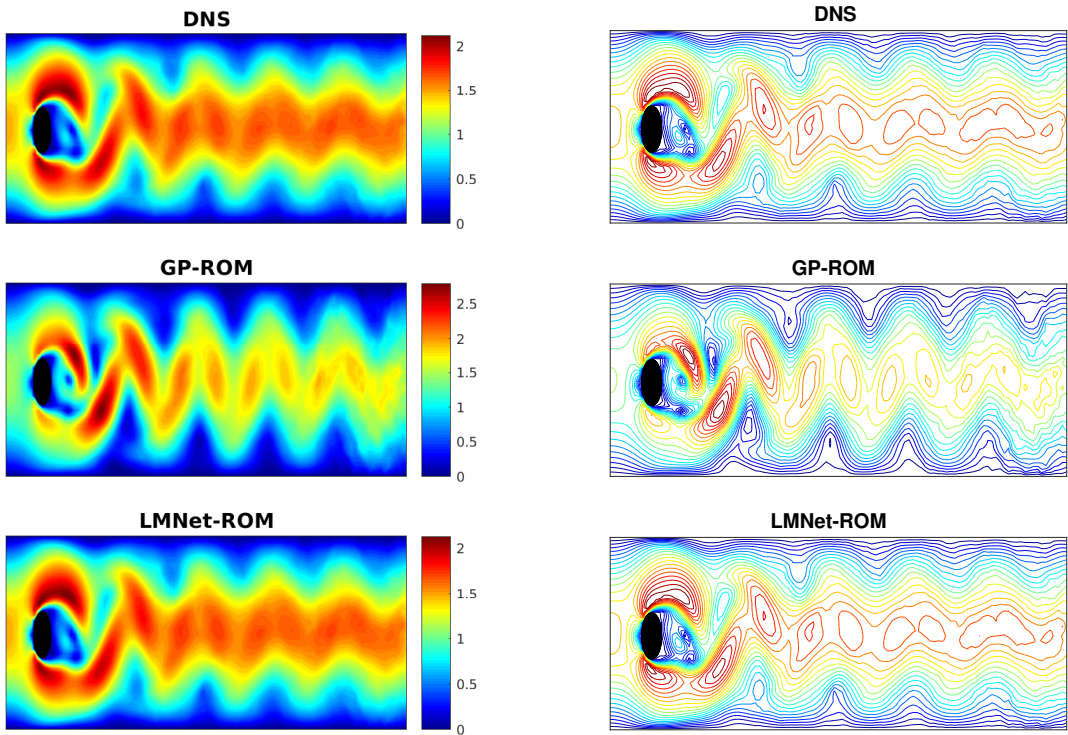
Figure 5. Vorticity prediction plots from the solution of LMNet-ROM (right) and GP-ROM (middle) with dimension $r = 8$. The exact data (DNS) is plotted on the left.

Table 4. Average L^2 error from different models

model \ dimension	r=4	r=6	r=8
EF-ROM	1.23e-01	7.31e-02	1.84e-02
DDF-ROM	2.27e-01	1.14e-02	1.22e-02
LMNet-ROM	**1.91e-03**	**3.57e-03**	**7.22e-03**

from each ROM and the associated speed-up factor. The LMNet-ROM time is only computed from one step AM method with one layer and 128 neurons, given the fact that this network architecture achieves good accuracy in the previous test. The result in Table 5 reveals that the LMNet-ROM is more efficient compared to other models. This is a huge advantage of our new method since reducing the computational cost while maintaining good accuracy is the primary goal of reduced order modeling.

Table 5. offline cost (second) and speed up factor from each ROM with dimension $r = 8$

model	cost	speed-up factor $(\frac{DNS}{ROM})$
GP-ROM	855.52s	43.05
EF-ROM	867.20s	42.47
DDF-ROM	6373.97s	5.78
LMNet-ROM	**445.25(s)**	**82.71**

5. Conclusions and Outlook

In this paper, we proposed a novel learning reduced order model framework for the numerical simulation of fluid flows. This framework was based on the recent development of linear multistep

network architecture. We numerically studied the LMNet-ROM in the simulation of a 2D flow past a cylinder. The numerical results demonstrate that the LMNet-ROM was significantly more accurate than the GP-ROM in system approximation and long-term prediction. Furthermore, we compared the new model with the two most recent stabilization closure models, EF-ROM and DDF-ROM. The results show that our new method outperforms the two closure models. Overall, the LMNet-ROM beats the aforementioned models both in accuracy and computational efficiency, which provides a promising and encouraging approach in model reduction of fluid dynamics. However, the LMNet-ROM's potential still needs to be explored. We outline some research directions that could be pursued.

