

Basis Identification for Reduced Order Modeling of Unsteady Flows Using Sparse Coding

Rohit Deshmukh*, Zongxian Liang†, Abhijit Gogulapati‡ and Jack J. McNamara§

Department of Mechanical and Aerospace Engineering

The Ohio State University, Columbus, OH, 43210, USA

J. Zico Kolter¶

School of Computer Science

Carnegie Mellon University, Pittsburgh, PA, 15213, USA

Basis identification is a critical step in the construction of accurate reduced order models using Galerkin projection. This is particularly challenging in unsteady flow fields due to the presence of multi-scale phenomena that cannot be ignored, and are not well captured using the ubiquitous Proper Orthogonal Decomposition. This study focuses on this issue by exploring an approach known as sparse coding for the basis identification problem. Compared to Proper Orthogonal Decomposition, which seeks to truncate the basis spanning an observed data set into a small set of dominant modes, sparse coding seeks a compact basis that best spans the entire data set. Thus, the resulting bases are inherently multi-scale, enabling improved reduced order modeling of unsteady flow fields. The approach is demonstrated for two canonical problems — a 2-D incompressible flow past a moving cylinder, and an incompressible flow inside a 2-D lid-driven cavity. Results indicate that a set of sparse modes generalize better to unseen flow compared to a truncated set of the traditional proper orthogonal modes. Furthermore, Galerkin reduction of the governing equations using sparse modes yields significantly improved fluid predictions.

Nomenclature

| | |
|-----------|---|
| A | = boundary of the computational domain |
| E | = percent error in the fluctuating component of the predicted flow |
| m | = total number snapshots in a snapshot matrix |
| N | = total number of modes used in projection |
| \hat{n} | = outer-pointing normal vector of A |
| p | = pressure |
| Q | = snapshot matrix |
| q | = fluctuating component of the flow field corresponding to the full-order model |
| Re | = Reynolds number |
| S | = coefficient matrix |
| s^i | = i^{th} coefficient corresponding to the i^{th} mode at any time instant |
| s_k | = k^{th} column of the coefficient matrix |
| t | = non-dimensional time |
| \bar{U} | = mean velocity component |
| u_1 | = velocity in the streamwise direction |

*Ph.D candidate, deshmukh.17@osu.edu, AIAA Student Member

†Post Doctoral Researcher, liang.593@osu.edu, AIAA Member

‡Post Doctoral Researcher, gogulapati.1@osu.edu, AIAA Member

§Associate Professor, mcnamara.190@osu.edu, AIAA Senior Member

¶Assistant Professor, zkcolter@cs.cmu.edu

| | |
|-------------------|--|
| u_2 | = velocity in the cross-flow direction |
| \tilde{u} | = prediction for the fluctuating component of the flow |
| x | = position vector |
| x | = coordinate in the streamwise direction |
| y | = coordinate in the cross-flow direction |
| β | = sparsity coefficient |
| η | = time-averaged accumulated energy |
| Ω | = computational domain |
| Φ | = matrix containing modes in its columns |
| Φ_i | = i^{th} mode |
| ∇ | = the gradient operator |
| $\ \cdot\ $ | = L2 norm, equal to the sum of squares of the entries in (\cdot) |
| $\ \cdot\ _0$ | = L0 norm, equal to the number of non-zero entries in (\cdot) |
| $\ \cdot\ _{1,1}$ | = L1 norm, equal to the sum of the absolute values of entries in (\cdot) |
| <i>Subscripts</i> | |
| i, j | = indices corresponding to the streamwise and cross-flow directions |

I. Introduction

The advancement of computational fluid dynamics (CFD), parallel computing algorithms, and computing hardware, has enabled unprecedented insight into complex flow physics through numerical analysis. However, the computational expense associated with such analysis tools has restricted their application to relatively small spatial and temporal scale studies on simplistic configurations. Thus, the advances in CFD methodologies have not yet enabled the broad consideration of nonlinear, multi-scale, unsteady flows, in systems level studies - e.g. fluid-structure interaction, flow control, aerodynamic design, structural design, etc. Yet, the critical need for high fidelity flow modeling in such problems motivates the pursuit of tractable and robust reduced order models (ROMs).

A common reduced order modeling approach is to project the governing equations onto a reduced dimensional space comprised of characteristic bases.¹⁻⁶ These can be based on orthogonal (e.g., Galerkin) or non-orthogonal (e.g., Petrov-Galerkin) projections.¹⁻⁶ However, the accuracy of such approaches is intimately bound to the quality of the chosen bases. This is challenging for nonlinear problems, due to the need to carry out basis identification from post execution data of representative system dynamics. In the context of highly unsteady nonlinear flows, this is further complicated by the fact that generally: 1) the gathering of the data is computationally expensive, 2) the data is very large and high dimensional, and 3) the scope of the data is narrow. Thus, identifying a compact set of prominent and dynamically important flow features to fundamentally characterize the fluid dynamics is a non-trivial problem.

Proper Orthogonal Decomposition (POD), or Principal Components Analysis (PCA),^{1,7-10} is a widely used and explored technique aimed at meeting this need. The approach is based on identifying and ordering principal components in observed data. The POD modes are optimal in terms of capturing the energy of an observed flow response, thus a reduced dimensional basis of the system is often identified by truncating the modes based on energy contribution to the response.^{1,7-10} However, there are several issues with this approach. First, the POD modes are only optimal in the sense of reconstructing the observed flow responses.^{1,11,12} Thus, they may not generalize well for model predictions that deviate from observed conditions. Also, from a fluid physics perspective, a truncated set of POD modes is biased towards the high-energy, large-scale, dominant structures and ignores the small-scale, low-energy structures.^{1,11,13} The large-scale structures are formed as a result of disturbances in the flow; obtaining energy from the mean flow, and then subsequently breaking down into smaller scales.¹⁴ The small-scale structures then cause energy dissipation and result in viscosity in the fluid-flow. Thus, POD based Galerkin ROMs do not account for sufficient energy dissipation, resulting in over-prediction of kinetic energy.^{1,11,13} Moreover, the energy accumulation over a period of time may also cause the ROM to become unstable.^{11,13,15} Related to this, it is now well-established that POD modes are ineffective in capturing the local dynamics (or transience) of

full-order systems.^{11,12} This is because the POD modes are always active (global) for dynamical systems. For these reasons it is clear that the optimality of POD modes, in terms of energy capture, is non-ideal for model reduction of nonlinear, unsteady flows.

Techniques such as Balanced Truncation,¹⁶ Balanced POD (BPOD),² and Eigensystem Realization Algorithm (ERA)¹⁷ have addressed some of the limitations identified earlier. However, balanced truncation is intractable for large data (for more than 10,000 degrees of freedom),³ BPOD is only applicable to response data of linear systems as it requires adjoint system information,^{2,17} and modes generated by ERA cannot be used for projection of non-linear dynamics.¹⁷ In a recent study, a technique is developed to generate a stable Galerkin projection based ROM.¹¹ However, building the ROM is an iterative process, and requires multiple time-integrations until an energy-balance is achieved. These issues highlight the need to explore alternative basis identification techniques that not only generalize well to changing flow conditions, but also accurately capture essential multi-scale features.

Olshausen and Field¹⁸ argue that most naturally occurring phenomena are conveniently represented using non-Gaussian distributions, whereas the PCA approach is suitable when the structure of the data can be represented using Gaussian distributions. In Gaussian distributions, the linear correlation between statistical structures is the most important relation. Observations from a naturally occurring phenomena, such as natural images, contain higher order statistics. To this end, a technique based on sparse coding was proposed to extract the higher order features from natural image data.¹⁸ This approach, which is also referred to as sparse dictionary learning,¹⁹ generates a finite dictionary of modes in which only a subset is active - i.e. has nonzero coefficients — at a given time. Furthermore, sparse coding describes a nonlinear system in a locally linear manner by tailoring the modes to local behavior of the system.²⁰ Thus, compared to the POD approach — where the principal components of the observed data are identified, ordered and then truncated to a compact set — sparse coding is formulated as a procedure to identify a compact representation that best spans the entire observed data. The sparse coding approach has been successfully applied in a number of topics, such as in image processing,²¹ audio analysis,²² neuroscience,^{18,23,24} and electrical power disaggregation.²⁵ However, to the authors' knowledge, this method has not been examined in the context of reduced order modeling of dynamical systems.

The overall objective of this study is to apply the concepts of sparse coding to identify coherent global and local structures in unsteady flows, and develop robust and accurate ROMs to predict the flow field behavior. The specific objectives of the paper are as follows:

1. Develop a framework that combines concepts of sparse coding with a Galerkin based projection methodology in order to generate predictive ROM for unsteady flow fields.
2. Investigate the performance of the ROMs for canonical incompressible unsteady flows.

The remainder of this paper is organized as follows. The POD and the sparse coding approaches are presented in Section II. Case studies describing the application of POD and sparse modes to model the unsteady flow fields are presented in Section III. Concluding remarks are presented in Section IV.

II. Method of Solution

High resolution data are computed by solving the 2-D incompressible Navier Stokes (NS) equations using Unsteady Reynolds Averaged Navier-Stokes (RANS) and Direct Numerical Simulation (DNS) CFD solvers. The POD approach is based on the method of snapshots developed by Sirovich.¹⁰ Sparse modes are evaluated using the algorithm developed in Ref. 26. These approaches are detailed next.

II.A. Full Order Models

The RANS and DNS solutions to the non-dimensionalized NS equations given by Eq.(1) are generated using CFL3D²⁷ and PICar3D,²⁸ respectively.

$$\frac{\partial u_i}{\partial x_i} = 0, \quad \frac{\partial u_i}{\partial t} + \frac{\partial u_i u_j}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{Re} \frac{\partial^2 u_i}{\partial x_i \partial x_j} \quad (1)$$

The PICar3D code uses a second order central difference spatial scheme and a second-order fractional-step method for time marching. The equations are discretized on a Cartesian mesh and boundary conditions are imposed using a ghost-cell procedure.²⁸ The CFL3D code solves the incompressible Unsteady

RANS on a cylindrical grid.²⁷ The turbulent viscosity is modeled using the Spalart-Allmaras turbulence model.

II.B. Generation of the Reduced Order modes

The full order solution is decomposed into mean ($\bar{U}(\mathbf{x})$) and fluctuating components ($\mathbf{q}(\mathbf{x}, t)$) of velocity. The component $\bar{U}(\mathbf{x})$ is evaluated as time average of velocity field. The component ($\mathbf{q}(\mathbf{x}, t)$) is approximated as a linear combination of the reduced order modes:

$$\mathbf{q}(\mathbf{x}, t) \approx \sum_{i=1}^N s^i(t) \Phi_i(\mathbf{x}) = \tilde{\mathbf{u}}(\mathbf{x}, t) \quad (2)$$

The computation of POD and sparse modes are discussed next.

