

Reduced Order Modeling of Turbulent Flows using Sparse Coding

Rohit Deshmukh*, Zongxian Liang[†] and Jack J. McNamara[‡]

Department of Mechanical and Aerospace Engineering

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

Basis identification is a critical step in the construction of accurate reduced order models using Galerkin projection. This is particularly challenging for turbulent flow fields due to the presence of multi-scale phenomena that cannot be ignored when building a reduced order model. The ubiquitous proper orthogonal decomposition approach seeks to truncate the basis spanning an observed data set into a small set of dominant modes, leading to loss of small scale information in turbulent flow fields. Ignoring the small scale information results in under-resolved rate of dissipation of energy, and consequently, over-prediction of kinetic energy by constructed reduced order models. This study focuses on this issue by exploring an approach known as sparse coding for the basis identification problem. The sparse coding approach seeks the best compact basis to span the entire data set, and capture the multi-scale features present in the turbulent flows. These approaches are demonstrated for a canonical problem of an incompressible flow inside a 2-D lid-driven cavity. Results indicate that Galerkin reduction of the governing equations using only a few sparse modes produces reasonably accurate predictions of second order statistics of the fluid flow. Additionally, the sparse modes based models are found to maintain balance between the production and dissipation of energy. Whereas, models constructed using the same numbers of proper orthogonal decomposition modes are found to perform poorly.

Nomenclature

A	= boundary of the computational domain
$[a_i]$	= Galerkin matrix, constant terms
$[b_{ij}]$	= Galerkin matrix, linear terms
$[c_{ijk}]$	= Galerkin matrix, quadratic terms
$[d_{ij}]$	= matrix containing correlations of reduced order modes
m	= total number snapshots in a snapshot matrix
N	= total number of modes used in projection
n	= number of variables, equal to number of grid points times number of physical variables
\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
R	= covariance
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

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

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

[‡]Associate Professor, mcnamara.190@osu.edu, AIAA Associate Fellow

u_1	= velocity in the streamwise direction
u_2	= velocity in the cross-flow direction
$\tilde{\mathbf{u}}$	= prediction for the fluctuating component of the flow
\mathbf{x}	= position vector
x	= coordinate in the streamwise direction
y	= coordinate in the cross-flow direction
z	= random signal
β	= sparsity coefficient
\mathcal{D}	= rate of energy dissipation
\mathcal{E}	= turbulent kinetic energy
\mathcal{G}	= energy balance terms that are linear in fluctuating components
\mathcal{L}	= rate of convection of energy
Ω	= computational domain
\mathcal{P}	= rate of energy production
Φ	= matrix containing modes in its columns
Φ_i	= i^{th} mode
∇	= the gradient operator
\mathcal{T}	= rate of transfer of energy between different modes
ξ	= time averaged accumulated energy in ROM predictions
ζ	= non-dimensional time lag
$\ \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
$1, 2$	= indices for random signals

I. Introduction

The advancement of computing hardware and the advent of parallel computing algorithms has enabled unprecedented insight into complex flow physics using computational fluid dynamics (CFD) simulations. This insight is often crucial in developing a detailed understanding of multi-disciplinary interactions with complex flow environments. However, the computational cost associated with high fidelity analysis tools has prevented their integration into broad systems level and multi-disciplinary studies. This critical need has motivated the pursuit of tractable, accurate, and robust reduced order models (ROMs).

A commonly used ROM approach involves projecting the governing equations onto a reduced dimensional space comprising of characteristic bases.¹⁻⁶ The projections can be orthogonal (e.g., Galerkin) or non-orthogonal (e.g., Petrov-Galerkin).¹⁻⁶ The accuracy of such approaches is intimately bound to the quality of the chosen bases. The data from which appropriate bases for unsteady flow fields can be extracted is typically limited to small spatial and temporal domains due to the previously mentioned computational expense of high fidelity CFD simulation. Furthermore, such data is high-dimensional due to ultrafine spatial and temporal resolution. Thus, identifying a compact set of dynamically important flow features that fundamentally characterize the fluid dynamics is a non-trivial exercise. In this context, 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, and a reduced dimensional basis of the system is typically 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. 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 that use a truncated basis may over-predict kinetic energy due to insufficient energy dissipation.^{1,11,13} Moreover, the energy accumulation over a period of time may also cause the ROM to become unstable.^{11,13,15} These issues are a significant concern for fluid-structure interactions problems, since over-prediction of the kinetic energy in the fluid will cascade into the structural response. Additionally, 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 all the POD modes are active in a response, i.e., all have nonzero coefficients for any time window. For these reasons the optimality of POD is generally not a well-suited criteria for predictive reduced order modeling of nonlinear, unsteady flows.

Techniques such as Balanced Truncation,¹⁶ Balanced POD (BPOD),² and Eigensystem Realisation 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 an accurate and 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 continued need to explore alternative basis identification techniques.

Olshausen and Field [18] 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 — for a given time window. 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 the best compact representation to approximately span 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 Recently, a sparsity-promoting Dynamic Mode Decomposition (DMD) algorithm was used to carry out efficient and accurate reconstruction of fluid-flow data.²⁶

This study is an extension of work carried out by the authors,²⁷ in which Galerkin projection models based on sparse modes were developed for the first time and compared with POD based models. This previous work indicated a strong potential for sparse modes to yield improved predictive ROMs. The goal of the present study is to develop a deeper assessment and understanding of this approach. To this end, POD and sparse ROMs are compared to the full order system using metrics such as, root mean square of flow fields, turbulent kinetic energy (TKE), power spectral densities (PSDs) of the TKE, local flow dynamics based on auto-covariances, two-point covariances, etc. Additionally, an energy balance equation^{28,29} is used to assess the abilities of POD and sparse ROMs in preserving and maintaining the energy in the reduced system.

The remainder of this paper is organized as follows. The POD and sparse coding approaches are presented in Section II. The Galerkin projection framework, and formulation of energy balance equations are also discussed. Case study describing the application of POD and sparse modes to model the unsteady flow field is presented in Section III. Concluding remarks are given in Section IV.

II. Method of Solution

High resolution data are computed by solving the 2-D incompressible Navier Stokes (NS) equations using a Direct Numerical Simulation (DNS) CFD solver. The POD approach is based on the method of snapshots developed by Sirovich.¹⁰ Sparse modes are computed using the algorithm developed by Friedman et al. [30]. These approaches are detailed next.

II.A. Full Order Models

The DNS solution to the non-dimensionalized NS equations given by Eq. 1 is generated using the immersed boundary method.³¹

$$\nabla \cdot \mathbf{u} = 0, \quad \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \frac{1}{Re} \nabla^2 \mathbf{u} \quad (1)$$

where ∇ is the gradient operator, t is non-dimensional time, Re is Reynolds number, p is pressure, and \mathbf{u} is velocity.

The DNS approach 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.³¹

II.B. Generation of the Reduced Order modes

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

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

where Φ_i is the i^{th} mode, $\tilde{\mathbf{u}}$ is the fluctuating component of the velocity field, N is the number of modes, $s^i(t)$ is the coefficient corresponding to the i^{th} mode at time t , and $\mathbf{x} \in \Omega$, Ω being the flow domain. The computation of POD and sparse modes are discussed next.

II.B.1. Proper Orthogonal Decomposition

First, a snapshot matrix $\mathbf{Q} = [\mathbf{q}_1 \mathbf{q}_2 \cdots \mathbf{q}_k \cdots \mathbf{q}_m]$ is constructed by stacking the flow observations \mathbf{q}_k in its columns. The procedure to extract POD modes from a snapshot matrix has been extensively documented^{8,10} and therefore is not repeated here. However, for comparative discussion with the sparse coding procedure, note that the POD modes fundamentally represent the solution to the following minimization problem:³²

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

where the columns of Φ are the POD modes, \mathbf{S} is the coefficient matrix, and $\|\cdot\|_F^2$ is the square Frobenius norm of \cdot . 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:

$$\min_{S, \Phi} \frac{1}{2} \|\mathbf{Q} - \Phi \mathbf{S}\|_F^2 + \beta \|\mathbf{S}\|_0, \quad \|\Phi_i\| \leq 1 \quad \text{for all } i \quad (4)$$

where the columns of Φ are the sparse modes; \mathbf{S} is a matrix of activation coefficients; each column s_k of \mathbf{S} is encouraged to be sparse by a penalty on the L0 “norm” of \mathbf{S} (the count of the non-zero elements of the matrix); and $\beta > 0$ is the regularization (or penalty) parameter, denoted here as the *sparsity coefficient*. Note that if the penalty term is ignored, Eq. 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 \mathbf{S} . Yang et al. [21] recommend 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,³³ both due to the fact that the objective is not jointly convex in Φ and \mathbf{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.³³ Optimization problem is then given as:

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

where $\|\cdot\|_1$ is the L1 norm. Although this problem is still not jointly convex in \mathbf{S} and $\mathbf{\Phi}$, 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 \mathbf{S} is known as a least absolute shrinkage and selection operator (LASSO) problem,³⁴ and has been widely studied.^{30,34,35} A number of algorithms have been devised to solve Eq. 5. In the current study, an algorithm³⁰ based on the coordinate descent technique is adopted. Likewise, optimizing Eq. 5 over $\mathbf{\Phi}$ is a constrained least-squares problem, and is solved by using the Method of Optimal Directions (MOD) approach.³⁶ The MOD approach is equivalent to post-multiplying \mathbf{Q} with the pseudo-inverse of matrix \mathbf{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. Moreover, the sparse modes are not ‘ordered’ in terms of energy content. Therefore, computation of a completely new set of sparse modes is needed when building a new sparse ROM; unlike POD, a set of sparse modes from a lower order ROM cannot be retained when building a higher order ROM. The process of sparse coding is summarized in Algorithm 1.

The desired number of sparse modes are generated by varying β between 0 and 0.5. For any given value of β , the process of computing the sparse modes is initiated with a guess for the sparse modes. The modes and corresponding coefficient matrix are updated iteratively as explained earlier. The sparse coding process is stopped once the modes are converged, such that the absolute change in the value of the objective function Eq. 5 between two consecutive iterations is less than the convergence criterion $\epsilon = 10^{-3}$. Due to the non-convexity of the overall minimization problem, numerical experimentation using random initial guesses was found in a high percentage of cases to yield locally optimal solutions that produced poor results. In order to avoid this, a deterministic procedure was developed for the initial guess of the sparse modes. First, as many POD modes as the desired number of sparse modes are computed. Next, the initial guess for sparse modes is computed as an *ad hoc* linear combination of lower order and higher order POD modes:

$$\phi_i = 0.6\psi_i + 0.4\psi_{N-i}, \quad i = 1, 2, \dots, N \quad (6)$$

where ϕ_i is the initial guess for the i^{th} sparse mode, ψ_i is the i^{th} POD mode, and N is the number of sparse modes to be computed. Note that simply using the POD modes as the initial guess yields sparse bases that are highly biased towards the lower order POD modes. However, the linear combination of the POD modes Eq. 6 as an initial guess removes this bias by introducing both low and high order features.