Probably the most important next step is to study the parametrized system prediction of the LMNet-ROM. The current neural network is trained under the data-set given by one parameter value (*Re*) from the full order model of NSE. How does the LMNet-ROM predict systems with different parameter values, such as initial conditions and boundary conditions? Parametrized system prediction is a challenging problem in engineering applications. We hope to provide a systematic investigation with the new LMNet-ROM in later research.

Another important research direction is to improve the robustness of the model with respect to noise data. Table 3 shows the drawback of this model for high magnitude noisy ($\geq 1\%$) data. We plan to address this issue by improving the neural network architecture. Also, regularization for preventing over-fitting needs to be fully studied in future research.

Finally, the generality study of LMNet-ROM is worth investigation. Although we constructed and tested the LMNet-ROM in a fluid dynamics setting, the LMNet-ROM framework can be applied to any type of nonlinear partial differential equations (PDE) that is amenable to reduced order modeling. The only input needed in the LMNet-ROM framework is the data from the FOM of any system, see Algorithm 1. The LMNet-ROM procedure does not restrict it to the particular physical system modeled by the given nonlinear PDE. Since the LMNet-ROM is built by fully data-driven learning, we expect it to be successful in the numerical simulation of general mathematical models (e.g., from elasticity or bioengineering).

Author Contributions: Investigation, Xuping Xie; Methodology Xuping Xie and Guannan Zhang; Project administration, Clayton G. Webster; Writing – original draft, Xuping Xie; Writing – review & editing, Guannan Zhang and Clayton G. Webster.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ROM	Reduced-Order Model
LMNet	Linear Multistep Neural Network
DNS	Direct Numerical Simulation
GP-ROM	Galerkin Projection Reduced-Order Model
EF-ROM	Evolve-then-Filter Reduced-Order Model
DDF-ROM	Data-Driven Reduced-Order Model

1. Lumley, J.L. The structure of inhomogeneous turbulent flows. *Atmospheric turbulence and radio wave propagation* **1967**.

2. Noack, B.R.; Morzynski, M.; Tadmor, G. *Reduced-Order Modelling for Flow Control*; Vol. 528, Springer Verlag, 2011.

3. Obinata, G.; Anderson, B.D. *Model reduction for control system design*; Springer Science & Business Media, 2012.

4. Carlberg, K.; Farhat, C.; Cortial, J.; Amsallem, D. The GNAT method for nonlinear model reduction: effective implementation and application to computational fluid dynamics and turbulent flows. *J. Comput. Phys.* **2013**.

5. Rowley, C.W.; Dawson, S.T. Model reduction for flow analysis and control. *Annual Review of Fluid Mechanics* **2017**, *49*, 387–417.

6. Amsallem, D.; Farhat, C. Stabilization of projection-based reduced-order models. *Int. J. Num. Meth. Eng.* **2012**, *91*, 358–377.

7. Xie, X.; Wells, D.; Wang, Z.; Iliescu, T. Approximate deconvolution reduced order modeling. *Comput. Methods Appl. Mech. Engrg.* **2017**, *313*, 512–534.

8. Kutz, J.N.; Brunton, S.L.; Brunton, B.W.; Proctor, J.L. *Dynamic mode decomposition: data-driven modeling of complex systems*; Vol. 149, SIAM, 2016.

9. San, O.; Maulik, R. Extreme learning machine for reduced order modeling of turbulent geophysical flows. *Physical Review E* **2018**, *97*, 042322.

10. Xiao, D.; Fang, F.; Buchan, A.; Pain, C.; Navon, I.; Muggeridge, A. Non-intrusive reduced order modelling of the Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering* **2015**, *293*, 522–541.

11. Peherstorfer, B.; Willcox, K. Data-driven operator inference for nonintrusive projection-based model reduction. *Computer Methods in Applied Mechanics and Engineering* **2016**, *306*, 196–215.

12. Noack, B.R.; Stankiewicz, W.; Morzyński, M.; Schmid, P.J. Recursive dynamic mode decomposition of transient and post-transient wake flows. *Journal of Fluid Mechanics* **2016**, *809*, 843–872.

13. Loiseau, J.C.; Brunton, S.L. Constrained sparse Galerkin regression. *Journal of Fluid Mechanics* **2018**, *838*, 42–67.