II.B.1. Proper Orthogonal Decomposition

The procedure to extract POD modes from a snapshot matrix $\mathbf{Q} = [\mathbf{q}_1 \mathbf{q}_2 \cdots \mathbf{q}_k \cdots \mathbf{q}_m]$ has been extensively documented^{8,10} and therefore is not repeated here. However, for comparative discussion with the sparse coding procedure, it is convenient to note that the POD modes fundamentally represent the solution to the following minimization problem:²⁹

$$\min_{\Phi, S} \frac{1}{2} \|\mathbf{Q} - \Phi S\|_F^2 \quad \text{such that } \|\Phi_i\| \leq 1 \text{ for all } i \quad (3)$$

The POD modes are arranged in the descending order of eigenvalues, where the first few modes represent the most energetic structures in the snapshots data.⁸ Note that the complete set of POD modes exactly reproduces the snapshot matrix.

II.B.2. Sparse Coding

In its general formulation, sparse coding aims to solve the following minimization problem:

$$\min_{S, \Phi} \frac{1}{2} \sum_{k=1}^m \left(\|\mathbf{q}_k - \Phi s_k\|_F^2 + \beta \|s_k\|_0 \right), \quad \|\Phi_i\| \leq 1 \text{ for all } i \quad (4)$$

where the columns of Φ are the sparse modes; S is a matrix of activation coefficients; each column s_k is encouraged to be sparse by a penalty on the L0 “norm” of s_k (the count of the non-zero elements of the vector); and $\beta > 0$ is the regularization (or penalty) parameter, denoted here as the *sparsity coefficient*. Note that if the penalty term is ignored, Eqs.(4) reduces to the PCA problem (Eq.(3)) and yields the POD modes. The L0 penalty forces some of the entries in s_k to be zero, thereby resulting in a sparse coefficient matrix. The level of sparsity is controlled using β , where increasing the value of β increases the number of zero entries in s_k , thus producing a sparser coefficient matrix S . Reference 21 recommends that a value of $0 < \beta < 0.5$ is typically adequate for most cases.

In practice, the solution of Eq.(4) is challenging due to its non-convexity,³⁰ both due to the fact that the objective is not jointly convex in Φ and S , and due to the non-convexity of the L0 norm. To address the latter problem, the L0 penalty in Eq.(4) is replaced with a L1 penalty, making the problem convex in nature.³⁰ Optimization problem is then given as:

$$\min_{S, \Phi} \frac{1}{2} \sum_{k=1}^m \left(\|\mathbf{q}_k - \Phi s_k\|_F^2 + \beta \|s_k\|_1 \right), \quad \|\Phi_i\| \leq 1 \text{ for all } i \quad (5)$$

where $\|\cdot\|_1$ is the L1 norm. Although this problem is still not jointly convex in S and Φ , it can be approximately optimized by alternating minimization over these two matrices, which lead to convex problems in both cases. In particular, optimizing Eq.(5) over S is known as a least absolute shrinkage and selection operator (LASSO) problem,³¹ and has been widely studied by the Statistics and Machine Learning communities, among others.^{26,31,32} A number of algorithms have been devised to solve Eq.(5). In the current study, an algorithm²⁶ based on the coordinate descent technique is adopted. Likewise, optimizing Equation (5)

over Φ is a constrained least-squares problem, and is solved by using the Method of Optimal Directions (MOD) approach.³³ The MOD approach is equivalent to post-multiplying Q with the pseudo-inverse of matrix S to obtain the updated dictionary elements.

It is important to recognize that the use of L1 penalty eliminates the condition of orthogonality of the modes. Consequently, the sparse modes are not ‘ordered’ in terms of energy content. Therefore, the desired number of sparse modes are generated by varying β between 0 and 0.5 in a trial and error manner. Arbitrary selection of the modes from a larger set is not recommended.

II.B.3. Galerkin Projection

The reduced order solution to the unsteady fluid system is obtained by computing the time histories of the modal weights (also called prediction coefficients) using a Galerkin projection framework. In Galerkin projection,¹⁻⁴ the governing partial differential equations are projected onto the space spanned by a set of basis functions to yield a system of ordinary differential equations. The implementation of Galerkin projection involves the following steps. First, the fluctuating component of each snapshot (q_k) is expanded as a linear combination of reduced order modes as shown in Eq.(2). Next, the expansion is substituted into the governing equations, Eq.(1). Subsequently, the residual term is minimized by constraining it to be orthogonal to the space spanned by the modes. This process results in following set of ordinary differential equations:

$$\left(\Phi_i, \frac{\partial(\tilde{u} + \bar{U})}{\partial t} + ((\tilde{u} + \bar{U}) \cdot \nabla) (\tilde{u} + \bar{U}) \right) = -\frac{1}{Re} (\nabla \Phi_i, \nabla(\tilde{u} + \bar{U})) + \frac{1}{Re} [\Phi_i \nabla(\tilde{u} + \bar{U})] \quad (6)$$

where $i = 1, 2, \dots, N$. The dot product between any vectors f and g is given as,

$$(f, g) = \int_{\Omega} f \cdot g d\Omega \quad (7)$$

and the operation $[.]$ is defined as,

$$[\Phi_i \nabla(\tilde{u} + \bar{U})] = \oint_A \Phi_i \frac{\partial(\tilde{u} + \bar{U})}{\partial \hat{n}} dA \quad (8)$$

The resulting system of ordinary differential equations is then time marched using the fourth order Runge-Kutta scheme to obtain the prediction coefficients. In the case studies presented in this paper, including the pressure terms in the Galerkin projection approach is established to have a negligible effect on the computed velocity field.³⁴⁻³⁶ Therefore, the projection of pressure was ignored when computing the solution.

III. Applications

Three case studies describing the application of POD and sparse approaches to model 2-D incompressible unsteady flow are presented. The first study examines the suitability of sparse and POD modes to capture unseen flow conditions. The second and third case studies compare the performance of ROMs generated by using POD and sparse modes within a Galerkin projection framework. Galerkin projection models are developed for a periodic flow past a stationary cylinder and a chaotic flow inside a lid-driven cavity.

III.A. Robustness to Changing Flow Conditions

The capability of POD and sparse modes in predicting *unseen* flow behavior is examined. Here, unseen implies a flow condition that is outside the observed range of parameters extracted from the available data. This is viewed as an important property in the context of highly unsteady flows, since it is expected that observed data for extracting bases will generally be limited in scope (i.e., sparse parameter sweeps and/or time records of flow response). Thus, the predictions involve extrapolation using the basis vectors to the unseen flow conditions. The POD and sparse modes are evaluated over a set of observed flow states, i.e., the snapshot matrix. An unseen flow snapshot, which is not present in the snapshot matrix, and is significantly

different from the observed flow states, was projected onto the evaluated modes. Errors in projection were compared for the POD and sparse modes.

The Unsteady RANS equations were solved using CFL3D code for two cases: flow past a stationary cylinder, and flow past a cylinder oscillating in the y (vertical) direction. The freestream velocity is of Mach 0.2 and the $Re = 100$ for both the cases. The snapshot matrix consists of a total of 100 snapshots, 50 of which are chosen from the stationary cylinder case, and 50 from the first half cycle of the cylinder oscillation. Note that the snapshot matrix was constructed using absolute velocity as the state variable. A total of 100 POD modes and 30 sparse modes were evaluated from the snapshot matrix, a few of which are shown in Fig. 1. The POD approach produces a set of orthogonal modes, that are ordered from highest to lowest energy. The spatial variance of the POD modes increases with increasing the mode number. Note that this behavior is not physically representative of the higher order features that occur in the flow field, and is purely a consequence of orthogonality in the modes. In contrast, sparse modes resemble the snapshots of the original flow field and multi-scale features.

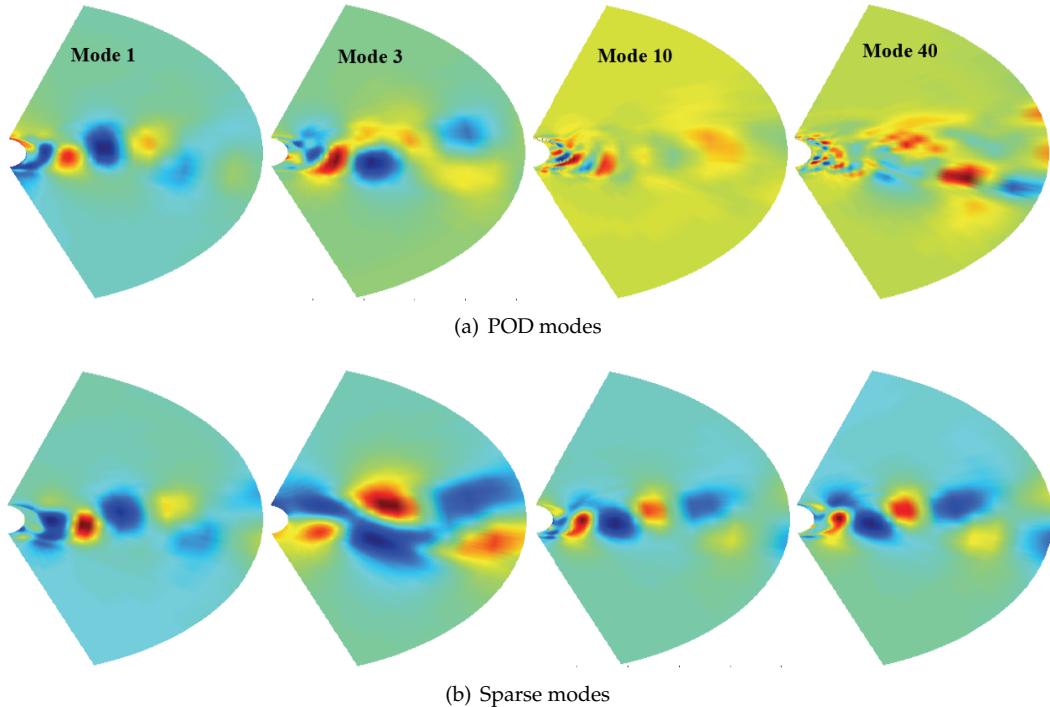


Figure 1. Comparison between a subset of POD and sparse modes.

An unseen flow snapshot is extracted at an arbitrary time during the second half cycle of the oscillating cylinder. This snapshot is then projected onto the POD and sparse modes. The modes are chosen in the following manner,

1. Truncating the POD modes to a desired number.
2. Using regression with an added L1 penalty (L1 regularization) for both POD and sparse modes by solving the optimization problem given in Eq. (5).