Another factor that is found to affect the performance of a set of sparse modes in the context of Galerkin projection is the sparsity of the coefficient matrix, defined as:

$$\text{Sparsity} = 100 \times \frac{\text{number of zero elements in } \mathbf{S}}{m \times N} \% \quad (7)$$

A coefficient matrix with a low sparsity is associated with modes that do not capture the local dynamics (transience) of the full order system. Moreover, a highly sparse coefficient matrix may result in modes that provide a poor representation of structure in the observed data. Thus, an ‘optimal’ sparsity level exists to capture both the structure and local dynamics in the observed data. Sparsity levels of 50 to 80 % are found to result in stable and reasonably accurate sparse ROMs. The sparsity in the coefficient matrix is ensured to fall in this range by varying the β value before building a ROM. Note that the selection of a β value that satisfies the sparsity constraints is currently a trial and error process.

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.¹⁻⁴ In Galerkin projection, 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:

Algorithm 1: Sparse coding algorithm

Input : Snapshot matrix \mathbf{Q} of size $n \times m$, number of sparse modes to be compute (N), sparsity coefficient (β), convergence criterion ($\epsilon \geq 10^{-3}$)

Output: Sparse modes (Φ), sparse coefficient matrix (\mathbf{S})

- 1 Compute first N POD modes ($\psi_1, \psi_2, \dots, \psi_N$)
- 2 Initialize sparse modes as:
- 3 **for** $i = 1, 2, \dots, N$ **do**
- 4 $\phi_i = 0.6\psi_i + 0.4\psi_{N-i}$
- 5 **end**
- 6 Iteration number, $i = 1$
- 7 **while** $\epsilon \geq 10^{-3}$ **do**
- 8 Solve for sparse coefficient matrix \mathbf{S} one column \mathbf{s}_k at a time by solving the LASSO problem:
- 9 **for** $k = 1, 2, \dots, m$ **do**
- 10 $\min_{\mathbf{s}_k} \left(\frac{1}{2} \|\mathbf{q}_k - \Phi \mathbf{s}_k\|_F^2 + \beta \|\mathbf{s}_k\|_1 \right)$
- 11 **end**
- 12 Update the sparse modes Φ by solving the following minimization problem using the MOD approach:
$$\min_{\Phi} \sum_{k=1}^m \frac{1}{2} \left(\|\mathbf{q}_k - \Phi \mathbf{s}_k\|_F^2 \right), \quad \|\Phi_i\| \leq 1 \text{ for all } i$$
- 13 Evaluate the objective function:
$$\Gamma(i) = \sum_{k=1}^m \left(\frac{1}{2} \|\mathbf{q}_k - \Phi \mathbf{s}_k\|_F^2 + \beta \|\mathbf{s}_k\|_1 \right), \quad \|\Phi_i\| \leq 1 \text{ for all } i$$
- 14 Check for convergence, $\epsilon = |\Gamma(i) - \Gamma(i-1)|$
- 15 Update iteration number $i = i + 1$
- 16 **end**
- 17 Evaluate sparsity, $\text{Sparsity} = 100 \times \frac{\text{number of zero elements in } \mathbf{S}}{m \times N} \%$

1. Fluctuating component of each snapshot (\mathbf{q}_k) is expanded as a linear combination of reduced order modes as shown in Eq. 2.
2. The basis expansion is substituted into the governing equations, Eq. 1.
3. The residual term is minimized by constraining it to be orthogonal to the space spanned by the modes:

$$\left(\Phi_i, \frac{\partial(\tilde{\mathbf{u}} + \bar{\mathbf{U}})}{\partial t} + ((\tilde{\mathbf{u}} + \bar{\mathbf{U}}) \cdot \nabla) (\tilde{\mathbf{u}} + \bar{\mathbf{U}}) \right) = \frac{1}{Re} (\Phi_i, \nabla^2(\tilde{\mathbf{u}} + \bar{\mathbf{U}})) \quad (8)$$

where $i = 1, 2, \dots, N$, and $\tilde{\mathbf{u}}$ is expressed as a linear combination of the reduced order modes, as shown in (2). The dot product between any vectors \mathbf{f} and \mathbf{g} is given as:

$$(\mathbf{f}, \mathbf{g}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{g} d\Omega \quad (9)$$

where Ω is the flow domain. Equation 8 yields the following set of evolution equations for the modal amplitudes $s^i(t)$:

$$\begin{bmatrix} \frac{ds^1(t)}{dt} \\ \vdots \\ \frac{ds^i(t)}{dt} \\ \vdots \\ \frac{ds^N(t)}{dt} \end{bmatrix} = [d_{ij}]^{-1} \begin{bmatrix} a_1 + \sum_{j=1}^N b_{1j} s^j(t) + \sum_{j=1}^N \sum_{k=1}^N c_{1jk} s^j(t) s^k(t) \\ \vdots \\ a_i + \sum_{j=1}^N b_{ij} s^j(t) + \sum_{j=1}^N \sum_{k=1}^N c_{ijk} s^j(t) s^k(t) \\ \vdots \\ a_N + \sum_{j=1}^N b_{Nj} s^j(t) + \sum_{j=1}^N \sum_{k=1}^N c_{Njk} s^j(t) s^k(t) \end{bmatrix} \quad (10)$$

where the coefficients are given as:

$$a_i = -(\Phi_i, (\bar{U} \cdot \nabla) \bar{U}) + \frac{1}{Re} (\Phi_i, \nabla^2 \bar{U}) \quad (11)$$

$$b_{ij} = -(\Phi_i, (\Phi_j \cdot \nabla) \bar{U}) - (\Phi_i, (\bar{U} \cdot \nabla) \Phi_j) + \frac{1}{Re} (\Phi_i, \nabla^2 \Phi_j) \quad (12)$$

$$c_{ijk} = -(\Phi_i, (\Phi_j \cdot \nabla) \Phi_k) \quad (13)$$

$$d_{ij} = (\Phi_i, \Phi_j) \quad (14)$$

The system of ordinary differential equations is then time marched using the fourth order Runge-Kutta scheme to obtain the prediction coefficients. Moreover, the projection of pressure term is written as:

$$(\Phi_i, \nabla p) = \int_{\Omega} \Phi_i \cdot \nabla p d\Omega. \quad (15)$$

The reduced order modes (Φ_i) are divergence free for incompressible flow;¹ therefore:

$$(\Phi_i, \nabla p) = \int_{\Omega} \nabla \cdot (p \Phi_i) d\Omega. \quad (16)$$

Using divergence theorem on Eq. 16:

$$(\Phi_i, \nabla p) = \int_A (p \Phi_i) \cdot \hat{n} dA. \quad (17)$$

As the reduced order modes (Φ_i) satisfy the homogeneous boundary conditions,¹ the integral of Φ_i over the boundary of the computational domain is 0. Thus, the projection of the pressure term is zero.

II.B.4. Energy Balance Equations

As a metric of comparison for the full and reduced order models, the following energy balance equation^{28,29} is used:

$$\frac{\partial \mathcal{E}}{\partial t} = \mathcal{L} + \mathcal{P} + \mathcal{T} + \mathcal{D} + \mathcal{G} \quad (18)$$

where \mathcal{E} is the turbulent kinetic energy, and $\mathcal{L}, \mathcal{P}, \mathcal{T}, \mathcal{D}, \mathcal{G}$ are time rates of change of convection, production, transfer, dissipation, and terms that are linear in fluctuating components, respectively. These terms are given as:

$$\frac{\partial \mathcal{E}}{\partial t} = \frac{1}{2} \frac{\partial (\tilde{u}, \tilde{u})}{\partial t} \quad (19)$$

$$\mathcal{L} = -(\tilde{u}, (\bar{U} \cdot \nabla) \tilde{u}) \quad (20)$$

$$\mathcal{P} = -(\tilde{u}, (\tilde{u} \cdot \nabla) \bar{U}) \quad (21)$$

$$\mathcal{T} = -(\tilde{u}, (\tilde{u} \cdot \nabla) \tilde{u}) \quad (22)$$

$$\mathcal{D} = \frac{1}{Re} (\tilde{u}, \nabla^2 \tilde{u}) \quad (23)$$

$$\mathcal{G} = \frac{1}{Re} (\tilde{u}, \nabla^2 \bar{U}) - (\tilde{u}, (\bar{U} \cdot \nabla) \bar{U}). \quad (24)$$

The time rate of change of TKE (\mathcal{E}) accounts for the changes in TKE at any instant in time, \mathcal{L} accounts for variations in TKE due to convection, \mathcal{P} is the rate at which energy is being added to the system from the

mean flow, \mathcal{D} is the rate of dissipation of energy due to the small scale turbulent structures, and \mathcal{T} is the rate of transfer of energy between different turbulent structures.^{28,29} For analysis of the energy, Eq. 18 is integrated in time:

$$\int_{t=0}^{t_0} \frac{\partial \mathcal{E}}{\partial t} dt = \int_{t=0}^{t_0} (\mathcal{L} + \mathcal{P} + \mathcal{T} + \mathcal{D} + \mathcal{G}) dt \quad (25)$$

where t_0 is the length of time window over which the data are collected. Note that for a statistically stationary process, the net rate of change of TKE over time is:

$$\int_{t=0}^{t_0} \frac{\partial \mathcal{E}}{\partial t} dt \approx 0. \quad (26)$$

Moreover, because \mathcal{G} is linear in fluctuating component, and time integral of fluctuating component is zero,

$$\int_{t=0}^{t_0} \mathcal{G} dt = 0. \quad (27)$$

Thus, for a statistically stationary process, the time integral of energy balance equation Eq. 18 reduces to:

$$\int_{t=0}^{t_0} (\mathcal{L} + \mathcal{P} + \mathcal{T} + \mathcal{D}) dt \approx 0. \quad (28)$$

Finally, the energy balance equation for a ROM is obtained by substituting the basis expansion, Eq. 2, into Eq. 18. As an accurate ROM for a statistically stationary system will not add or deduct TKE, thus satisfying Eq. 28.

III. Modeling of a Lid-driven Cavity

Application of POD and sparse approaches to model 2-D incompressible unsteady flow inside a lid-driven cavity is considered. The unsteady response of a fluid enclosed in a cavity is actuated using a moving surface (lid). This configuration, widely referred to as ‘the lid-driven cavity flow’ has been used in several studies^{11,37,38} 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 50,000 uniformly sampled observations are collected from 500 units of non-dimensional time once the flow had reached a statistically stationary state. The first step in carrying out reduced order modeling is construction of a snapshot matrix, this process is discussed next.

III.A. Choice of Snapshot Matrix

A subset of snapshots from the large set of observations is used to construct a snapshot matrix. An ‘optimal’ snapshot matrix should capture most of the important flow features while keeping the number of snapshots to minimum. Convergences in the mean flow and instantaneous turbulent kinetic energy (TKE) quantities are used as criteria for choosing an optimal snapshot matrix. Four snapshot matrices constructed using different sampling frequencies and spanning different periods of time are considered, as summarised in Table 1. The snapshot matrix numbered 1 with sampling frequency of 100 and time span of 50 units is used as the benchmark case, against which the rest of the three snapshot matrices are compared. Normalized L2 percent error in the estimated mean flow field for matrices 2, 3, and 4 are also provided in Table 1. The errors in mean flow for matrices 2 and 3 are less than half of that for matrix 4. Next, different sets of fluctuating velocity components are computed using the estimated mean flow fields. Time-histories of TKEs are computed over 100 units of time using the sets of fluctuating components, as shown in Fig. 1. The normalized L2 errors in estimates of TKE obtained from matrix numbered 2 and 3 are only 3%, whereas for matrix 4 the error is as high as 13.32 %. This convergence study indicates that the first and second order statistics, i.e., mean flow and instantaneous TKE, respectively, are converged for matrices 2 and 3. Moreover, matrix 3 contains 1250 snapshots as compared to 2500 snapshots in matrix 2. Therefore, snapshot matrix numbered 3 with sampling rate of 50 units and time span of 25 units is used to construct ROMs.