14. Balajewicz, M.J.; Dowell, E.H.; Noack, B.R. Low-dimensional modelling of high-Reynolds-number shear flows incorporating constraints from the Navier–Stokes equation. *J. Fluid Mech.* **2013**, *729*, 285–308.

15. 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.* **2015**, *102*, 1136–1161.

16. Carlberg, K.; Barone, M.; Antil, H. Galerkin v. discrete-optimal projection in nonlinear model reduction. *arXiv preprint arXiv:1504.03749* **2015**.

17. Xie, X.; Mohebujjaman, M.; Rebholz, L.; Iliescu, T. Data-driven filtered reduced order modeling of fluid flows. *SIAM Journal on Scientific Computing* **2018**, *40*, B834–B857.

18. Protas, B.; Noack, B.R.; Östh, J. Optimal nonlinear eddy viscosity in Galerkin models of turbulent flows. *J. Fluid Mech.* **2015**, *766*, 337–367.

19. San, O.; Maulik, R. Neural network closures for nonlinear model order reduction. *Advances in Computational Mathematics* **2018**, pp. 1–34.

20. Chang, B.; Meng, L.; Haber, E.; Tung, F.; Begert, D. Multi-level residual networks from dynamical systems view. *arXiv preprint arXiv:1710.10348* **2017**.

21. Lu, Y.; Zhong, A.; Li, Q.; Dong, B. Beyond finite layer neural networks: Bridging deep architectures and numerical differential equations. *arXiv preprint arXiv:1710.10121* **2017**.

22. Brunton, S.L.; Proctor, J.L.; Kutz, J.N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **2016**, p. 201517384.

23. Rudy, S.H.; Brunton, S.L.; Proctor, J.L.; Kutz, J.N. Data-driven discovery of partial differential equations. *Science Advances* **2017**, *3*, e1602614.

24. Raissi, M.; Karniadakis, G.E. Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics* **2018**, *357*, 125–141.

25. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Machine learning of linear differential equations using Gaussian processes. *Journal of Computational Physics* **2017**, *348*, 683–693.

26. Chen, T.Q.; Rubanova, Y.; Bettencourt, J.; Duvenaud, D. Neural Ordinary Differential Equations. *arXiv preprint arXiv:1806.07366* **2018**.

27. Parish, E.J.; Duraisamy, K. A paradigm for data-driven predictive modeling using field inversion and machine learning. *Journal of Computational Physics* **2016**, *305*, 758–774.

349 28. Gouasmi, A.; Parish, E.J.; Duraisamy, K. A priori estimation of memory effects in reduced-order models of
350   nonlinear systems using the Mori–Zwanzig formalism. *Proc. R. Soc. A* **2017**, *473*, 20170385.
351 29. Wells, D.; Wang, Z.; Xie, X.; Iliescu, T. An evolve-then-filter regularized reduced order model for
352   convection-dominated flows. *International Journal for Numerical Methods in Fluids* **2017**, *84*, 598–615.
353 30. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Multistep Neural Networks for Data-driven Discovery of
354   Nonlinear Dynamical Systems. *arXiv preprint arXiv:1801.01236* **2018**.
355 31. Ascher, U.M.; Petzold, L.R. *Computer methods for ordinary differential equations and differential-algebraic*
356   *equations*; Vol. 61, Siam, 1998.
357 32. Schäefer, M.; Turek, S. The benchmark problem "flow around a cylinder". *Flow Simulation with*
358   *High-Performance Computers II* **1996**, *52*, 547566.
359 33. Brunton, S.; Tu, J.; Bright, I.; Kutz, J. Compressive sensing and low-rank libraries for classification of
360   bifurcation regimes in nonlinear dynamical systems. *SIAM Journal on Applied Dynamical Systems* **2014**,
361   *13*, 1716–1732.
362 34. Mohebujjaman, M.; Rebholz, L.G.; Xie, X.; Iliescu, T. Energy balance and mass conservation in reduced
363   order models of fluid flows. *Journal of Computational Physics* **2017**, *346*, 262–277.
364 35. Caiazzo, A.; Iliescu, T.; John, V.; Schyschlowa, S. A numerical investigation of velocity-pressure reduced
365   order models for incompressible flows. *J. Comput. Phys.* **2014**, *259*, 598–616.

366 **Sample Availability:** Samples of the compounds ...... are available from the authors.