The L1 regularization ensures that the best modes from the dictionary are selected for a given unseen snapshot. The number of selected modes is inversely proportional to the value of the sparsity parameter (β). The RMS errors in projection of the unseen flow snapshot onto the chosen subset of modes are provided in Table 1. The error values converge to approximately 2% as the number of modes is increased. The unseen flow snapshot is described as a linear combination of selected modes in Fig. 2. Note that the error in representation is 27.57% when using 3 sparse modes, 84.09% and 52.52 % for 3 POD modes with direct truncation and L1 regularization, respectively. Furthermore, when using L1 regularization, POD modes 1, 5 and 6 are found to be active, and 2, 3 and 4 are inactive. This is in contrast to *a priori* truncation wherein the modes 1-3 are active and rest are inactive. This demonstrates that *a priori* truncation of POD modes is not ideal for flow conditions that deviate from the observed set. The sparse modes produce lower errors

when compared to the POD modes, both when the POD modes are truncated, and when the POD modes are selected using L1 regularization. Furthermore, adequate resolution of the flow field was achieved using only a limited number of sparse modes, indicating that these modes may provide better representation of the unseen flow fields over a wide range of parameters.

Table 1. Comparison between RMS errors in projection of an unseen flow snapshot- direct truncation of POD modes, and L1 regularization over the POD and sparse modes.

| Number of modes | POD error, direct truncation (%) | POD error, L1 regularization (%) | Sparse error, L1 regularization (%) |
|-----------------|----------------------------------|----------------------------------|-------------------------------------|
| 3 | 84.09 | 52.52 | 27.57 |
| 5 | 62.07 | 32.22 | 19.14 |
| 7 | 24.20 | 19.90 | 3.02 |
| 10 | 20.94 | 5.23 | 2.96 |
| 15 | 2.26 | 2.22 | 2.08 |

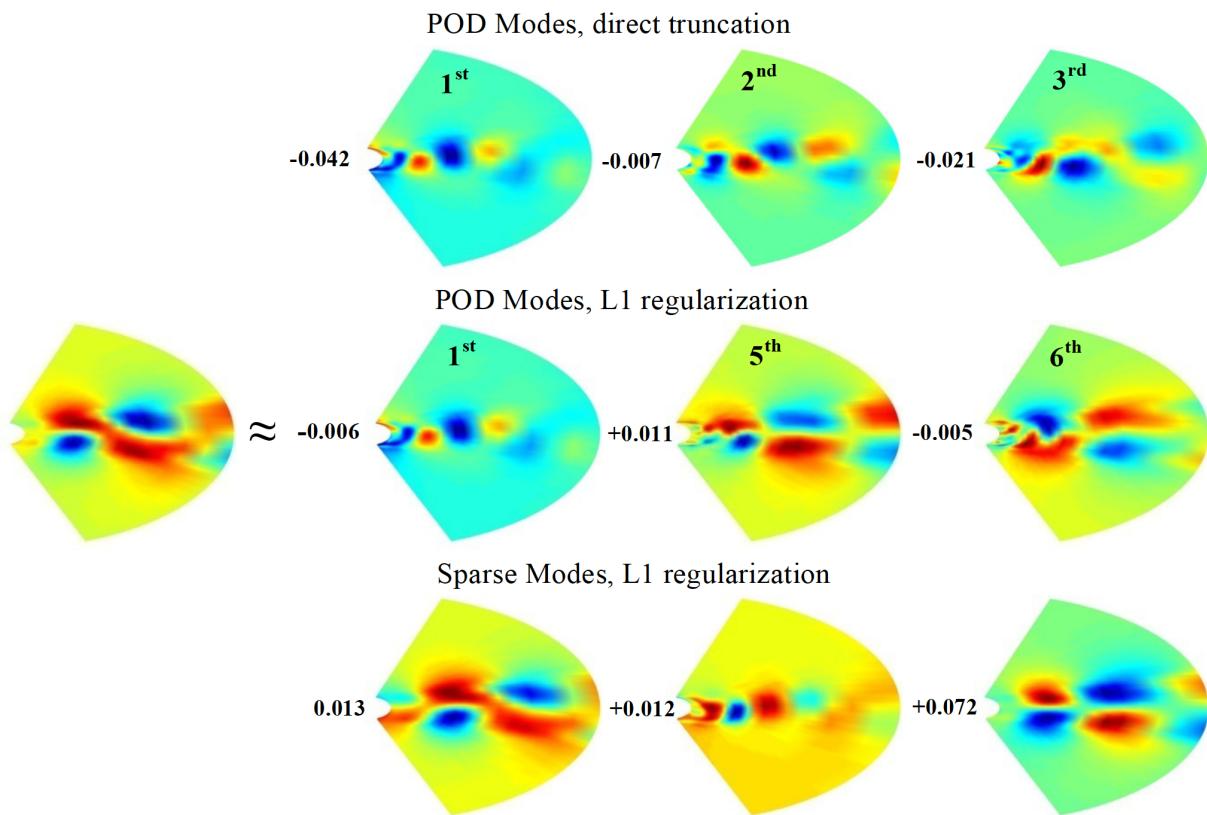


Figure 2. Unseen flow snapshot as a linear combination of POD and sparse modes for direct truncation and L1 regularization.

III.B. Flow Past a Cylinder

The application of sparse and POD modes within a Galerkin Projection framework is discussed next for the prediction of unsteady flow past a stationary cylinder. The full order solution is obtained from the PICar3D DNS solver using a 434×218 stretching grid at a Re of 200. A total of 480 uniformly sampled observations are collected from 20 units of non-dimensional time once the flow establishes a periodic response. The

snapshot matrix is constructed using observations numbered 1 to 240. Thus, the ROMs are constructed using the data from the first 10 time units in the flow solution after periodicity is achieved.

The velocity components in the POD modes numbered 1, 3, 5, 8 and 10, and five sparse modes taken arbitrarily from a set of 10 sparse modes are shown in Figs. 3 and 4, respectively. For reference, the velocity components from the full order solutions are displayed in Fig. 5. It is evident that the sparse modes are qualitatively similar to the full-order solutions, whereas the spatial variation in the POD modes increases as the mode number increases due to the orthogonality property.

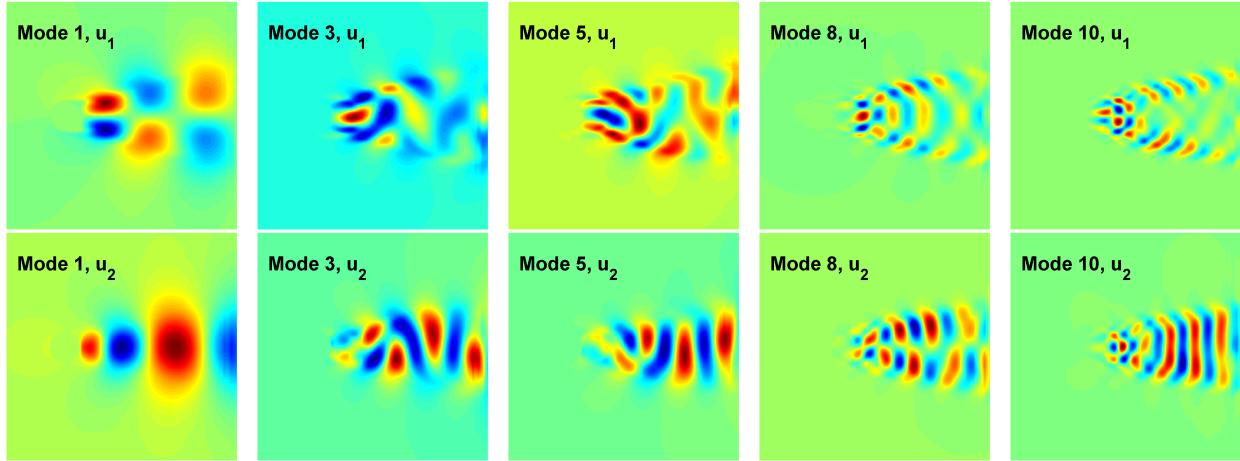


Figure 3. The u_1 and u_2 components of the POD modes numbered 1, 3, 5, 8 and 10 of the flow past a cylinder.

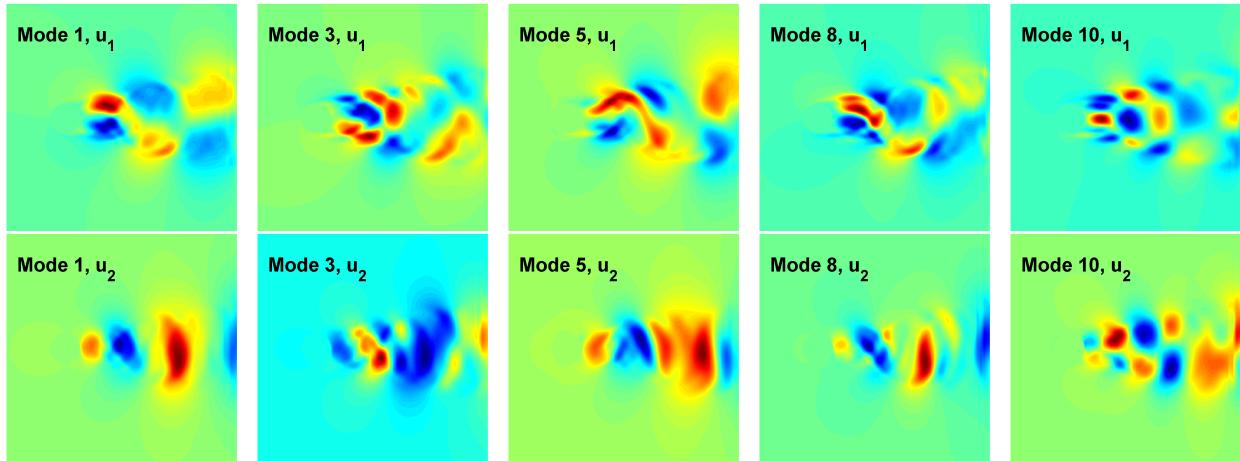


Figure 4. The u_1 and u_2 components of five sparse modes from a 10-mode sparse model of the flow past a cylinder.

The energy content in the modes is computed as the sum of the squares of the projection coefficients corresponding to the snapshot matrix. The energies of the modes are shown in Fig. 6. The energy contained in the POD decreases rapidly with increase in mode number, with modes 1 and 2 accounting for over 90% of the total energy. In comparison, the energy distribution in the sparse modes is much more uniform, or "leveled", and does not exhibit any monotonic trend. This is because the POD modes are arranged in the order of decreasing energy content; consequentially, the lower order POD modes contain majority of the dominant, high-energy structures. On the other hand, all the sparse modes contain multi-scale features, and therefore capture relatively equal amount of flow energy.