Table 1. Normalized L2 errors in estimated mean flows obtained from four different snapshot matrices, and estimated instantaneous TKE over 100 time units.

Snapshot matrix number	Non-dimensional time spanned	Non-dimensional sampling frequency	Number of snapshots	Error in mean flow (%)	Error in TKE (%)
1	50	100	5000	0.00	0.00
2	25	100	2500	1.16	3.00
3	25	50	1250	1.16	3.00
4	12.5	100	1250	2.77	13.32

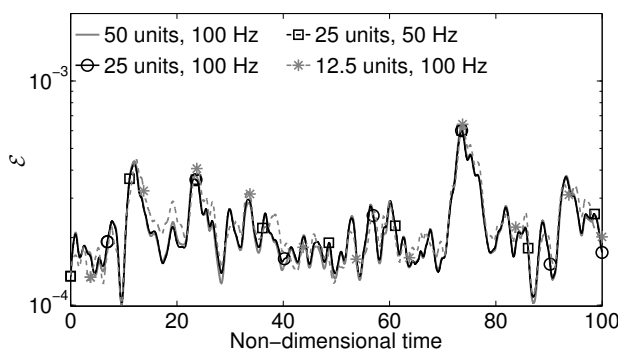


Figure 1. The TKE of flow predicted by DNS over first 100 time units, computed using mean flow obtained from different snapshot matrices.

III.B. Comparison Between POD and Sparse Modes

Comparison between POD and sparse modes, and original snapshots is provided next. The velocity components of the POD modes numbered 1, 10, 50 and 100, and four modes taken arbitrarily from a set of 100 sparse modes are shown in Figs. 2 and 3, respectively. For reference, the velocity components from the full order solution are displayed in Fig. 4. 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, where lower order modes capture the high energy structures. Thus, a truncated set of POD modes discards the low-energy structures. In contrast, sparse modes resemble the snapshots of the original flow field and yield qualitatively similar multi-scale features.

The energy content in the modes is estimated as the sum of the squares of the projection coefficients corresponding to the snapshot matrix. The energies of the modes are shown in Fig. 5. The energy contained in the POD modes decreases rapidly with increase in mode number, with first 17 modes accounting for approximately 90% of the total energy. In comparison, the energy distribution in the sparse modes is more uniform, or leveled.

The time varying projection coefficients of the POD and sparse modes for first 25 units of time history are shown in Fig. 6. The POD coefficients are almost always non-zero, i.e., active. 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 a 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.

The ability of POD and sparse approaches to carry out effective dimensionality reduction is assessed next. For this purpose, the snapshot matrix is approximated using different sets of POD and sparse modes. The root mean square (RMS) values of the approximated flow velocities u_1 and u_2 are computed over time, at each of the grid points, for different sets of POD and sparse modes. The normalized L2 errors in representation of the RMS flow fields by different sets of modes are listed in Table 2. The errors corresponding to

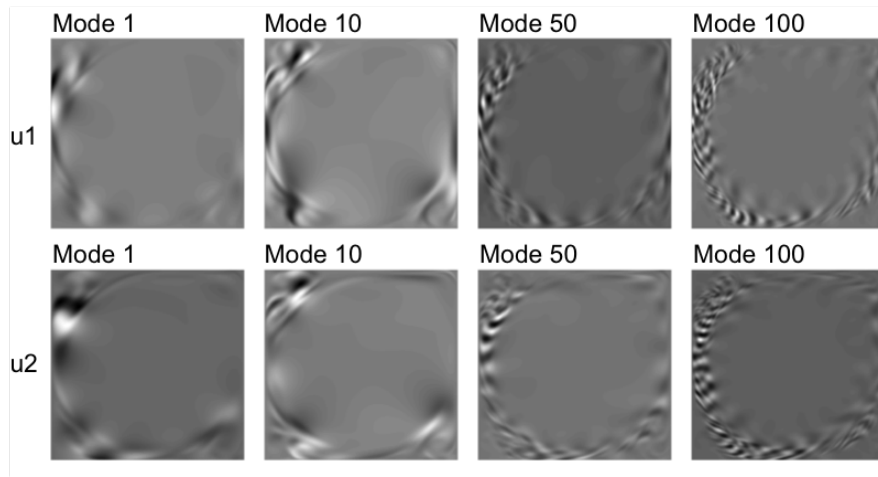


Figure 2. The u_1 and u_2 components of the POD modes numbered 1, 10, 50 and 100 of the lid-driven cavity.

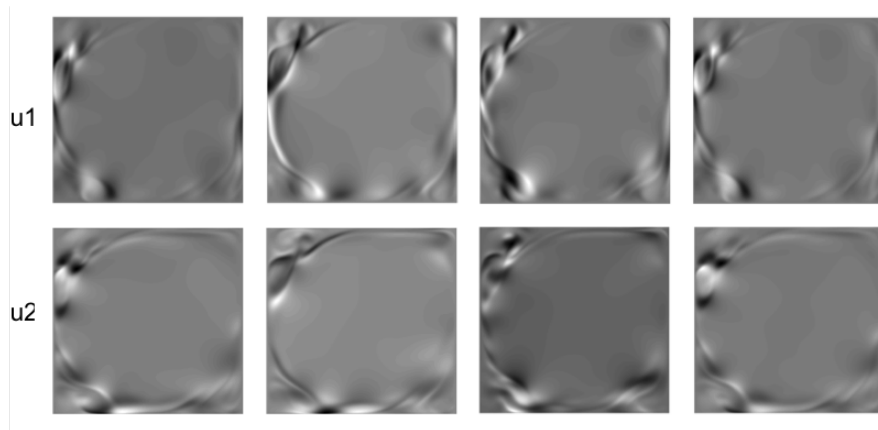


Figure 3. The u_1 and u_2 components of four of the sparse modes of the lid-driven cavity.

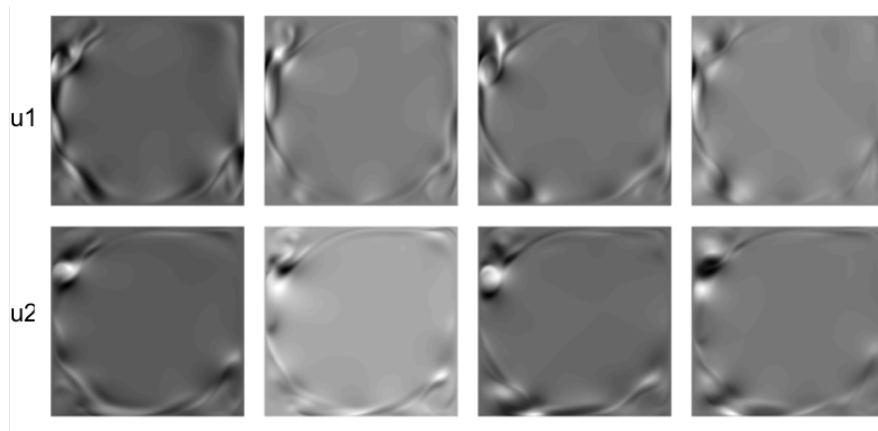


Figure 4. The u_1 and u_2 components of four different snapshots of the lid-driven cavity.

the POD sets are smaller than those of sparse sets, for same number of modes. This is due to the optimality of the POD modes in reconstructing the snap-shot matrix. However, the differences between the errors in representation using POD and sparse sets are small, with less than 1.5% of maximum difference.

Next, the property of sparse modes to capture multi-scale features is further studied. The transformation matrix between 20 sparse modes and 200 POD modes is computed. Next, \log_{10} of the absolute value of the

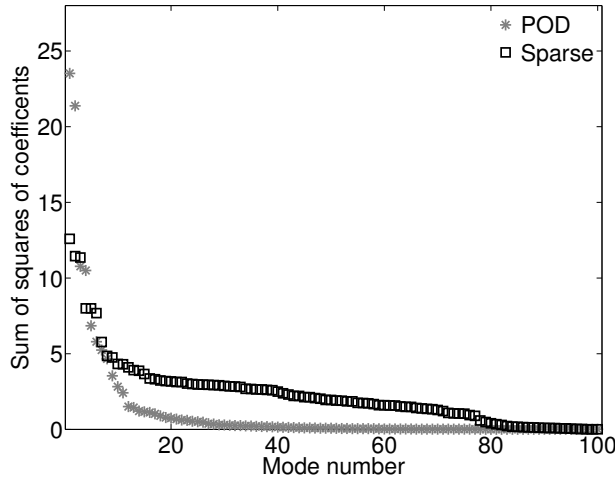


Figure 5. Estimated energy content in the first 100 POD modes and the sparse modes of a 100-mode model of the lid-driven cavity, expressed in terms of the sum of squares of projection coefficients evaluated over 25 time units. The energy content in sparse modes is rearranged in terms of decreasing sums of squares of coefficients.

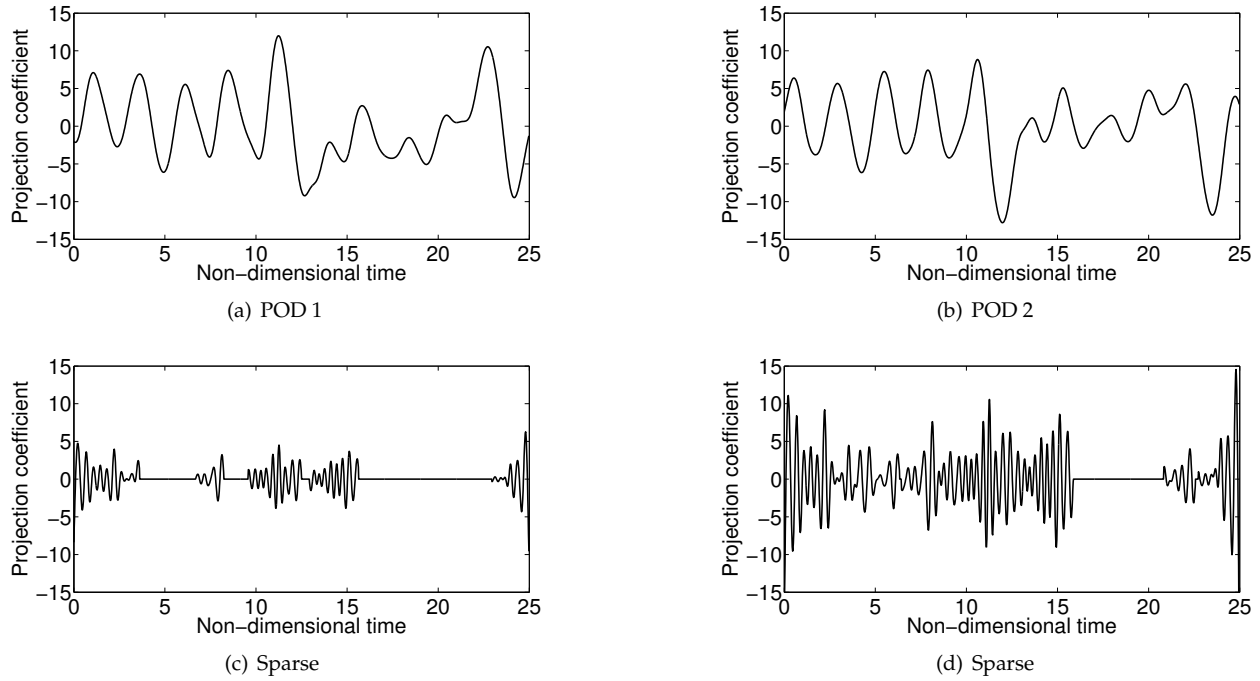


Figure 6. The time history of projection coefficients corresponding to the 100-mode POD and 100-mode sparse models of the lid-driven cavity.

transformation matrix is computed to obtain the ‘modified’ transformation matrix, as shown in Fig. 7. A value of 0 in the modified transformation matrix represents 100% correlation between corresponding pair of POD and sparse modes, a value of $-\infty$ represents zero correlation. Note that none of the entries in the modified transformation matrix is smaller than -9. Thus, the 20 sparse modes contain features of all the 200 POD modes. Consequentially, a set of sparse modes contain both the large energy (low order POD modes) and small energy (high order POD modes) information.

