

A Randomized Proper Orthogonal Decomposition Technique

Dan Yu and Suman Chakravorty

Abstract—In this paper, we consider the problem of model reduction of large scale systems, such as those obtained through the discretization of PDEs. We propose a randomized proper orthogonal decomposition (RPOD) technique to obtain the reduced order models by randomly choosing a subset of the inputs/outputs of the system to construct a suitable small sized Hankel matrix from the full Hankel matrix. It is shown that the RPOD technique is computationally orders of magnitude cheaper when compared to techniques such as the Eigensystem Realization Algorithm (ERA)/Balanced proper orthogonal decomposition (BPOD) while obtaining the same information in terms of the number and accuracy of the dominant modes. The method is tested on a linearized channel flow problem.

I. INTRODUCTION

In this paper, we consider the problem of model reduction of systems that are governed by partial differential equations (PDE). We propose a randomized version of the snapshot proper orthogonal decomposition technique that allows us to form a reduced order model (ROM) of the PDE of interest in terms of the eigenfunctions of the PDE operator by randomly choosing a subset of the input/output snapshot ensembles, and as a consequence, constructing a sub-matrix of the full Hankel matrix. The RPOD procedure requires orders of magnitude less computation when compared to the Balanced proper orthogonal decomposition (BPOD)/ Eigensystem Realization Algorithm (ERA) procedure applied to the full-order Hankel matrix resulting from the discretization of a PDE with a large number of inputs and outputs. The technique is applied to a linearized channel flow problem to illustrate the procedure.

Model reduction has attracted considerable attention in the past several decades. It is a technique that constructs a lower-dimensional subspace to approximate the original higher-dimensional dynamic system. Balanced POD [1], [2] is a model reduction technique based on the balanced truncation [3] and the snapshot POD technique [4]. Balancing transformations are constructed using the impulse responses of both the primal and adjoint system, and hence, the most controllable and observable modes can be kept in the ROM. In 1978, Kung [5] presented a new model reduction algorithm in conjunction with the singular value decomposition (SVD) technique, and the ERA [6] was developed based on this technique. The BPOD is equivalent to the ERA procedure [7], and forms the Hankel matrix using the primal and adjoint system simulations as opposed to the input-output data as in ERA. More recently, there has been work on obtaining information regarding the dominant modes of a system, based on the snapshot POD followed by a diagonalization of the ROM matrix to extract the modes, called the dynamic mode decomposition (DMD) [8], [9].

The primary drawback of BPOD and ERA is that for a large scale system, such as that obtained by discretizing a PDE, with a large number of inputs/outputs, the computational burden incurred is very high. There are two main parts to the computation: first is to collect datasets from computationally expensive primal and adjoint simulation in order to generate the Hankel matrix. The second part is to solve the singular value decomposition (SVD) problem for the resulting Hankel matrix. Thus, our primary goal in this paper is to reduce the computation required to obtain these ROMs without losing accuracy.

Improved algorithms based on BPOD have been proposed to reduce the computational cost of constructing full Hankel matrix. For example, [2] proposed an output projection method to address the problem when the number of outputs is large. The outputs are projected onto a small subspace via an orthogonal projection P_r that minimizes the error between the full impulse response and the projected impulse response. However, the method cannot make any claim regarding the closeness of the solution to one that is obtained from the full Hankel matrix, and is still faced with a very high computational burden when the number of inputs is large. There have also been methods proposed [10] to reduce the number of snapshots, however, the primary problem regarding large number of inputs/ outputs remains the same. In contrast, we show that by randomly sampling the inputs and outputs, we solve a much smaller problem while extracting almost the same information about the system as would be from the full Hankel matrix.

There are two major classes of randomization algorithms used for low-rank matrix approximations and factorizations: random sampling algorithms and random projection algorithms. For a large scale matrix H , random sampling algorithms construct a rank k approximation matrix \hat{H} by choosing and rescaling some columns of H according to certain sampling probabilities [11], so the error satisfies $\|H - \hat{H}\|_F \leq \|H - H_{(k)}\|_F + \epsilon \|H\|_F$, with high probability, where $H_{(k)}$ is a best rank k approximation of H , ϵ is a specified tolerance, and $\|H\|_F$ denotes the Frobenius norm of H . This is not a suitable error bound when $\|H\|_F$ is large. Thus, in [12], columns are sampled according to leverage scores, where the leverage scores are calculated by performing the SVD of H , so that the error satisfies $\|H - \hat{H}\|_F \leq (1 + \epsilon) \|H - H_{(k)}\|_F$, with high probability. A direct application of both algorithms would require the full Hankel matrix to be constructed, however, such a construction of the Hankel matrix is computationally prohibitive when the number of inputs/outputs is large. Further, the leverage scores are calculated by performing the SVD of the Hankel matrix,

which is also computationally prohibitive owing to the size of the problem.