The time varying projection coefficients of the sparse and POD modes for first ten units of time history, are shown in Fig. 7. The POD coefficients are almost always non-zero. However, a subset of the sparse coefficients are zero for a given snapshot. This indicates that only a subset of sparse modes is active at

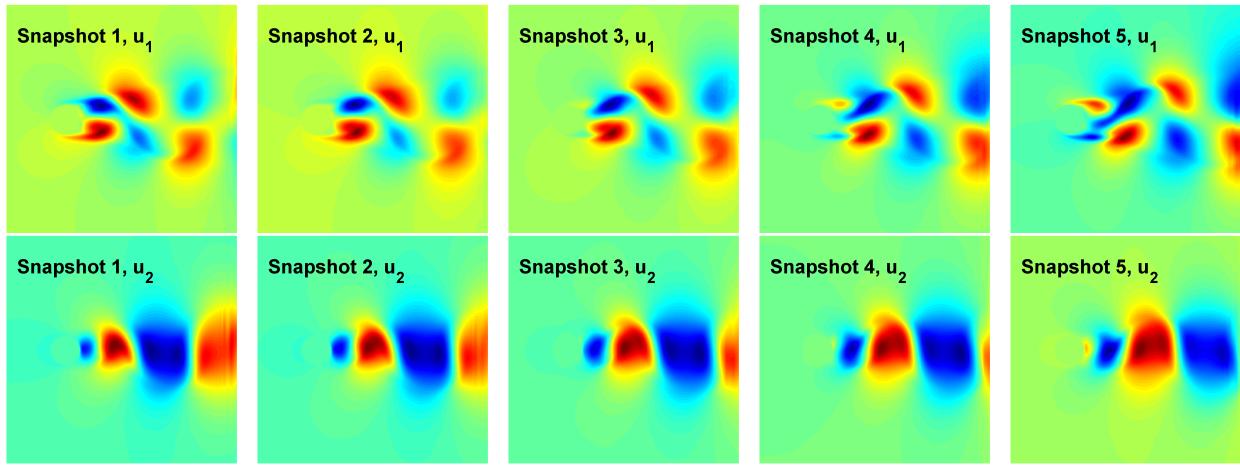


Figure 5. The u_1 and u_2 components of five different snapshots of the flow past a cylinder.

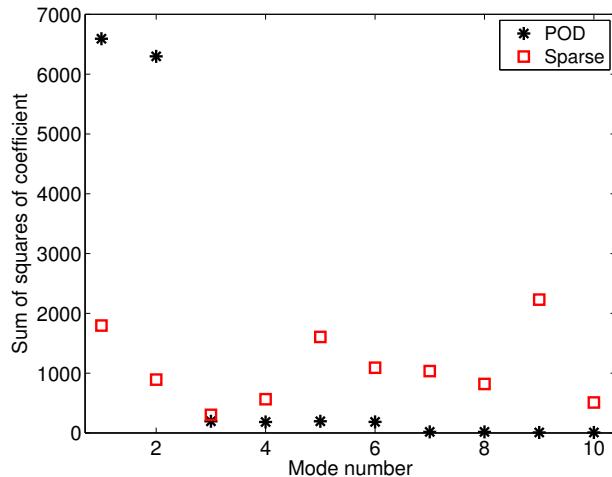


Figure 6. Energy content in the first 10 POD modes and the sparse modes of a 10-mode model of the flow past a cylinder, expressed in terms of the sum of squares of projection coefficients evaluated over 10 time units.

given time. The active sparse modes within a small window in time are therefore more tailored to the dynamics of the system in that time window. On the contrary, the POD approach is not designed to capture the local dynamics; this is because the POD modes are optimal in terms of capturing energy, and are always active.

The flow predictions obtained from the ROMs are compared to the full order solutions for the 20 time units over which the high-fidelity data are available. The errors in the fluctuating component are quantified as:³⁵

$$E(t) = 100 \frac{\|\mathbf{q}(\mathbf{x}, t) - \tilde{\mathbf{u}}(\mathbf{x}, t)\|}{\|\mathbf{q}(\mathbf{x}, t)\|} \% \quad (9)$$

where $\|\cdot\|$ is the L2 norm. The errors in POD and sparse predictions are shown in Fig. 8. Additionally, the maximum and time-averaged errors in the ROMs are listed in Table 2. The E_{max} for the 5-mode POD model is 29.26%, which is approximately three times that of the 5-modes sparse model. Moreover, the error in the 5 mode sparse model is comparable to errors in the 10 to 50-mode POD models. When using the same number of modes, the POD prediction deviate from the original solution much more rapidly compared to the sparse predictions. It is observed that the errors in both POD and sparse predictions converge as the number of the modes is increased. The accuracy of the two sets of predictions deteriorates as the solution is marched forward in time. However, within the region of available data, the sparse ROMs yielded better

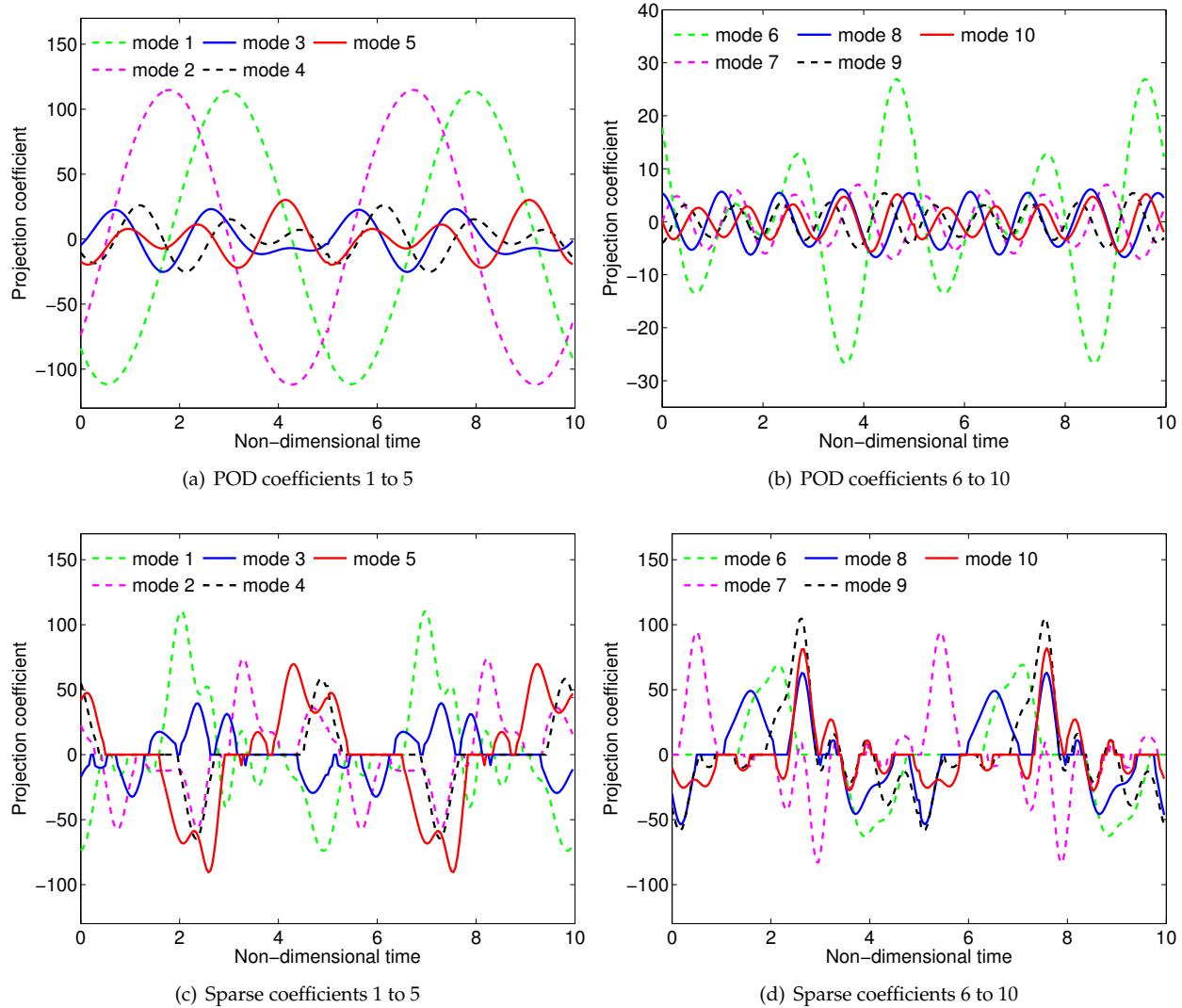


Figure 7. The time history of projection coefficients corresponding to the 10-mode POD and 10-mode sparse model of the flow past a cylinder.

agreement with the original solution compared to the POD ROMs.

The instantaneous kinetic energy due to the fluctuating component of velocity is shown in Fig. 9. In general, the POD ROMs over-predict the energy whereas the sparse ROMs under-predict. In particular, the 5-mode POD model yields qualitatively different behavior compared to the other ROMs. It over-predicts the energy by approximately 30%, whereas other ROMs predict the energy within a $\pm 5\%$ of discrepancy, averaged over first 20 time units. Furthermore, the phase angle in the instantaneous fluctuating kinetic energy evolution for the POD models deviates faster as compared to the sparse models with same number of modes. The faster deviation in phase along with over-prediction of kinetic energy results in larger errors in predicted values from the POD ROMs. The power spectral densities (PSD) of the instantaneous fluctuating kinetic energy are shown in Fig. 10. The ROMs capture the location of peaks accurately. As expected, the POD models over-predict the energy values, in particular for frequencies less than 0.2 Hz, which represents the frequency of vortex shedding. This behavior is attributed to the absence of the low-energy, small-scale structures which are responsible for energy dissipation. The sparse ROMs marginally over-predict the energy in the low-frequency range and under-predict the energy at higher frequencies, which implies that the sparse ROMs are over-dissipative in the high frequency ranges.

The computational costs associated with the POD and sparse ROMs are listed in Table 3 and Table 4, respectively. The simulations are performed on a Intel(R) Xeon(R) E5-2620 v2 @ 2.10GHz processor. The

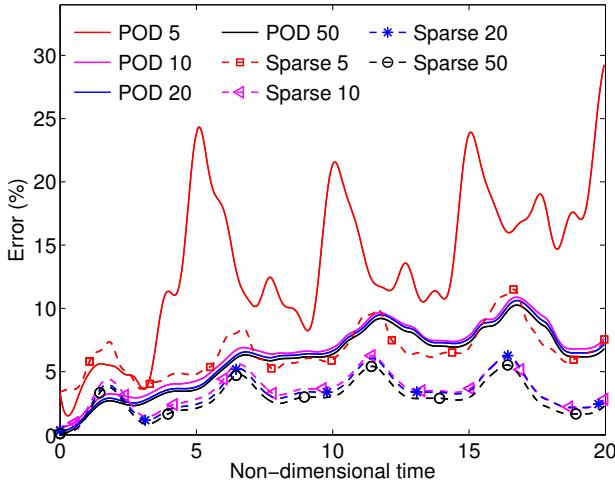


Figure 8. Prediction errors of the POD and sparse-Galerkin ROMs for the flow past a cylinder. The snapshot matrix used to compute the modes spans first 10 time units, and the errors are computed over first 20 units.