III.C. Predicted RMS Velocity Fields

The flow predictions obtained from the ROMs are compared to the full order solutions for the 500 time units. The turbulent flows are chaotic in nature, therefore point-wise accurate solutions are not reasonably

Table 2. Normalized L2 error in representation of RMS of u_1 and u_2 velocity components for the snapshot matrix by POD and sparse modes.

Number of modes	POD, u_1 (%)	POD, u_2 (%)	Sparse, u_1 (%)	Sparse, u_2 (%)
10	1.09×10^1	1.19×10^1	1.23×10^1	1.25×10^1
20	4.89	4.79	5.95	5.71
50	5.86×10^{-1}	5.66×10^{-1}	7.21×10^{-1}	8.37×10^{-1}
80	9.76×10^{-2}	9.18×10^{-2}	1.03×10^{-1}	1.08×10^{-1}
100	3.42×10^{-2}	3.13×10^{-2}	3.82×10^{-2}	3.83×10^{-2}
200	4.03×10^{-4}	4.04×10^{-4}	4.06×10^{-4}	4.06×10^{-4}

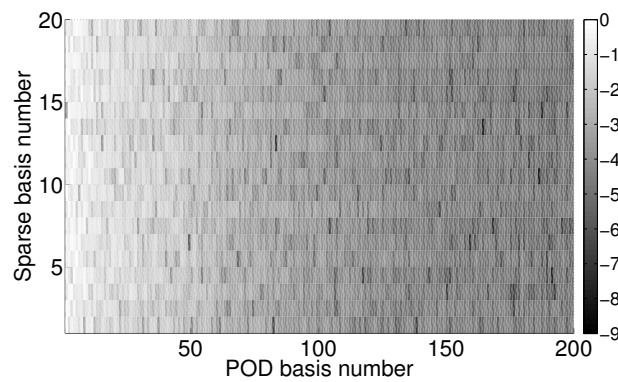


Figure 7. Transformation matrix between the 20 sparse modes and 200 POD modes.

expected from a ROM; instead, accurate prediction of second order statistics is sought. The RMS of flow field is one of the metrics used in this study to assess ROM accuracy. The normalized L2 errors in the RMS of u_1 and u_2 components computed over time, as predicted by POD and sparse ROMs, are listed in Table 3. The errors in predicting the RMS velocities are smaller for sparse ROMs as compared to POD ROMs when less than 80 modes are used. The errors converge to approximately 10 % when 80 or more modes are used in ROM construction. Moreover, the convergence in error values is significantly faster for sparse ROMs compared to POD ROMs.

Table 3. Normalized L2 error in prediction of RMS of u_1 and u_2 velocity components by several POD and sparse ROMs.

Number of modes	POD, u_1 (%)	POD, u_2 (%)	Sparse, u_1 (%)	Sparse, u_2 (%)
10	324.91	302.56	33.07	29.37
20	832.48	838.71	17.26	18.35
50	51.26	53.00	10.14	10.08
80	11.20	11.68	11.73	11.21
100	11.21	11.53	10.12	11.17
200	10.97	11.30	11.42	11.40

III.D. Predicted Instantaneous TKE

The instantaneous TKE predicted by the ROMs is shown in Fig. 8. The sparse ROMs generally capture the energy levels with greater accuracy compared to the POD ROMs for the same number of modes used. The 10- and 20-mode POD models over-predict the energy by two orders of magnitude, averaged over first 500 time units, whereas, the 10-mode sparse model over-predicts the energy by a factor of 1.2 over the same duration. Similar to the observations made for the RMS velocities, as many as 80 POD modes are required to capture the TKE levels; whereas, all the sparse models reasonably capture the TKE levels. These results are consistent with the previous findings^{1,11,13,29} 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.

Next, the power spectral densities (PSDs) of the instantaneous turbulent kinetic energy predicted by 20-, 80-, and 200-mode POD and sparse ROMs are shown in Fig. 9. Bendat and Piersol [39] recommend using long time-histories with integral powers of 2 instances for accurate PSD estimations. Therefore, the PSDs are computed over 2^{15} data points corresponding to the last 327.68 time units using Welch's method.⁴⁰ Moreover, Hanning window is used to minimize truncation errors. As demonstrated in the previous set of results, the 20-mode POD model significantly over-predicts the energy values over the entire range of frequencies. Whereas, 20-mode sparse and higher order ROMs provide a good overall prediction.

III.E. Predicted Local Flow Dynamics

Next, the accuracy of POD and sparse ROMs in capturing the local dynamics of the flow within the lid-driven cavity is studied. For this purpose the velocity data is considered at two probe locations denoted by P1 and P2, as shown in Fig. 10. Again, the probe data are saved for last 327.68 time units, which correspond to 2^{15} data points.

First, the PSDs of the $u1$ and $u2$ velocity components are computed for probe P1, Fig. 11. The 20-mode POD model over-predicts the PSD of both $u1$ and $u2$ components over the first 20 units of frequency. Whereas, 20-mode sparse, 80-, and 200-mode POD and sparse models provide accurate predictions for the PSDs. Notice that the PSDs are plotted in the logarithmic scale, and over 99% of the power is contained in the first 3 units of frequency.

Next, the autocovariances of predicted velocities are computed for probe P1. The autocovariance of a random signal $z1(t)$ is given as:³⁹

$$R_{z1,z1}(\zeta) = \langle (z1(t) - \mu_{z1}), (z1(t + \zeta) - \mu_{z1}) \rangle_t \quad (29)$$

where $R_{z1,z1}$ is autocovariance of a random signal $z1(t)$, μ_{z1} is mean of $z1(t)$, ζ is the non-dimensional time lag, and $\langle \cdot \rangle_t$ is convolution operation over variable t . The autocovariance is maximum at time lag $\zeta = 0$, i.e. when the signal is convoluted with itself without any time lag. Moreover, autocovariance of a random signal will sharply drop to zero as the $|\zeta|$ is increased from zero. Whereas, for a constant signal, the autocovariance will acquire a constant value for all the values of ζ . Thus, autocovariance provides a good measure of dynamics at any given point in the flow field. Autocovariances of $u1$ and $u2$ velocities at the P1 probe location are provided in the Fig. 12. The 20-mode POD model over-predicts the peak value of autocovariance for $u1$ by 2 orders of magnitude, and autocovariance for $u2$ by an order of magnitude. Whereas, the 20-mode sparse, and 80- and 200-mode POD and sparse models capture the peak value, width of the peak, and overall magnitude of autocovariance with reasonable accuracy. Furthermore, 20-mode sparse, and 80- and 200-mode POD and sparse ROMs predict the autocovariances of $u1$ component with better accuracy than autocovariances of $u2$ component. However, notice that the magnitudes of autocovariances of $u2$ component are 4 orders of magnitude smaller than those of $u1$ component. Therefore, these ROMs are able to capture the overall dynamics of the flow, which corresponds to the $u1$ component for the considered probes, with good accuracy.

In order to further assess the capabilities of the ROMs in capturing the local dynamics, the relationships between the flow quantities at two probe locations are studied. Two-point covariance between two random signals $z1$ and $z2$, recorded at two different probe locations, is given by:³⁹

$$R_{z1,z2}(\zeta) = \langle (z1(t) - \mu_{z1}), (z2(t + \zeta) - \mu_{z2}) \rangle_t \quad (30)$$

where $R_{z1,z2}$ is the two-point covariance of random signals $z1(t)$ and $z2(t)$, and μ_{z2} is mean of $z2(t)$. Two-point covariance provides an estimate of relationship between velocity signals at two probe locations. The

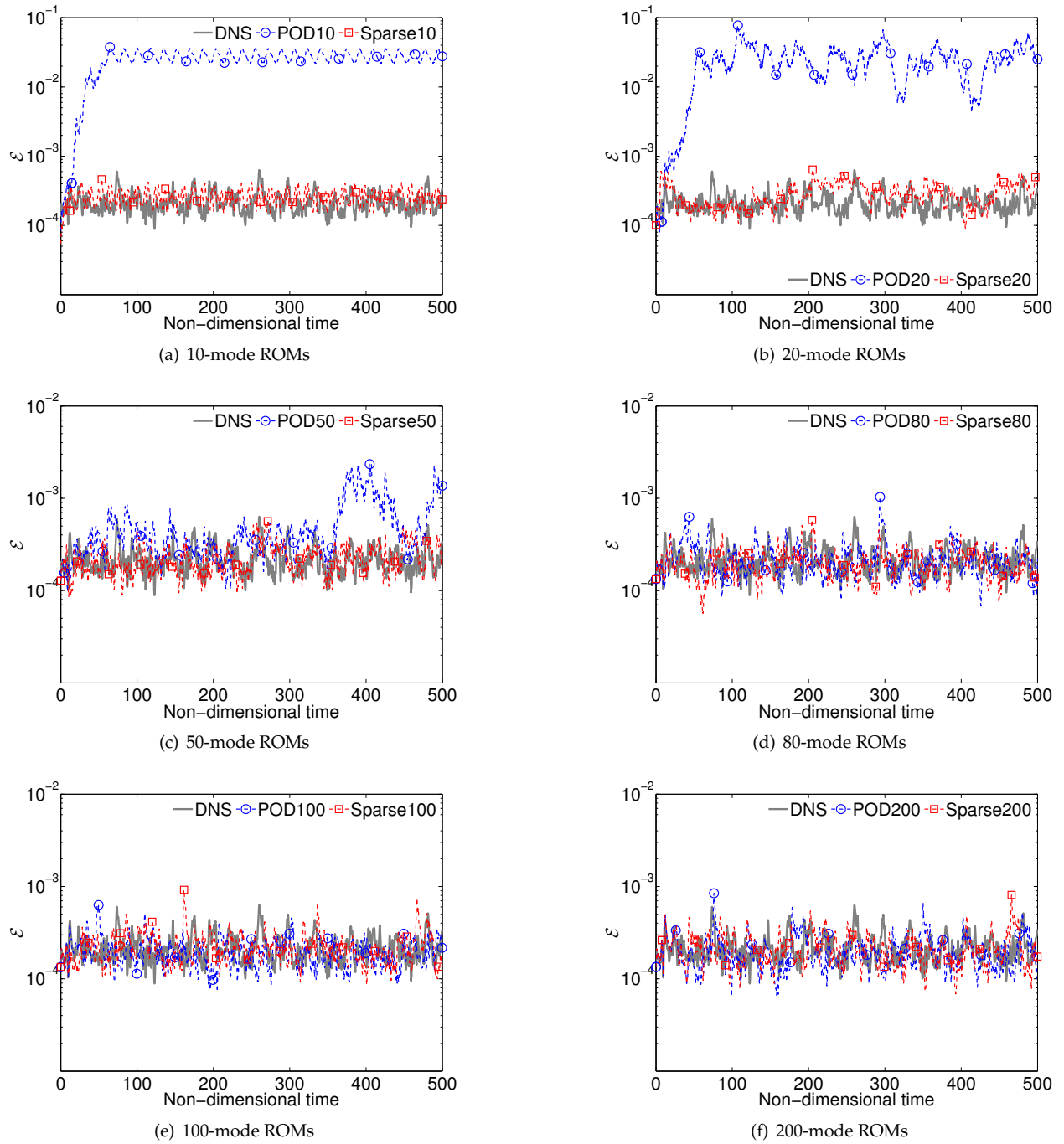


Figure 8. Time history of the instantaneous TKE of the lid-driven cavity as predicted by DNS, and the computed POD and sparse ROMs. The snapshot matrix used to compute the modes spans first 25 time units, and ROMs are integrated for 500 time units.

two-point covariances between the flow quantities at two probe locations, $R_{u1(p1),u1(p2)}$ and $R_{u2(p1),u2(p2)}$ are shown in Fig. 13. All the ROMs capture the location of the peak at $\zeta = 0.03$ time units. However, the 20-mode POD ROM significantly over-predicts the magnitudes of the two-point covariances. Similar to the previous observations, the 20-mode sparse, and the higher order ROMs provide reasonably accurate predictions for the two-point covariances.