In random projection method [13], the large matrix H is projected on to an orthonormal basis Q such that the error satisfies $\|H - QQ^*H\| \leq (1 + \epsilon)\|H - H_{(k)}\|$ with high probability, where $\|H\|$ denotes the spectral 2-norm of H . A gaussian test matrix Ω is generated, and the orthonormal basis Q is constructed by performing a QR factorization of the matrix product $H\Omega$. The bottleneck of this algorithm remains, as above, the construction of the full Hankel matrix, which is prohibitively expensive. Moreover, the results are based on the Gaussian property of the sampling matrix Ω , which is not satisfied in our case, since we do not want to, and cannot sample every primal/dual trajectory.

There has been great interest in the Systems and Control community over the past several years in tractable randomized techniques to solve computationally difficult systems and control design problems [14], [15]. The RPOD technique can be construed as one such technique for the model reduction of large scale dynamical systems. In particular, it is perhaps most closely related to the ‘‘Scenario Method’’ for systems and control design [15]. The scenario method obtains bounds on the number of convex constraints that need to be sampled from an uncountable set of constraints such that the solution to an associated robust control problem can be guaranteed to satisfy an ϵ -fraction of the constraints, with probability greater than $1 - \beta$ where ϵ, β are design parameters. In RPOD, we derive a bound for the total number of columns that need to be sampled from a low rank matrix (say rank l) containing a large number of columns, given that the columns are spanned by modes $\{v_1, \dots, v_l\}$, such that the sampled matrix has the same rank as the large matrix with probability at least $1 - \beta$, given that the minimum fraction of the columns in which any of the spanning modes v_i is present is $\bar{\epsilon}$. The scenario method obtains the bound $\frac{2}{\epsilon}(\log(\frac{1}{\beta}) + d)$ where d is the dimension (size) of the problem, whereas our bound is $\frac{1}{\bar{\epsilon}}\log(\frac{1}{\beta})$ where the rank l is the size of our problem. The derivation of our bound, albeit different from the bound in [15], is nonetheless inspired by the developments in that reference.

We had introduced an iterative POD method (I-POD) in [16], [17] that recursively obtains eigenfunctions of a linear operator. We use one input trajectory and one output trajectory to construct a small sub-Hankel matrix, and extract the dominant modes corresponding to this data set. We extract all the relevant modes from all the input/output trajectories by doing I-POD recursively. The main practical problem associated with the I-POD is the orthogonality between different modes. Numerical error exists when extracting dominant modes from different trajectories, and therefore reorthogonalization is needed, which typically leads to numerical instability, and thus, the practical performance of the method is significantly worse than comparable methods such as BPOD.

The main contribution of the RPOD method is that by randomly choosing some of the input/output trajectories and snapshots, we only collect parts of the input/output

ensemble to form a suitable sub-matrix of the full Hankel matrix. Thus, the computations required to form the sub-Hankel matrix, and the subsequent SVD, is computationally orders of magnitude less expensive when compared to the construction/ SVD of the full Hankel matrix. We show that the sub-Hankel matrix we construct retains almost the same information as the full Hankel matrix in terms of the numbers and accuracy of the underlying modes.

The rest of the paper is organized as follows. In Section II, we briefly show how to construct the eigenfunctions of the original system using the cross-correlation matrix between the input and output impulse responses. In Section III, we introduce the randomized proper orthogonal decomposition (RPOD) method where we randomly choose a subset of the inputs/outputs of the system to construct a sub-Hankel matrix when the number of inputs/outputs are large. Then we show that such an approximation contains the same information that is contained in the full Hankel matrix in terms of the dominant modes, given that the number of sampled inputs/outputs satisfies a certain bound. In Section IV, we provide computational results comparing the RPOD with the BPOD for a linearized channel flow problem.

II. EIGENFUNCTION RECONSTRUCTION TECHNIQUE

Consider a stable linear input-output system

$$\begin{aligned} x_k &= Ax_{k-1} + Bu_k, \\ z_k &= Cx_k, \end{aligned} \quad (1)$$

where $x_k \in \mathbb{R}^N$, $u_k \in \mathbb{R}^p$, $z_k \in \mathbb{R}^q$ are the state, inputs, and outputs at discrete time instant t_k respectively. $B = [b_1, \dots, b_p]$ is the input influence matrix and $C = [c_1, \dots, c_q]'$ is the output influence matrix, x' denotes the transpose of x . The dimension of the state N is very large. In the case of a PDE, the above system is obtained via a suitable discretization of the PDE using techniques such as Finite Elements (FE)/ Finite Differences (FD).