Table 2. The maximum and time-averaged errors in the POD and sparse ROMs for the flow past a cylinder.

| Number of modes | POD ROM error (%) | | Sparse ROM error (%) | |
|-----------------|-------------------|---------------|----------------------|---------------|
| | Maximum | Time averaged | Maximum | Time averaged |
| 5 | 29.26 | 13.43 | 11.50 | 6.61 |
| 10 | 10.88 | 6.49 | 6.27 | 3.57 |
| 20 | 10.62 | 6.21 | 6.16 | 3.25 |
| 50 | 10.25 | 5.95 | 5.52 | 2.91 |

computations of the sparse modes is approximately five times more expensive than computation of the POD modes. The costs to generate the Galerkin matrices and to carry out the ROM integration are comparable for the POD and sparse ROMs. The cost for generating the full-order DNS data is 4000 s, which is approximately 17 times slower than the most expensive sparse model. An aspect to note from the computational costs of the ROMs is the nonlinear increase in cost with increasing number of modes. This highlights the importance of constructing a reduced order model with the fewest number of modes possible; which is obtained for this case using a sparse coding procedure.

Table 3. Computational costs associated with several POD ROMs developed for the flow past a cylinder case. Note that the modes are computed using snapshot matrix spanning first 10 time units; whereas, the ROMs are integrated for 100 time units.

| Number of modes | Modes calculation (s) | Galerkin matrices (s) | ROM integration (s) | Total (s) |
|-----------------|-----------------------|-----------------------|---------------------|-----------|
| 5 | 18.59 | 0.06 | 0.52 | 19.17 |
| 10 | 23.61 | 0.28 | 1.00 | 24.89 |
| 20 | 27.46 | 1.53 | 2.38 | 31.37 |
| 50 | 30.14 | 80.56 | 13.42 | 124.12 |

III.C. Lid-driven Cavity

The unsteady response of a fluid enclosed in a cavity and actuated using a moving surface (lid) is examined. This configuration, widely referred to as ‘the lid-driven cavity flow’ has been used in several studies^{11,37,38}

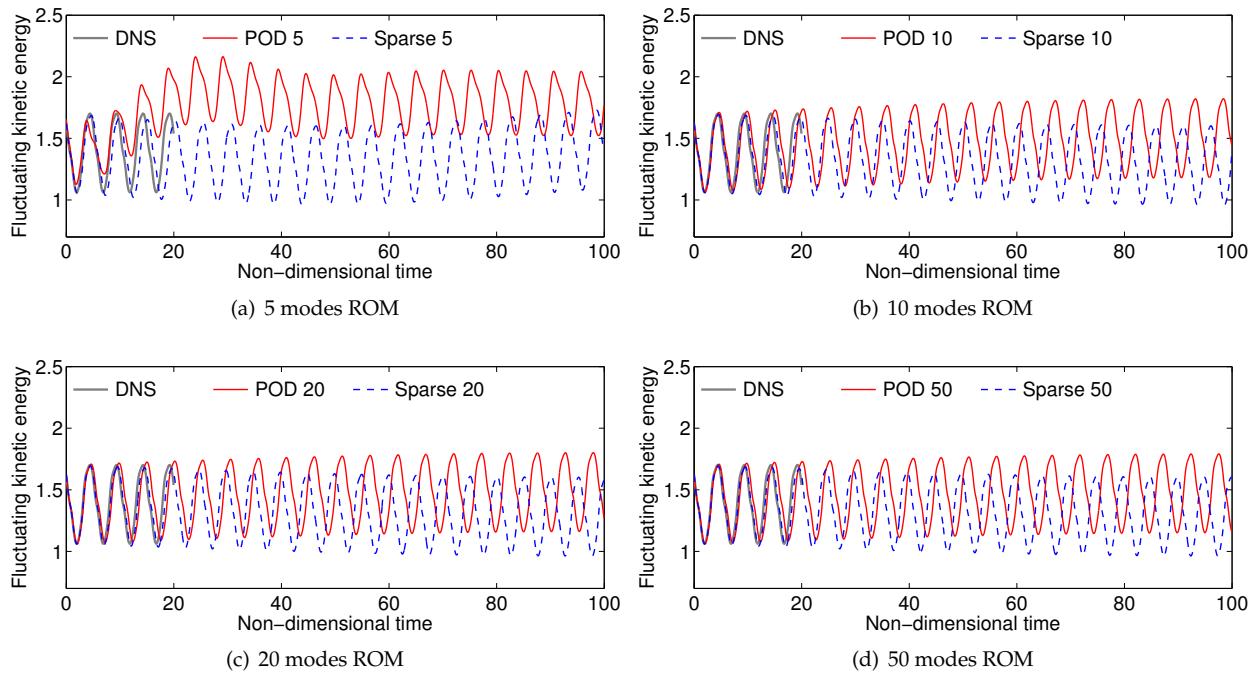


Figure 9. Time history of the instantaneous fluctuating kinetic energy of the flow past a cylinder as predicted by DNS, and the computed POD and sparse ROMs. The high-fidelity data are available for first 20 time units, and the ROMs are integrated for 100 units.

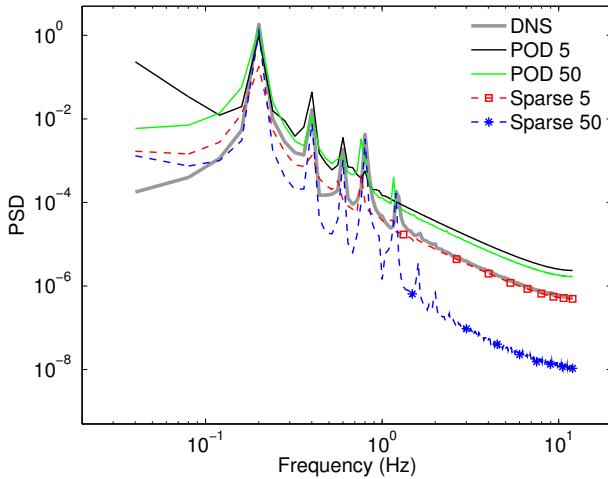


Figure 10. PSD of the fluctuating kinetic energy of the flow past a cylinder as predicted by DNS, and the 5 and 50 mode POD and sparse ROMs.

for benchmarking purposes. The cavity considered in this study is a 2-D square, enclosed region, with three rigid stationary walls and one rigid lid translating in x direction with constant velocity. The lid velocity is prescribed as $(1 - (2x - 1)^2)^2$, where x varies from 0 to 1. The DNS solutions are obtained using a 512×512 uniform grid at Re of 30,000, where the Re is computed based on the maximum lid velocity. A total of 10,000 uniformly sampled observations are collected from 100 units of non-dimensional time once the flow had reached a statistically stationary state. The snapshot matrix is constructed using first 1250 observations. Thus, the ROMs are constructed using the data from the first 12.5 time units.

The velocity components of the POD modes numbered 1, 5, 10 and 50, and four modes taken arbitrarily from a set of 10 sparse modes are shown in Figs. 11 and 12, respectively. For reference, the velocity components from the full order solution are displayed in Fig. 13. The characteristics of the POD and sparse modes obtained in this case are very similar to those in the previous cases of flow past a cylinder. The spatial wave

Table 4. Computational costs associated with several sparse ROMs developed for the flow past a cylinder case. Note that the modes are computed using snapshot matrix spanning first 10 time units; whereas, the ROMs are integrated for 100 time units.

| Number of modes | Modes calculation (s) | Galerkin matrices (s) | ROM integration (s) | Total (s) |
|-----------------|-----------------------|-----------------------|---------------------|-----------|
| 5 | 92.84 | 0.08 | 0.61 | 93.53 |
| 10 | 106.53 | 0.3 | 1.13 | 107.96 |
| 20 | 127.88 | 1.86 | 2.44 | 132.18 |
| 50 | 142.38 | 73.02 | 11.45 | 226.85 |

number in the POD modes increases with the mode number. Additionally, the small-scale flow structures in higher order modes are arranged in a localized spatial order. There is no particular order in the sparse modes in terms of spatial wave numbers. The small-scale structures captured in sparse modes more closely resemble with the original snapshots, and lack any spatially-localized order.

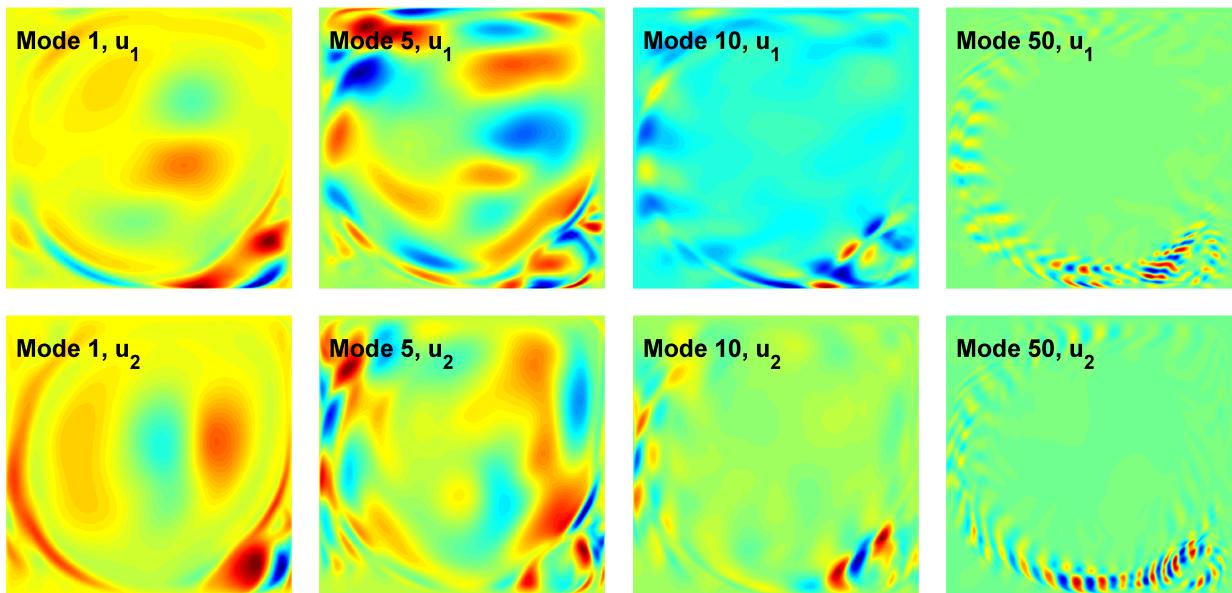


Figure 11. The u_1 and u_2 components of the POD modes numbered 1, 5, 10 and 50 of the lid-driven cavity.