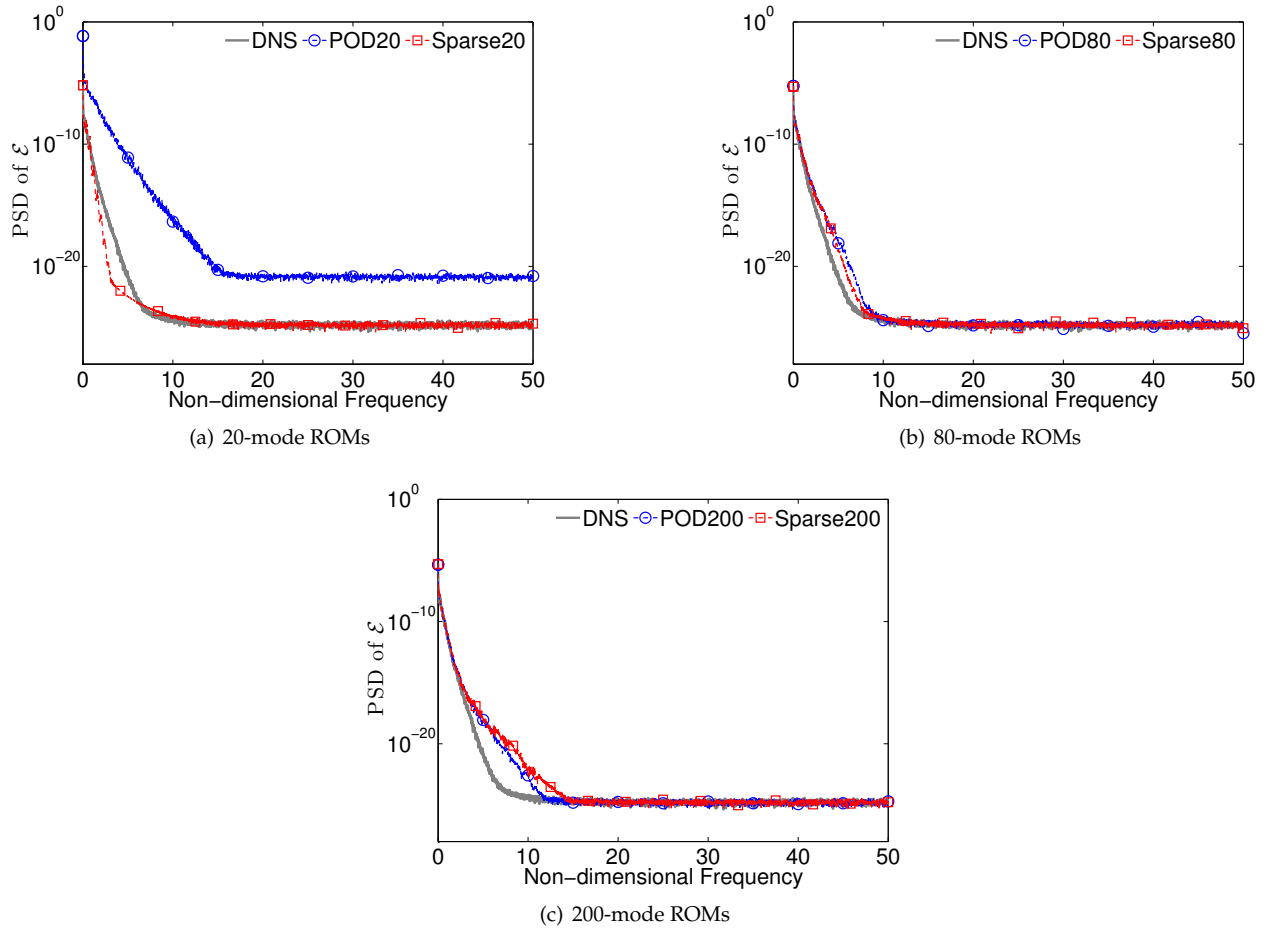


Figure 9. PSD of the turbulent kinetic energy of the lid-driven cavity as predicted by DNS, and the 20-, 80-, and 200-mode ROMs.

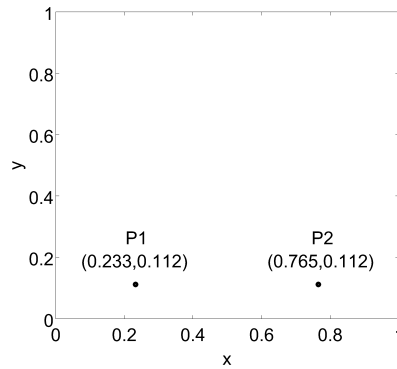
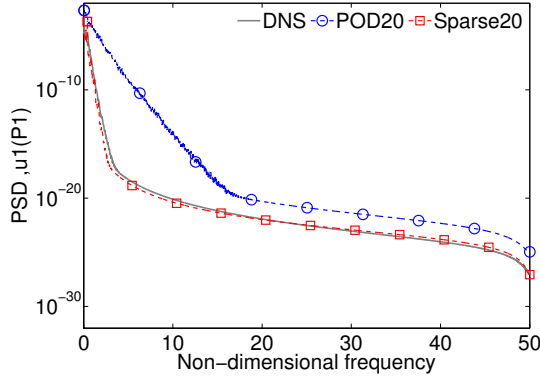


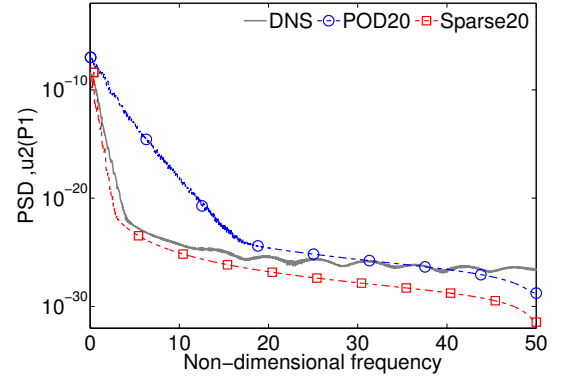
Figure 10. The two probe locations in the computational domain over which the local dynamics of the predicted and DNS flows are computed.

III.F. Energy Balance Analysis

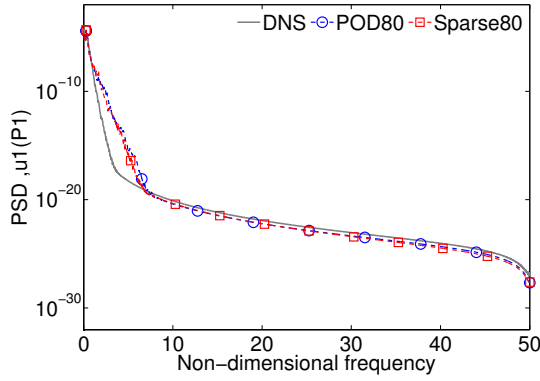
In order to further analyze the stabilizing property of sparse modes, energy balance analyses of DNS, and POD and sparse ROMs are carried out. The energy balance terms are computed using Eqs. 19-24. The energy balance terms for DNS are shown in Fig. 14. The TKE (\mathcal{E}), time rate of change of TKE ($\partial\mathcal{E}/\partial t$), production (\mathcal{P}), and dissipation (\mathcal{D}) are shown in Figs. 14(a) and 14(b). It is observed that there is a positive correlation between (\mathcal{P}) and ($\partial\mathcal{E}/\partial t$), and a negative correlation between (\mathcal{D}) and ($\partial\mathcal{E}/\partial t$). This result is due to the fact that an increase in rate of production will result in increase in rate of TKE; whereas, increase in



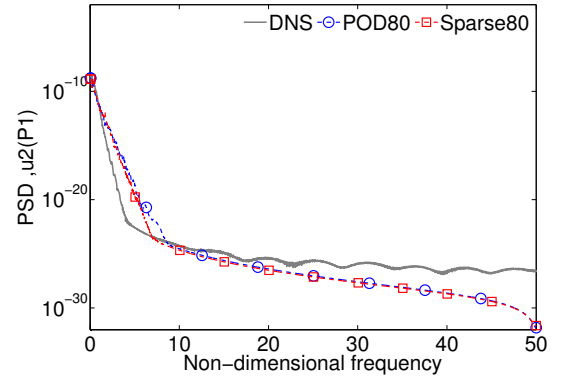
(a) PSD of u_1 by 20-mode models



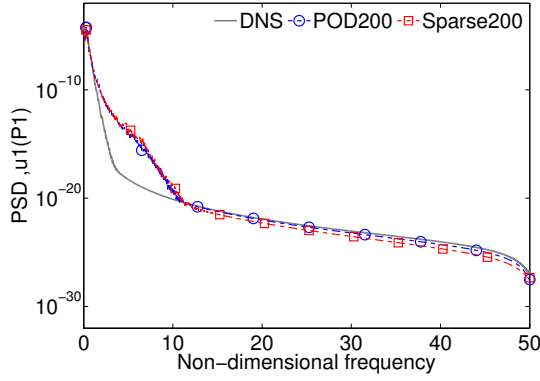
(b) PSD of u_2 by 20-mode models



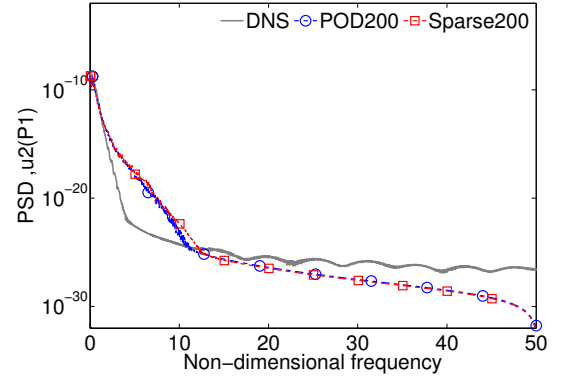
(c) PSD of u_1 by 80-mode models



(d) PSD of u_2 by 80-mode models



(e) PSD of u_1 by 200-mode models



(f) PSD of u_2 by 200-mode models

Figure 11. PSDs of the u_1 and u_2 velocity components at probe P1, as predicted by DNS, and 20-, 80-, and 200-mode ROMs.

magnitude of rate of dissipation will have opposite effect on rate of TKE. Moreover, positive values of TKE rate result in increase in TKE, and vice versa. Also notice that there is no net gain in TKE over time for the statistically stationary DNS data. Furthermore, terms \mathcal{L} , \mathcal{T} , and \mathcal{G} , shown in Fig. 14(c), are negligible compared to rest of the energy balance terms. The rate of convection (\mathcal{L}) is negligible as there is no net flow in or out of the considered control volume of a closed lid-driven cavity. The negligible transfer rate term (\mathcal{T}) is consistent with the study carried out by Noack et al.[29]. The \mathcal{G} term is small as expected for a statistically stationary flow.