In this section, we introduce an eigenfunction reconstruction technique based on BPOD. The eigenfunction of the POD operator that are present in input/output data are reconstructed and used as a reduced order basis. This helps us in distinguishing underlying invariant modes when we implement the RPOD algorithm introduced in section III.

The impulse response of the primal system is collected by using b_j , $j = 1, 2, \dots, p$, as initial conditions for the simulation of the system

$$x_k = Ax_{k-1}. \quad (2)$$

We take M_1 snapshots across the trajectories at time t_1, t_2, \dots, t_{M_1} , and construct the primal snapshot ensemble $X = [x_1(t_1), \dots, x_1(t_{M_1}), \dots, x_p(t_1), \dots, x_p(t_{M_1})] \in \mathbb{R}^{N \times pM_1}$, where $x_j(t_k)$ is the state snapshot at time t_k with b_j as the initial condition, $k = 1, 2, \dots, M_1$ and $j = 1, 2, \dots, p$.

Similarly, we use the transposed rows of the output matrix, c'_i , $i = 1, 2, \dots, q$, as the initial conditions for the simula-

tions of the adjoint system A' ,

$$y_k = A' y_{k-1}, \quad (3)$$

and M_2 snapshots are taken across trajectories at time $\hat{t}_1, \dots, \hat{t}_{M_2}$, leading to the adjoint snapshot ensemble $Y = [y_1(\hat{t}_1), \dots, y_1(\hat{t}_{M_2}), \dots, y_q(\hat{t}_1), \dots, y_q(\hat{t}_{M_2})] \in \mathbb{R}^{N \times q M_2}$, where $y_i(\hat{t}_k)$ is the state snapshot of the adjoint system at time \hat{t}_k with c'_i as the initial condition, $k = 1, 2, \dots, M_2$ and $i = 1, 2, \dots, q$. The Hankel matrix H is defined as:

$$H = Y'X. \quad (4)$$

First, we solve the SVD problem of the matrix H :

$$H = U\Sigma V'. \quad (5)$$

Assume that Σ_p consists of the first l non-zero singular values of Σ , and (U_p, V_p) are the corresponding left and right singular vectors from (U, V) , then the POD projection matrices can be defined as:

$$T_r = X V_p \Sigma_p^{-\frac{1}{2}}, T_l = Y U_p \Sigma_p^{-\frac{1}{2}}. \quad (6)$$

T_r and T_l are the BPOD bases, and the ROM constructed using BPOD is:

$$\tilde{A} = T_l' A T_r, \tilde{B} = T_l' B, \tilde{C} = C T_r. \quad (7)$$

We can see that the POD bases T_r and T_l change when the collected snapshots X and Y are changed, thus, T_r and T_l constructed using BPOD method are not invariant to the datasets X and Y . We want to construct a global set of POD bases which remains invariant to the particular snapshots X and Y .

Assume that \tilde{A} in (7) has a full set of distinct eigenvectors. Let (Λ_{ij}, P) represent the eigenvalue-eigenvector pair for \tilde{A} , i.e.,

$$\tilde{A} = T_l' A T_r = P \Lambda_{ij} P^{-1}, \quad (8)$$

Thus it follows that

$$\Lambda_{ij} = \underbrace{(P^{-1} T_l')}_\Phi A \underbrace{(T_r P)}_\Psi. \quad (9)$$

Here, T_r, T_l are the POD transformation bases and P is the ROM eigenfunction matrix. The transformation Ψ_{ij}, Φ_{ij} denote the composite transformation from the original state space to the POD eigenfunction space, and in turn to the ROM eigenfunction space. Thus, the reduced order model is:

$$\begin{cases} A_r = \Lambda_{ij} = \Phi_{ij}' A \Psi_{ij}, \\ B_r = \Phi_{ij}' B, \\ C_r = C \Psi_{ij}. \end{cases} \quad (10)$$

In the following, we relate the eigenvalues and eigenvectors of A to the diagonal form Λ_{ij} and the transformation Ψ_{ij}, Φ_{ij} .