The error in POD and sparse predictions are provided in Fig. 14. Additionally, the maximum and time-averaged errors in the ROMs are listed in Table 5. The sparse ROMs with 20 or more number of modes predict the flow states with maximum error (E_{max}) of less than 100%, whereas the maximum error in POD ROMs is greater than 100% for all the models. The 5-mode POD model provides the worst prediction with errors as high as 2282.29%, whereas the maximum error of the 5-mode sparse model is 294.55%. Moreover, the E_{max} for the 20-mode POD model is 333.62% which is higher than that of the 5-mode sparse model. There is a drastic increase in ROM prediction error after the first 12.5 units of time. This behavior is expected as the modes are computed using first 12.5 units of high-fidelity data. However, the errors in the POD ROMs increase at a much faster rate compared to the sparse ROMs. The rate decreases with increase of the number of modes. Adding modes to the ROMs produces a greater improvement in the POD ROM compared to the sparse ROM.

The instantaneous turbulent kinetic energy predicted by the ROMs is shown in Fig. 15. When the same number of modes are used, the sparse ROMs capture the energy levels with greater accuracy compared to the POD ROMs. The 5-mode POD model over-predicts the energy by a factor of 8.3, averaged over first 100 time units, whereas, the 5-mode sparse model over-predicts the energy by a factor of 1.8 over the

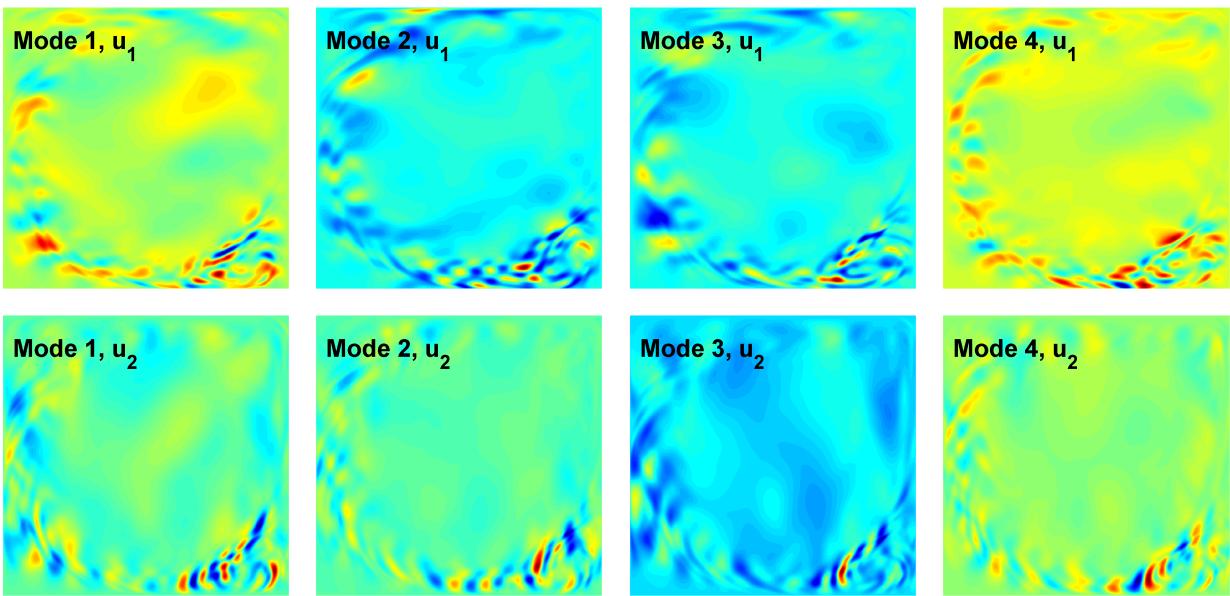


Figure 12. The u_1 and u_2 components of four of the sparse modes of the lid-driven cavity.

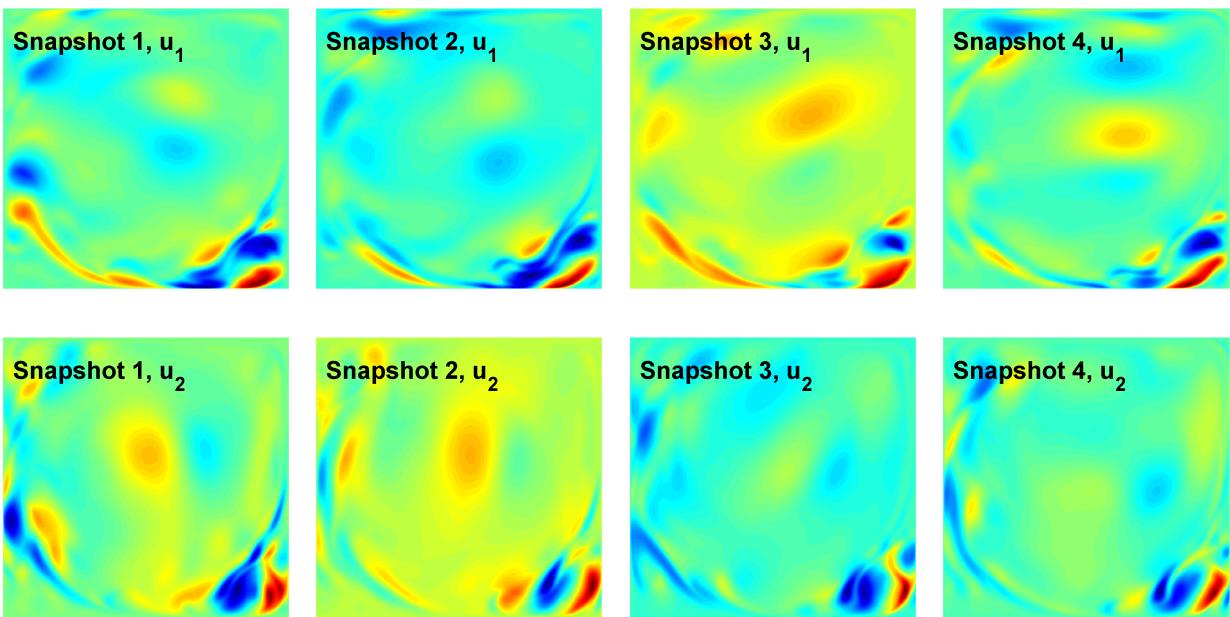


Figure 13. The u_1 and u_2 components of four different snapshots of the lid-driven cavity.

same duration. Similar to the trends observed in the error values, the 5-mode sparse model captures the energy levels more accurately as compared to 20-mode POD model. Furthermore, it is observed that the higher order sparse models under-predict the energy between $t = 20$ units to $t = 50$ units. The predicted energy values stabilize at the levels of the full-order model past $t = 50$ units. Overall, the 50 and 100 mode sparse, and 100-mode POD models capture the average energy levels with reasonable accuracy. These results indicate that ignoring small-scale, low-energy information in a truncated POD set may result in over-prediction of turbulent kinetic energy, and therefore result in an inaccurate flow prediction. The sparse modes include the low energy, dynamically important, features that provide the appropriate dissipation to

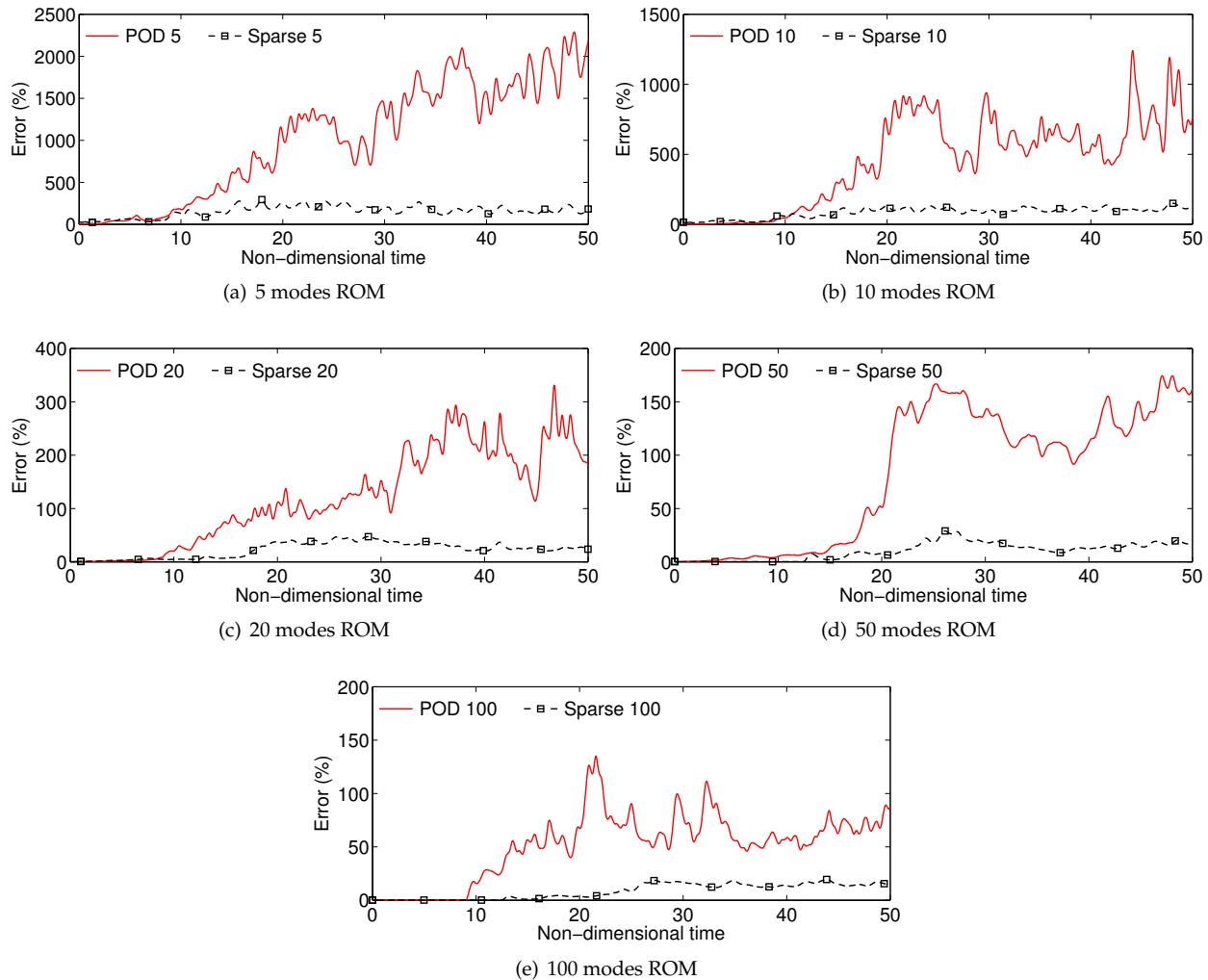


Figure 14. Error in prediction by the POD and sparse ROMs of the lid-driven cavity. The snapshot matrix used to compute the modes spans first 12.5 time units, and the errors are computed over first 50 units.

balance the large scale dominant structures.