Next, the energy balance analysis is carried out for POD and sparse ROM predictions. The instantaneous energy balance terms for the 10-, 20-, 80-, and 200-mode POD and sparse ROMs are shown in Figs. 15 and

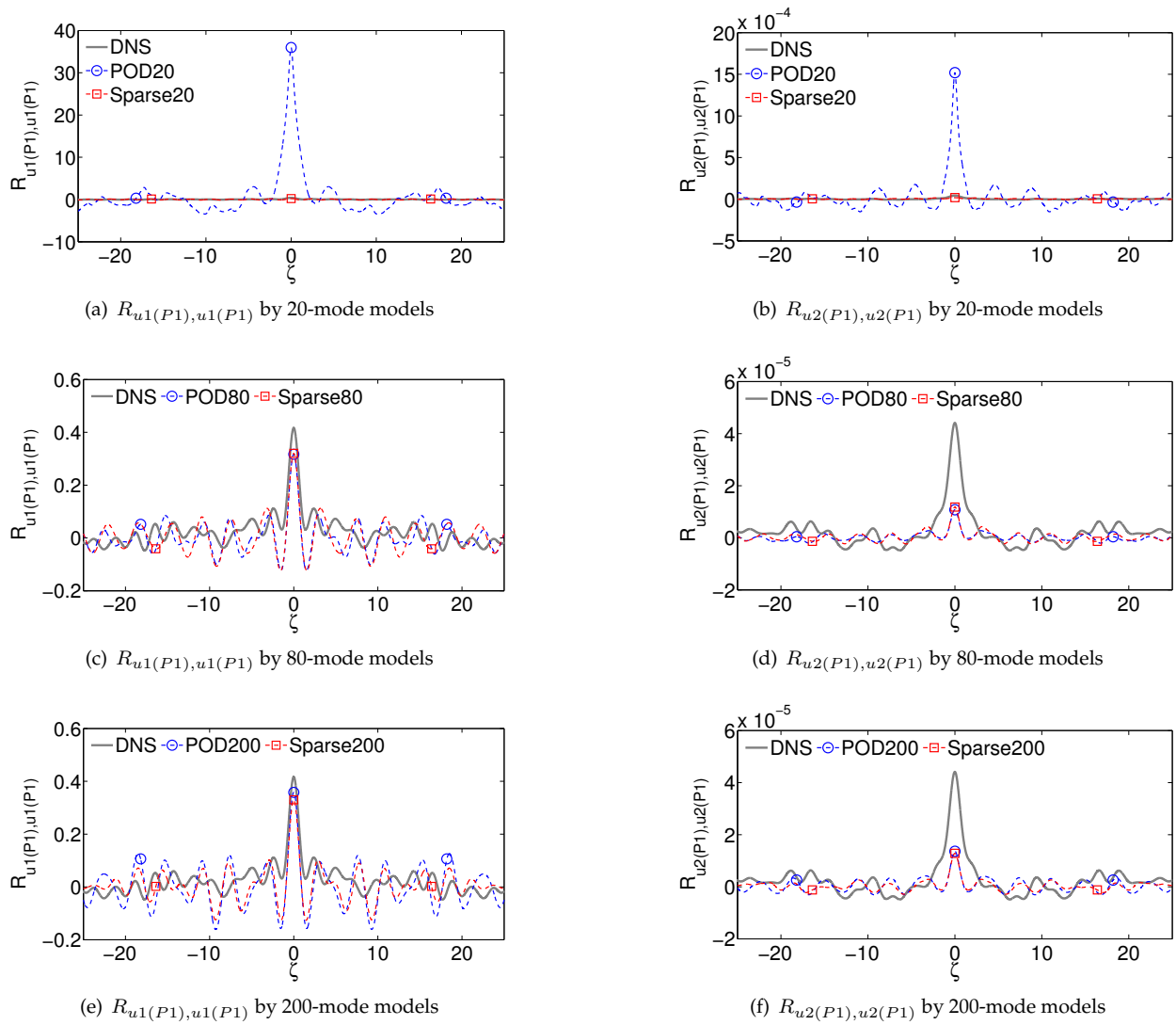


Figure 12. Autocovariances of the $u1$ and $u2$ velocity components at probe P1, as predicted by DNS, and 20-, 80-, and 200-mode ROMs.

16, respectively. Moreover, the average energy accumulation over different time spans starting at $t = 0$ units is computed for different ROMs as:

$$\xi = \frac{1}{t} \int_0^t (\mathcal{P} + \mathcal{D}) dt. \quad (31)$$

The average energy accumulation for different POD and sparse ROMs over first 50 time units are compared in Fig. 17. It is observed that the average energy accumulation increases for the 10- and 20-mode POD ROMs over the first 50 time unit; as a result the TKEs also increase over the same period (Figs. 15(a) and (b)) and saturate at much higher levels compared to DNS. The higher order POD and all the sparse ROMs maintain the energy balance by predicting the average energy accumulation in DNS solution with reasonable accuracy.

Next, the time-integrated energy balance terms (Eq. 28) computed from the flow approximations obtained from different sets of POD and sparse modes are analyzed. Here, the flow approximations are computed using direct projection of DNS flow solutions onto different sets of modes. The magnitudes of time-integrated energy balance terms for POD and sparse sets are provided in Figs. 18 and 19, respectively. The production terms for 10-, 20-, and 50-mode POD sets are significantly larger than the corresponding dissipation terms. On the other hand, the magnitudes of production and dissipation terms for the sets of

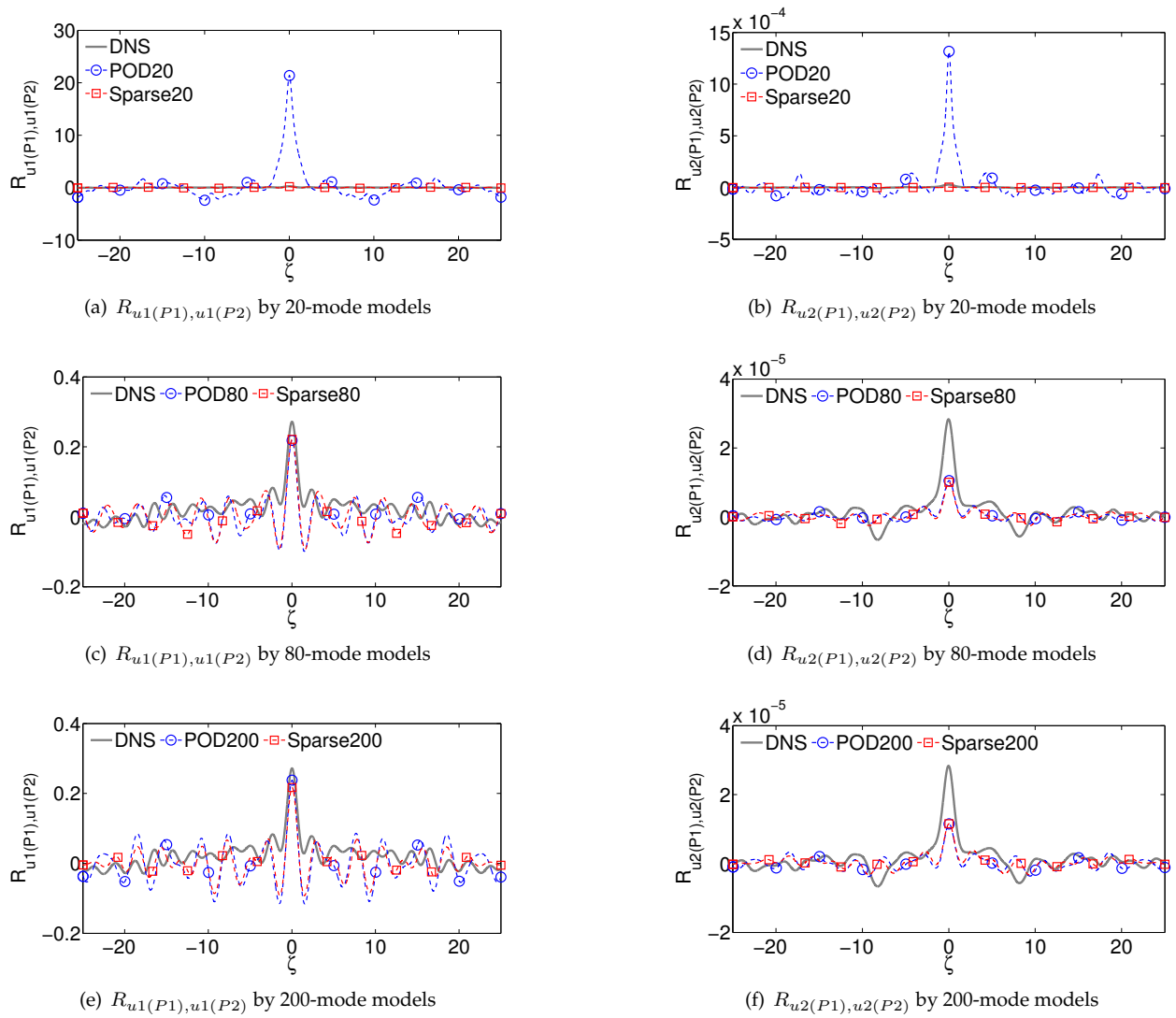


Figure 13. Two-point covariances of the $u1$ and $u2$ velocity components between probes P1 and P2, as predicted by DNS and 20-, 80-, and 200-mode ROMs.

sparse modes are approximately equal. These observations can be explained by considering the degree of spatial variance in POD and sparse modes. Recall that the lower order, high energy POD modes contain small spatial variance, and the spatial variance increases with mode number, as shown in Fig. 2. The magnitude of dissipation rate is directly proportional to the Laplacian (∇^2) of the fluctuating component (Eq. 23). Because the fluctuating component is approximated as linear combination of modes (Eq. 2), the modes with higher spatial variance contribute more towards the dissipation rate as compared to the modes with smaller spatial variance.^{11,29} Whereas, the production term is directly proportional to the Laplacian (∇^2) of the mean component. Thus, a truncated set of only lower order POD modes accurately capture the production term, but under-resolve the dissipation term. On the other hand, both the production and dissipation terms are captured by a set of sparse modes with reasonable accuracy due to their multi-scale nature. Furthermore, the time-integrated terms converge to the DNS values as the number of modes are increased, with negligible convection and transfer terms in all the cases. It is observed that the sets of modes that capture the production and dissipation terms with same level of accuracy, result in reasonably good predictions from corresponding ROMs. These trends suggest that there is a direct relationship between balance in energy in the approximated flow solutions using a set of modes, and performance of the corresponding ROM.