In order to investigate the general stability of the constructed ROMs, the time-average accumulated energy is computed and shown Fig. 16. The value of the time-averaged accumulated energy (η) at time t is given as:

$$\eta(t) = \frac{1}{t} \int_{t=0}^t \|\mathbf{q}(\mathbf{x}, t)\| dt \quad (10)$$

where $\|\cdot\|$ is the L2 norm. It is found that the POD models over-predict the $\eta(t)$, and the $\eta(t)$ decreases with an increase in the number of POD modes. On the contrary, the sparse predictions do not exhibit any identifiable trend. Overall, the time-averaged accumulated energy levels do not show a rapid increase over simulation times of 500 units, suggesting that the ROMs are stable over an extended time simulation.

Next, the PSDs of the instantaneous turbulent kinetic energy are shown in Fig. 17. As demonstrated in the previous set of results, except for 100-mode POD model, all the other POD models significantly over-predict the energy values over the entire range of frequencies, whereas the sparse ROMs provide a good overall prediction.

The computational costs associated with the POD and sparse ROMs are listed in Table 6 and Table 7, respectively. The computations of the sparse modes is one order of magnitude more expensive than generating the POD modes. The costs to generate the Galerkin matrices and to carry out the ROM integration for POD and sparse ROMs are comparable. The computational cost for generating the DNS data is 432,000 s

Table 5. The maximum and time-averaged errors in the POD and sparse ROMs for the lid-driven cavity.

| Number of modes | POD ROM error (%) | | Sparse ROM error (%) | |
|--------------------|----------------------|---------------|-------------------------|---------------|
| | Maximum | Time-averaged | Maximum | Time-averaged |
| 5 | 2282.29 | 1016.82 | 294.55 | 155.82 |
| 10 | 1240.91 | 462.23 | 149.92 | 84.19 |
| 20 | 333.62 | 119.92 | 47.27 | 22.63 |
| 50 | 174.44 | 83.82 | 28.98 | 10.27 |
| 100 | 134.97 | 51.31 | 19.43 | 8.51 |

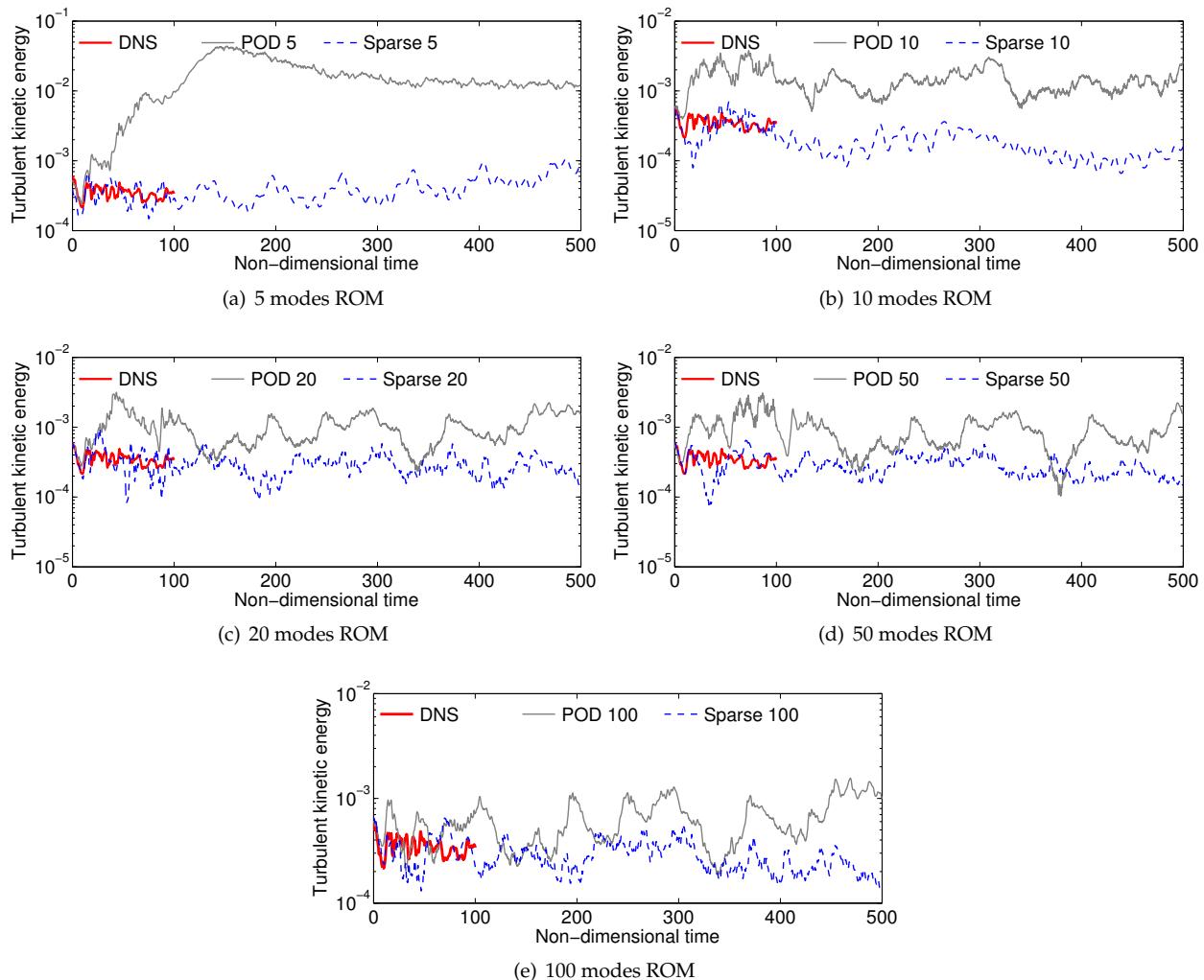


Figure 15. Time history of the instantaneous turbulent kinetic energy of the lid-driven cavity as predicted by DNS, and the computed POD and sparse ROMs. The snapshot matrix used to compute the modes spans first 12.5 time units, ROMs are integrated for 500 time units, and the high-fidelity data are available for first 100 time units.

(or 120 hours), which is approximately 20 times the computational cost associated with the most expensive sparse ROM. Note that, the cost of integrating a ROM is negligible as compared to the cost of building one. The ratio of the cost associated with a ROM to the one associated with the full order model is therefore inversely proportional to the physical integration time.

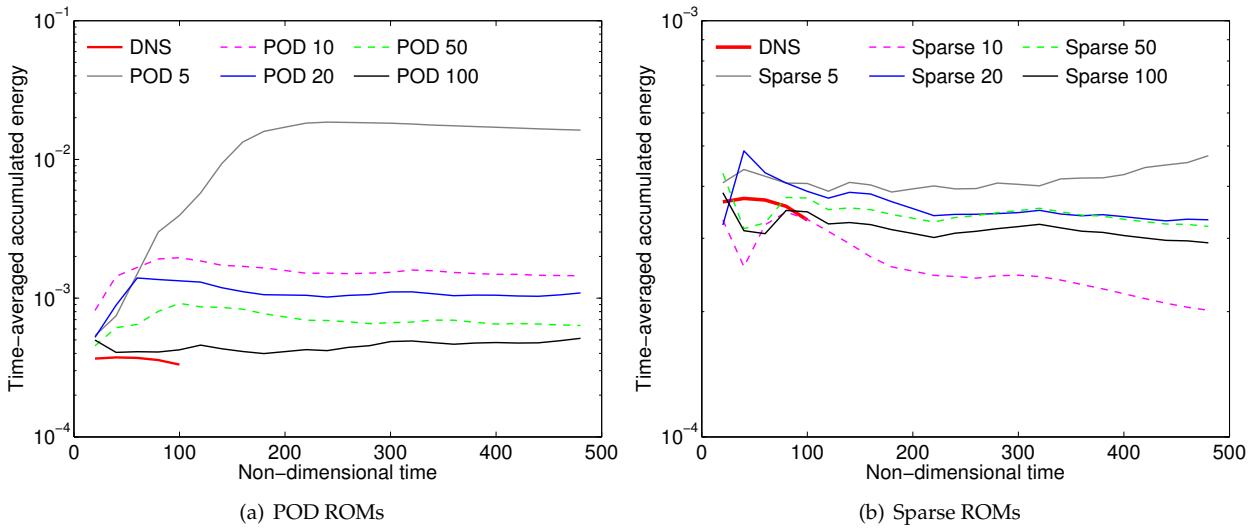


Figure 16. The time-averaged accumulative energy for several ROMs and the DNS.

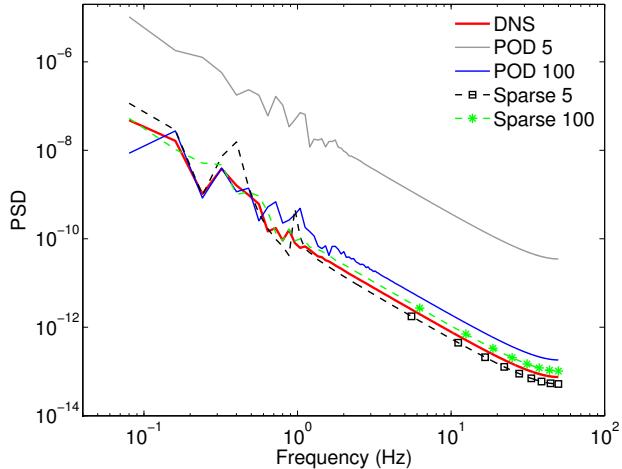


Figure 17. PSD of the turbulent kinetic energy of the lid-driven cavity as predicted by DNS, and the 5 and 100-mode POD and sparse ROMs.

Table 6. Computational costs associated with several POD ROMs developed for the lid-driven cavity. Note that the modes are computed using a snapshot matrix spanning first 12.5 time units; whereas, the ROMs are integrated for 500 time units.

| Number of modes | Modes calculation (s) | Galerkin matrices (s) | ROM integration (s) | Total (s) |
|-----------------|-----------------------|-----------------------|---------------------|-----------|
| 5 | 547.24 | 0.22 | 3.10 | 550.56 |
| 10 | 579.37 | 1.24 | 5.28 | 585.89 |
| 20 | 632.56 | 21.74 | 16.47 | 670.77 |
| 50 | 768.37 | 347.05 | 54.34 | 1169.76 |
| 100 | 936.35 | 3693.05 | 355.31 | 4987.71 |

IV. Conclusions

The generation of the sparse and POD modes is described with three examples, a URANS database of 2-D flows past stationary and a moving cylinders, a DNS database of a flow past a stationary cylinder,

Table 7. Computational costs associated with several sparse ROMs developed for the lid-driven cavity. Note that the modes are computed using a snapshot matrix spanning first 12.5 time units; whereas, the ROMs are integrated for 500 time units.

| Number of modes | Modes calculation (s) | Galerkin matrices (s) | ROM integration (s) | Total (s) |
|-----------------|-----------------------|-----------------------|---------------------|-----------|
| 5 | 3206.28 | 0.31 | 3.37 | 3209.96 |
| 10 | 3938.71 | 1.26 | 5.35 | 3945.32 |
| 20 | 5139.25 | 21.02 | 15.82 | 5176.09 |
| 50 | 14303.36 | 326.82 | 53.66 | 14683.84 |
| 100 | 17380.94 | 3289.23 | 347.30 | 21017.47 |

and a DNS database of a flow inside a 2-D lid-driven cavity. The computed POD and sparse modes are compared in terms of the energy content in the modes, time history of the projection coefficients, and spatial structure of the u_1 and u_2 components of the modes. Several Galerkin projection based ROMs are developed for the flow past a cylinder and lid-driven cavity problems. Performances of the POD and sparse ROMs are compared in terms of prediction error, evolution of the turbulent kinetic energy, and the PSDs of the time histories of the energy. Additionally, robustness of the sparse modes to changing flow conditions is investigated using the 2-D flows past a stationary and a moving cylinder database. The results of these studies allow one to reach several useful conclusions, namely:

1. Sparse modes are inherently multi-scale. Conversely, POD modes are biased towards high-energy, dominant structures. Furthermore, the orthogonality of the POD modes introduces physically unrepresentative fine scale features in higher order modes.
2. In the sparse coding approach, only a subset of sparse modes are active at a given instant. This property results in modes which are more tailored to the local dynamics of the system. On the contrary, the POD process yields a set of modes active at all times, rendering the approach incapable of capturing the local dynamics.
3. Direct projection results indicate that a set of sparse modes generalizes better to unseen flow conditions as compared to a truncated POD set of same size. Fewer sparse modes are required to represent an unseen flow snapshot than the number of POD modes.
4. The sparse-Galerkin models out-perform the POD-Galerkin models in predicting the flow states in both the case studies.
5. The sparse ROMs provide more accurate predictions for turbulent kinetic energy as compared to the POD ROMs. The multi-scale nature of the sparse modes results in a balance between destabilizing dominant energy features and dissipation due to small-scale, low-energy features. On the contrary, a truncated set of POD modes is biased towards the high-energy features, causing over-prediction in the energy levels when used in Galerkin ROMs .
6. The instantaneous turbulent kinetic energy and the time-averaged accumulated energy trends indicate that the POD and sparse ROMs developed for the lid-driven cavity are stable.
7. The computational cost of computing sparse modes is up to 20 times the cost of computing the same number of POD modes. Despite this, the cost of the most expensive sparse-Galerkin models remain significantly smaller than that of solving the full-order model. This is the only drawback observed for the sparse coding basis identification approach.

Acknowledgments

The authors gratefully acknowledge the support of ONR grant N00014-14-1-0018, under the direction of Dr. Judah Milgram, an HPCMPO Frontier PETT Project Grant, under the direction of David Bartoe, and an allocation of computing time from the Ohio Supercomputer Center.

References

- ¹Holmes, P., Lumley, J. L., Berkooz, G., and Rowley, C., *Turbulence, coherent structures, dynamical systems and symmetry*, Cambridge university press, New York, 2nd ed., 2012.
- ²Ilak, M., Bagheri, S., Brandt, L., Rowley, C. W., and Henningson, D. S., "Model reduction of the nonlinear complex Ginzburg-Landau equation," *SIAM Journal on Applied Dynamical Systems*, Vol. 9, No. 4, 2010, pp. 1284–1302.
- ³Kalashnikova, I., van Bloemen W. B., Arunajatesan, S., and Barone, M., "Stabilization of projection-based reduced order models for linear time-invariant systems via optimization-based eigenvalue reassignment," *Computer Methods in Applied Mechanics and Engineering*, Vol. 272, 2014, pp. 251–270.
- ⁴Rowley, C., "Model reduction for fluids, using balanced proper orthogonal decomposition," *International Journal of Bifurcation and Chaos*, Vol. 15, No. 03, 2005, pp. 997–1013.
- ⁵Noack, B. R., Afanasiev, K., Morzyński, M., Tadmor, G., and Thiele, F., "A hierarchy of low-dimensional models for the transient and post-transient cylinder wake," *Journal of Fluid Mechanics*, Vol. 497, 2003, pp. 335–363.
- ⁶Noack, B. R. and Eckelmann, H., "A low-dimensional Galerkin method for the three-dimensional flow around a circular cylinder," *Physics of Fluids (1994-present)*, Vol. 6, No. 1, 1994, pp. 124–143.
- ⁷Lucia, D. J., Beran, P. S., and Silva, W. A., "Reduced-order modeling: new approaches for computational physics," *Progress in Aerospace Sciences*, Vol. 40, No. 1, 2004, pp. 51–117.
- ⁸Berkooz, G., Holmes, P., and Lumley, J. L., "The proper orthogonal decomposition in the analysis of turbulent flows," *Annual review of fluid mechanics*, Vol. 25, No. 1, 1993, pp. 539–575.
- ⁹Chatterjee, A., "An introduction to the proper orthogonal decomposition," *Current science*, Vol. 78, No. 7, 2000, pp. 808–817.
- ¹⁰Sirovich, L., "Turbulence and the dynamics of coherent structures. I-Coherent structures. II-Symmetries and transformations. III-Dynamics and scaling," *Quarterly of applied mathematics*, Vol. 45, 1987, pp. 561–571.
- ¹¹Balajewicz, M. J., Dowell, E. H., and Noack, B. R., "Low-dimensional modelling of high-Reynolds-number shear flows incorporating constraints from the Navier-Stokes equation," *Journal of Fluid Mechanics*, Vol. 729, 2013, pp. 285–308.
- ¹²Ilak, M. and Rowley, C. W., "Modeling of transitional channel flow using balanced proper orthogonal decomposition," *Physics of Fluids (1994-present)*, Vol. 20, No. 3, 2008, pp. 034103.
- ¹³Amsallem, D. and Farhat, C., "Stabilization of projection-based reduced-order models," *International Journal for Numerical Methods in Engineering*, Vol. 91, No. 4, 2012, pp. 358–377.
- ¹⁴Pope, S. B., *Turbulent Flows*, Cambridge University Press, New York, sixth ed., 2009.
- ¹⁵Amsallem, D., Zahr, M. J., and Farhat, C., "Nonlinear model order reduction based on local reduced-order bases," *International Journal for Numerical Methods in Engineering*, Vol. 92, No. 10, 2012, pp. 891–916.
- ¹⁶Zhou, K., Doyle, J. C., and Glover, K., *Robust and optimal control*, Prentice Hall, 1st ed., 1996.
- ¹⁷Ma, Z., Ahuja, S., and Rowley, C. W., "Reduced-order models for control of fluids using the eigensystem realization algorithm," *Theoretical and Computational Fluid Dynamics*, Vol. 25, No. 1-4, 2010, pp. 233–247.
- ¹⁸Olshausen, B. and David, J., "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, Vol. 381, No. 6583, 1996, pp. 607–609.
- ¹⁹Kreutz-Delgado, K., Murray, J. F., Rao, B. D., Engan, K., Lee, T., and Sejnowski, T. J., "Dictionary learning algorithms for sparse representation," *Neural computation*, Vol. 15, No. 2, 2003, pp. 349–396.
- ²⁰Yang, J., Yu, K., and Huang, T., "Efficient highly over-complete sparse coding using a mixture model," *Computer Vision–ECCV 2010*, Springer, 2010, pp. 113–126.
- ²¹Yang, J., Yu, K., Gong, Y., and Huang, T., "Linear spatial pyramid matching using sparse coding for image classification," *Computer Vision and Pattern Recognition*, IEEE, 2009, pp. 1794–1801.
- ²²Grosse, R., Raina, R., Kwong, H., and Ng, A., "Shift-invariance sparse coding for audio classification," *arXiv preprint arXiv:1206.5241*, 2012.
- ²³Lee, H., Battle, A., Raina, R., and Ng, A., "Efficient sparse coding algorithms," *Advances in neural information processing systems*, 2006, pp. 801–808.
- ²⁴Olshausen, B. and Field, D., "Sparse coding of sensory inputs," *Current opinion in neurobiology*, Vol. 14, No. 4, 2004, pp. 481–487.
- ²⁵Kolter, J., Batra, S., and Ng, A., "Energy disaggregation via discriminative sparse coding," *Advances in Neural Information Processing Systems*, 2010, pp. 1153–1161.
- ²⁶Friedman, J., Hastie, T., and Tibshirani, R., "Regularization paths for generalized linear models via coordinate descent," *Journal of statistical software*, Vol. 33, No. 1, 2010, pp. 1–22.
- ²⁷Krist, S., Biedron, R., and Rumsey, C., *CFL3D user's manual (version 5.0)*, Citeseer, 1998.
- ²⁸Mittal, R., Dong, H., Bozkurttas, M., Najjar, F., Vargas, A., and von Loebbecke, A., "A versatile sharp interface immersed boundary method for incompressible flows with complex boundaries," *Journal of computational physics*, Vol. 227, No. 10, 2008, pp. 4825–4852.
- ²⁹Abdi, H. and Williams, L., "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 2, No. 4, 2010, pp. 433–459.
- ³⁰Zuo, W., Meng, D., Zhang, L., Feng, X., and Zhang, D., "A generalized iterated shrinkage algorithm for non-convex sparse coding," *Proceedings of the IEEE international conference on computer vision*, 2013, pp. 217–224.
- ³¹Tibshirani, R., "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, pp. 267–288.
- ³²Friedman, J., Hastie, T., and Tibshirani, R., "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, Vol. 9, No. 3, 2008, pp. 432–441.
- ³³Engan, K., Aase, S. O., and Hakon H., J., "Method of optimal directions for frame design," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, IEEE, 1999, pp. 2443–2446.

³⁴Ullmann, S. and Lang, J., "A POD-Galerkin reduced model with updated coefficients for Smagorinsky LES," *Proceeding of V. European conference on computational fluid dynamics, ECCOMAS CFD*, 2010.

³⁵Graham, W., Peraire, J., and Tang, K., "Optimal control of vortex shedding using low-order models. Part I – open-loop model development," *International Journal for Numerical Methods in Engineering*, Vol. 44, No. 7, 1999, pp. 945–972.

³⁶Bergmann, M., Cordier, L., and Brancher, J., "Optimal rotary control of the cylinder wake using proper orthogonal decomposition reduced-order model," *Physics of Fluids (1994-present)*, Vol. 17, No. 9, 2005, pp. 097–101.

³⁷Leblond, C., Allery, C., and Inard, C., "An optimal projection method for the reduced-order modeling of incompressible flows," *Computer Methods in Applied Mechanics and Engineering*, Vol. 200, No. 33, 2011, pp. 2507–2527.

³⁸Terragni, F., Valero, E., and Vega, J., "Local POD plus Galerkin projection in the unsteady lid-driven cavity problem," *SIAM Journal on Scientific Computing*, Vol. 33, No. 6, 2011, pp. 3538–3561.