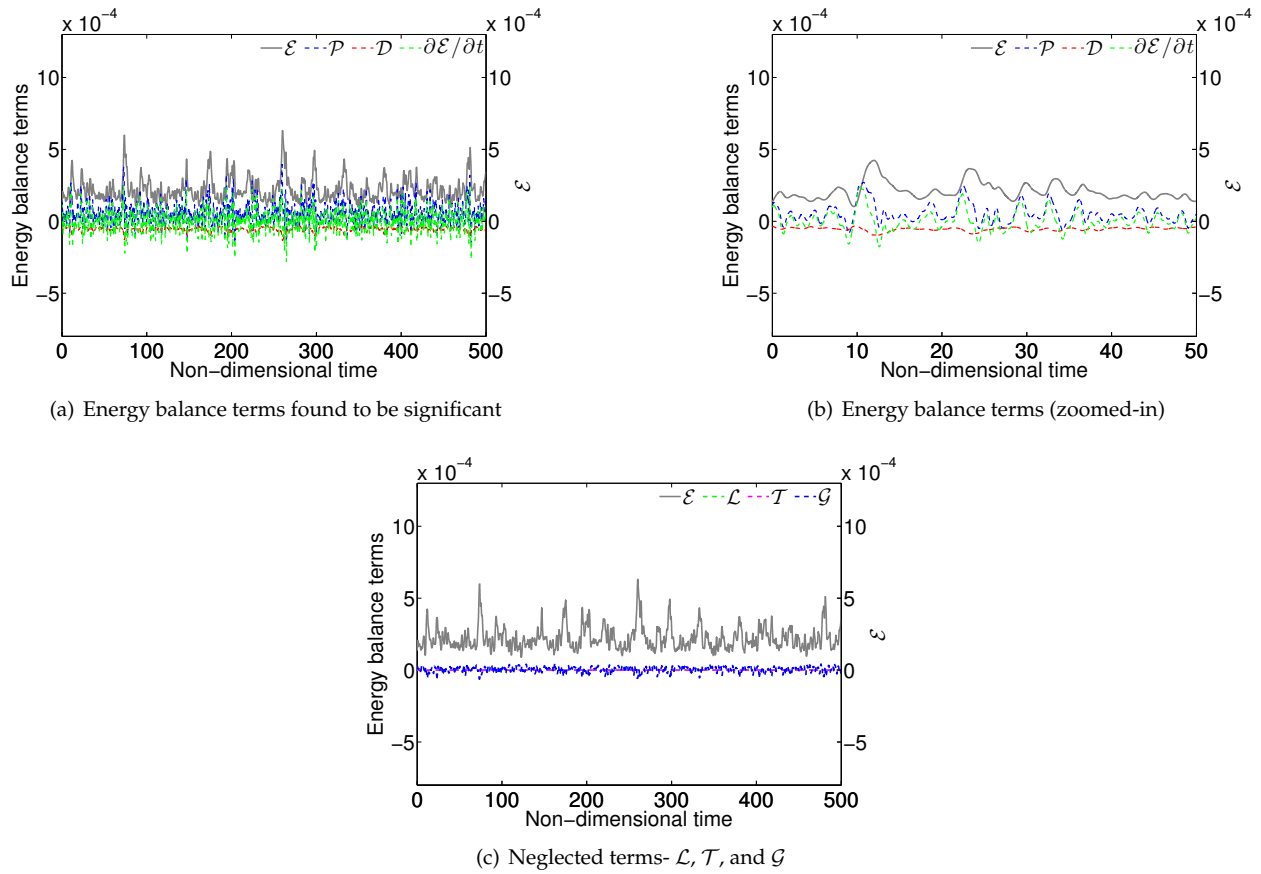


Figure 14. Instantaneous energy balance terms obtained for the DNS flow solutions.

III.G. Computational Expense

All the simulations are performed in serial on an Intel(R) Xeon(R) E5-2620 v2 @ 2.10GHz processor. The cost for generating the 500 units of DNS data is approximately 79.25 hours. Whereas, even the most expensive 200-mode ROMs provide an online speedup of 37 times; where the speed-up values are obtained by dividing the cost of CFD by the combined cost of ROM integration and output operations. The cost of generating the Galerkin matrices is proportional to the cube of the number of modes used. This results in a very high computational cost of generating a ROM when a large number of modes are used. Therefore, it is highly desirable to minimize the number of modes used in ROM construction, while maintaining reasonable accuracy in predictions. Reasonably accurate solutions are obtained using the 20-mode sparse ROM with an online speedup of 220 times. It is observed that the computations of the sparse modes is one order of magnitude more expensive than generating the POD modes.

IV. Conclusions

The generation of the POD and sparse modes is described with an example of 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. Performances of the POD and sparse ROMs are compared in terms of evolution of the turbulent kinetic energy, RMS of the velocities over the computational domain, and local dynamics of the predicted flow. Moreover, thorough energy balance analyses of the ROM predictions and approximations using different sets of POD and sparse modes are carried out. The results of these studies allow one to reach several useful conclusions, namely:

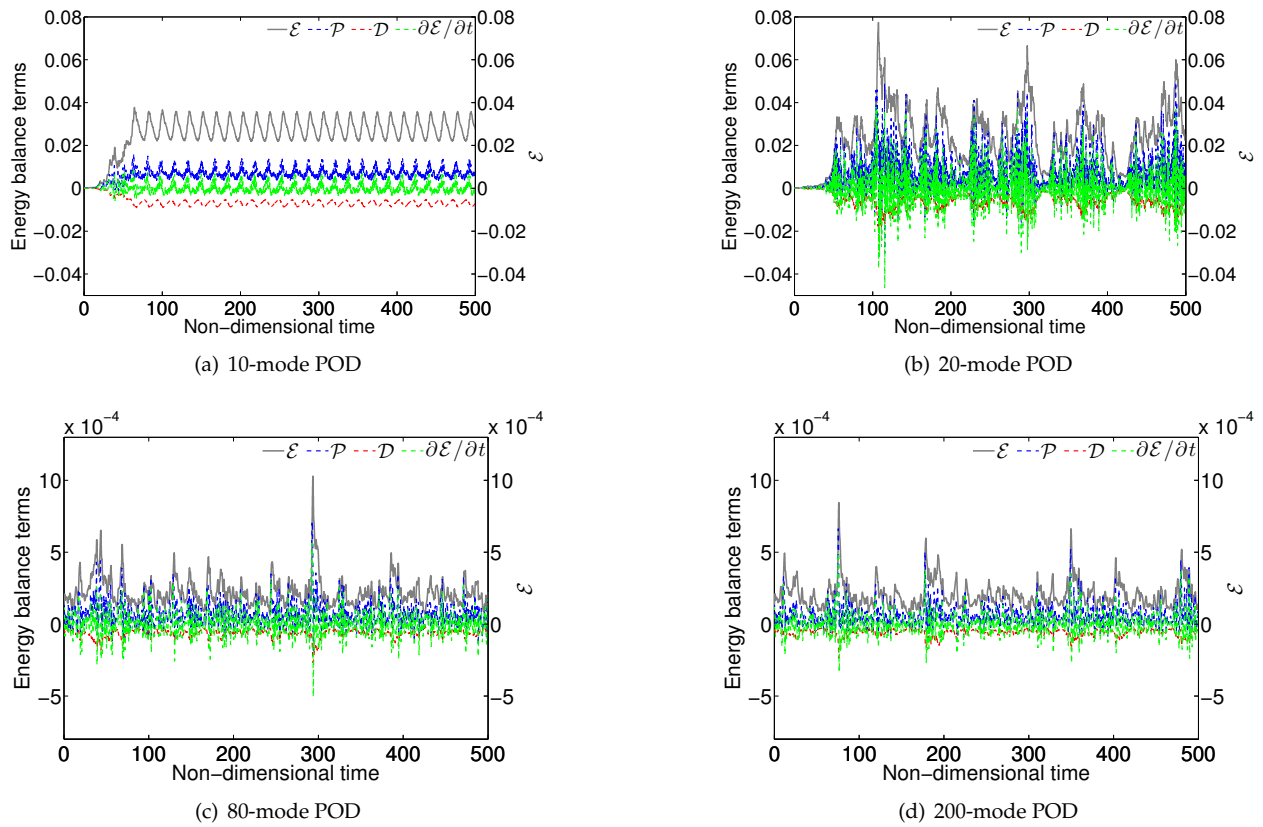


Figure 15. Instantaneous energy balance terms computed for the predicted solutions from POD ROMs.

1. Sparse modes are inherently multi-scale, which yields spatial features that balance destabilizing dominant energy features and dissipation due to small-scale, low-energy features.
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 temporally local dynamics (transience) of the system. On the contrary, the POD process yields a set of modes active at all times.
3. The low order sparse-Galerkin models out-perform the POD-Galerkin models in predicting the flow states in the presented study. The sparse ROMs provide more accurate predictions for TKE as compared to the POD ROMs of same order, when less than 80 modes are used. A ROM constructed with only 20 sparse modes is found to be adequate in capturing the instantaneous TKE, RMS of velocities, and spatially local dynamics of the flow within the lid-driven cavity. Whereas, as many as 80 POD modes are required to obtain reasonably accurate flow predictions.
4. The energy balance analyses of the ROM predictions show that lower order POD ROMs under-resolve the dissipation, resulting in over-prediction of TKE. All the sparse ROMs maintain the energy balance.
5. Energy balance analysis of the time-integrated energy balance terms obtained from the flow approximations using set of reduced order modes suggest that there is a relationship between the difference in magnitudes of time-integrated production and dissipation terms and the ROM performance. A large difference between magnitudes of production and dissipation terms results in over-prediction of TKE by the ROM.

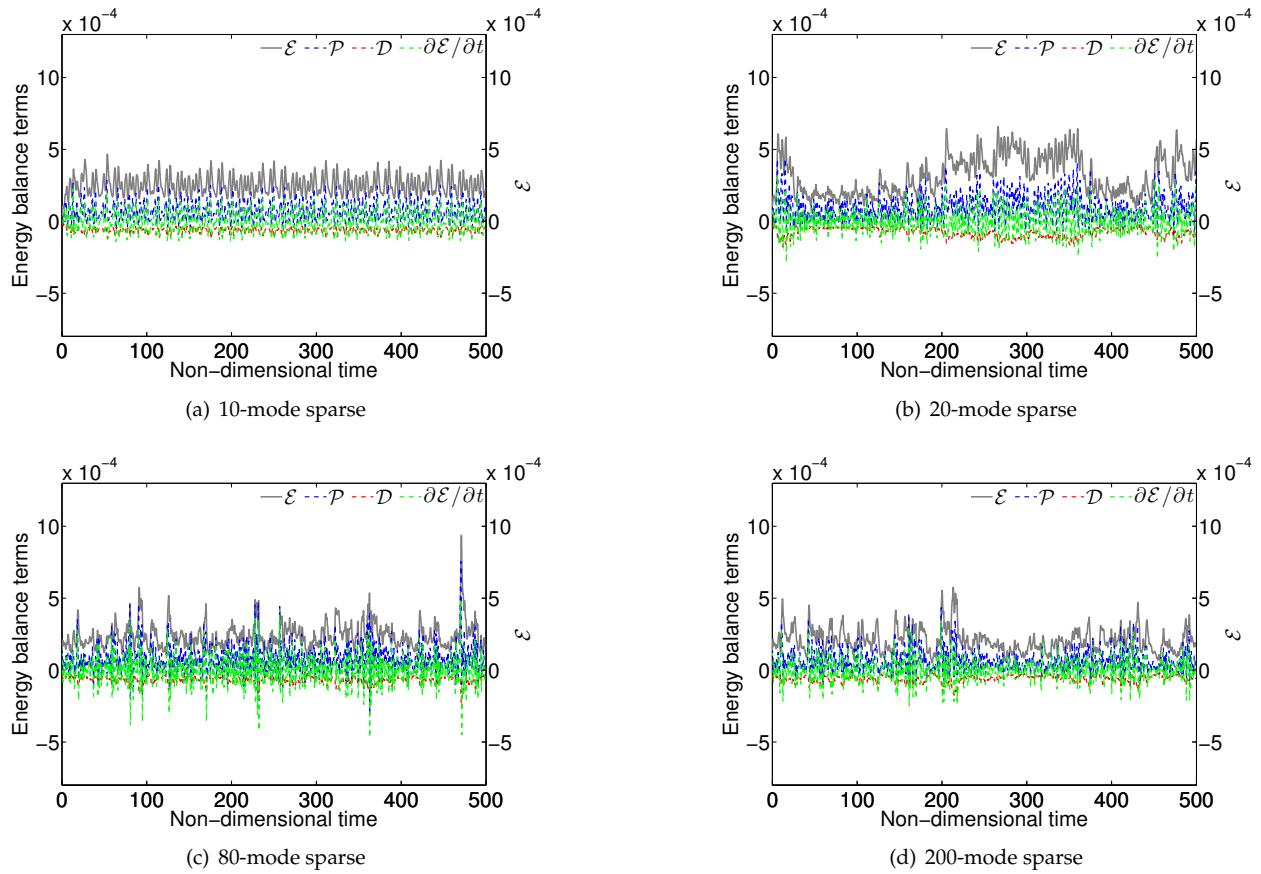


Figure 16. Instantaneous energy balance terms computed for the predicted solutions from sparse ROMs.

6. The computational cost of computing sparse modes is an order of magnitude higher than 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.
7. The increased computational cost of identifying the sparse modes, and the inability of the current approach in retaining the sparse modes used in a lower order ROM, when building a higher order one, are the two observed drawbacks of the method in this study relative to the POD approach.

Acknowledgments

The authors gratefully acknowledge the support of ONR grant N00014-14-1-0018, under the direction of Dr. Judah Milgram, an HPCMPPO Frontier PETTT Project Grant, under the direction of David Bartoe, and an allocation of computing time from the Ohio Supercomputer Center. The authors thank Dr. Lionel Agostini, Mr. Kalyan Goparaju, Mr. S. Unnikrishnan, and Dr. Datta Gaitonde, Mechanical & Aerospace Engineering, The Ohio State University for providing technical guidance in the study of local flow dynamics.

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*

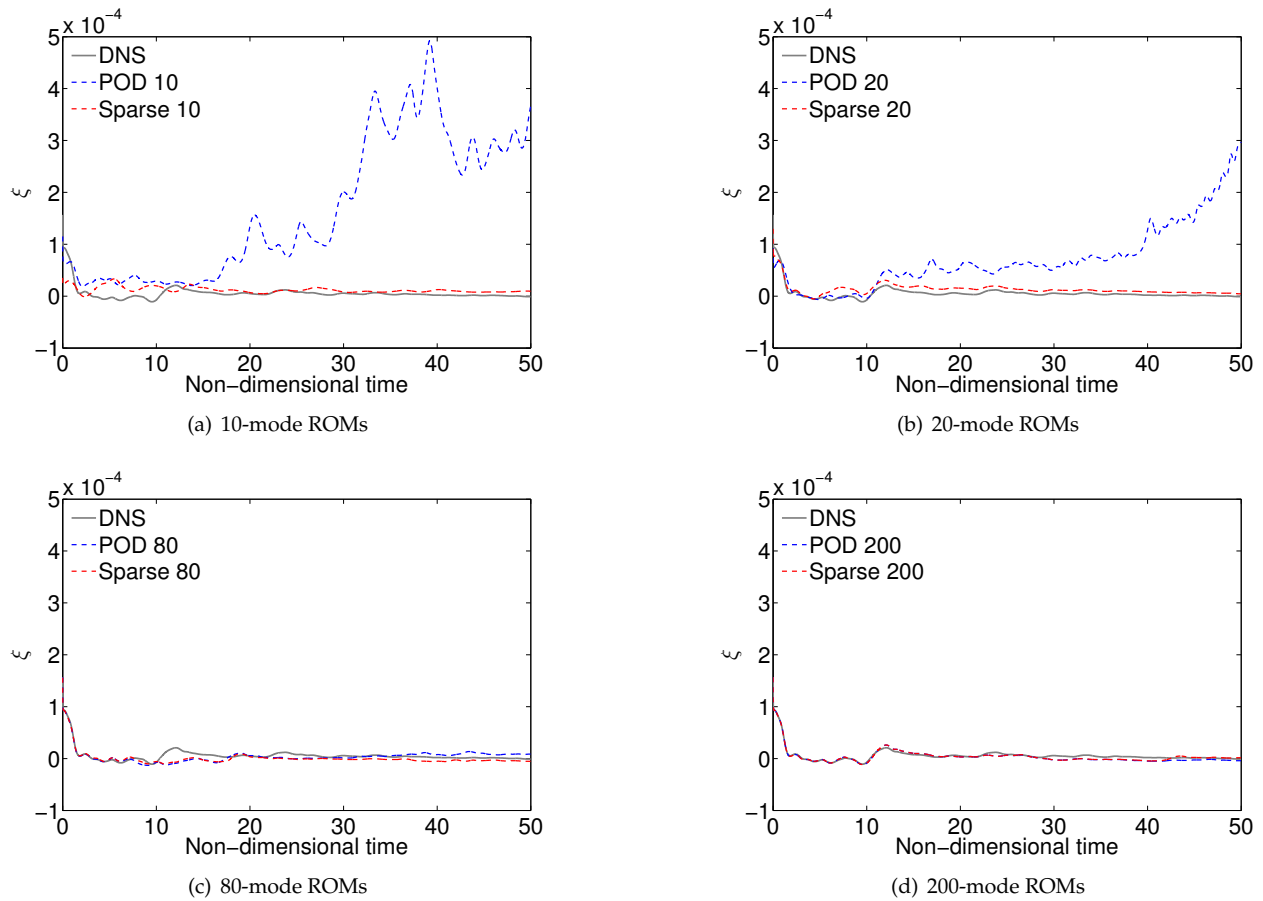


Figure 17. Time averaged $\mathcal{P} + \mathcal{D}$ quantities computed over different time spans starting from $t = 0$.

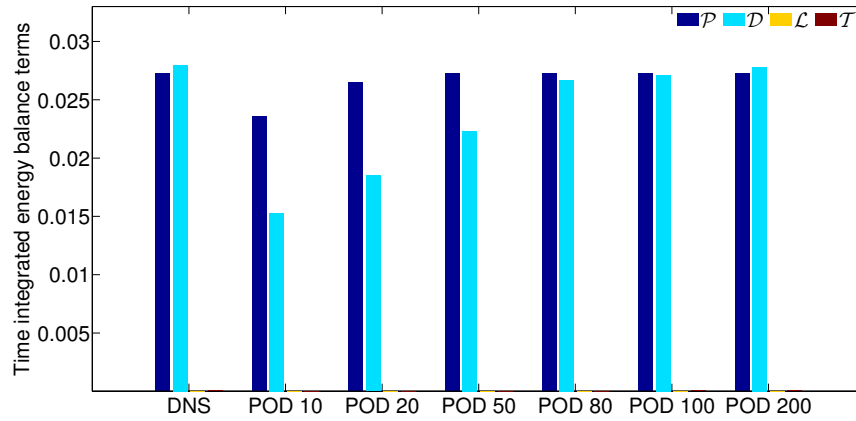


Figure 18. Magnitudes of time averaged energy balance terms computed for POD sets.

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., Morzynski, 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.

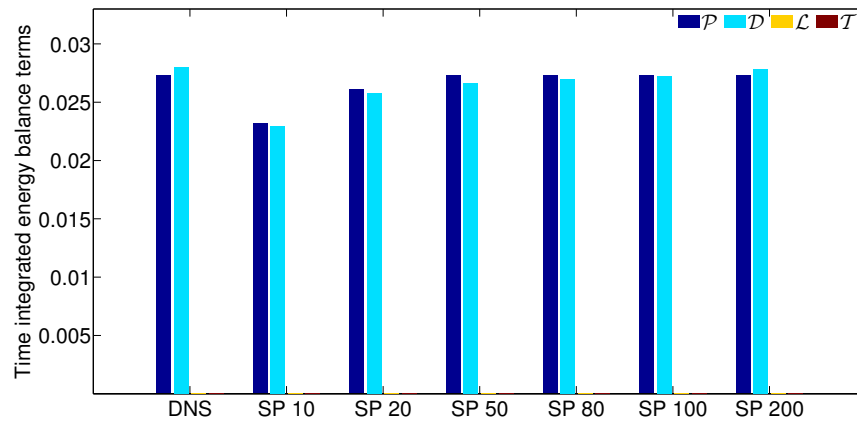


Figure 19. Magnitudes of time averaged energy balance terms computed for sparse sets.

- 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 Field, D., "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.
- ²⁶Jovanović, M. R., Schmid, P. J., and Nichols, J. W., "Sparsity-promoting dynamic mode decomposition," *Physics of Fluids (1994-present)*, Vol. 26, No. 2, 2014, pp. 024103–024124.
- ²⁷Deshmukh, R., Liang, Z., Gogulapati, A., McNamara, J. J., and Kolter, J. Z., "Basis Identification for Reduced Order Modeling of Unsteady Flows Using Sparse Coding," *Proceedings of the 56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference (SciTech 2015)*, No. AIAA 2015-0688, Kissimmee, FL, January 5–9 2015.
- ²⁸Monin, A. S. and Yaglom, A. M., *Statistical Fluid Mechanics: Mechanics of Turbulence. Vol. 1*, MIT press, 1971.
- ²⁹Noack, B. R., Papas, P., and Monkewitz, P. A., "The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows," *Journal of Fluid Mechanics*, Vol. 523, 2005, pp. 339–365.
- ³⁰Friedman, 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.

- ³¹Mittal, R., Dong, H., Bozkurtas, 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.
- ³²Abdi, H. and Williams, L., "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 2, No. 4, 2010, pp. 433–459.
- ³³Zuo, 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.
- ³⁴Tibshirani, R., "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, pp. 267–288.
- ³⁵Friedman, J., Hastie, T., and Tibshirani, R., "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, Vol. 9, No. 3, 2008, pp. 432–441.
- ³⁶Engan, 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.
- 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.
- ³⁹Bendat, J. S. and Piersol, A. G., *Random Data: Analysis and Measurement Procedures*, John Wiley & Sons, Hoboken, New Jersey, 4th ed., February 2010.
- ⁴⁰Welch, P., "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, 1967, pp. 70–73.