Suppose the snapshot ensembles $X \in \mathbb{R}^{N \times p M_1}$ and $Y \in \mathbb{R}^{N \times q M_2}$ are spanned by r_1 and r_2 right/left eigenvectors of A respectively. Since some of the eigenvectors will decay

fast, and are not dominant in the M_1 and M_2 snapshot ensembles, thus, $r_1 \leq N$ and $r_2 \leq N$. Also, we take enough snapshots so that $r_1 \leq M_1$ and $r_2 \leq M_2$. Notice that the active left and right eigenvectors in the snapshots Y and X may not be the same, so we denote $X = V_S \alpha_S + V_D \delta \alpha_D$, $Y = U_S \beta_S + U_D \delta \beta_D$, where (U_S, V_S) are the active left and right eigenvectors corresponding to the same eigenvalues Λ_S , (U_D, V_D) are the rest of the left and right eigenvectors, and $\alpha_S, \beta_S, \delta \alpha_D, \delta \beta_D$ are the coefficient matrices.

Assumption 1: We assume that the contributions of the left and right eigenvectors corresponding to the different eigenvalues are small, i.e., $\|\delta \alpha_D\| \leq C_1 \epsilon$, $\|\delta \beta_D\| \leq C_2 \epsilon$, where C_1, C_2 are some constant, and ϵ is sufficiently small.

Under Assumption 1, the following result holds.

Proposition 1: Denote (Λ_S, U_S, V_S) as the actual eigenvalues, left and right eigenvectors of A which are active in both sets of snapshots X and Y . The errors in eigenvalue and eigenvector reconstruction are $\|\Lambda_{ij} - \Lambda_S\| \leq k_1 \epsilon^2$, $\|\Phi_{ij} - U_S\| \leq k_2 \epsilon$, and $\|\Psi_{ij} - V_S\| \leq k_3 \epsilon$, i.e., $(\Lambda_{ij}, \Phi_{ij}, \Psi_{ij})$ are arbitrarily good approximation of the eigenvalues, left and right eigenvectors active in both sets of snapshots X and Y , where k_1, k_2, k_3 are some constants, ϵ is defined in Assumption 1, and is sufficiently small.

The proof uses the eigenvalue perturbation theory [18], and is shown in Appendix .

Remark 1: If Assumption 1 is not satisfied, we can still prove that $\|\Lambda_{ij} - \Lambda_S\| \leq k \epsilon$, where k is some constant. However, $\Psi_{ij} = V_S + V_{c\bar{o}} \Delta_{c\bar{o}}$, and $\Phi_{ij} = U_S + U_{\bar{c}o} \Delta_{\bar{c}o}$. Here, $V_{c\bar{o}}$ are the most controllable but not observable right eigenvectors, and $U_{\bar{c}o}$ are the most observable but not controllable left eigenvectors, and $\Delta_{c\bar{o}}$ and $\Delta_{\bar{c}o}$ are suitable coefficient matrices. Thus, the Markov parameters of the ROM are:

$$\begin{aligned} \hat{h}_k &= C \Psi_{ij} (\Lambda_{ij})^k \Phi_{ij}' B \\ &= C (V_S + V_{c\bar{o}} \Delta_{c\bar{o}}) (\Lambda_{ij})^k (U_S' + \Delta_{\bar{c}o}' U_{\bar{c}o}') B. \end{aligned} \quad (11)$$

However, $C V_{c\bar{o}} \approx 0$ because of the unobservability of $V_{c\bar{o}}$, and $U_{\bar{c}o}' B \approx 0$ because of the uncontrollability of $U_{\bar{c}o}$. Thus, the ROM Markov parameters

$$\hat{h}_k \approx C V_S (\Lambda_S)^k U_S' B, \quad (12)$$

are the same as those that would be constructed using only (V_S, U_S) without the spillover from $V_{c\bar{o}}$ and $U_{\bar{c}o}$, i.e., in terms of the impulse response, the spillover from $V_{c\bar{o}}$ and $U_{\bar{c}o}$ does not matter.

III. RANDOMIZED PROPER ORTHOGONAL DECOMPOSITION METHOD

From Section II, we see that we can construct POD bases, and extract the underlying eigenvectors of the original system, which are invariant to the particular primal and adjoint datasets X and Y . Assume that the rank of the full Hankel matrix $H = Y'X$ is l . Since the dimension of the systems governed by PDEs may be very large due to the discretization, the computation to construct the Hankel matrix and solve the SVD problem is very expensive, especially when there are a large number of inputs/ outputs.

The eigenfunction reconstruction technique from Section II suggests that if we can construct a sub-Hankel matrix \hat{H} which is still rank l , then the underlying l eigenmodes can be recovered from the sub-Hankel matrix. Thus, in this section, we introduce a randomized proper orthogonal decomposition (RPOD) method based on the eigenfunction reconstruction technique which randomly chooses a small subset of the inputs/ outputs, and constructs a sub-Hankel matrix from the full Hankel matrix such that the information encoded in the sub-Hankel matrix is almost the same as that in the full Hankel matrix, in terms of the number and accuracy of the underlying modes that can be extracted.

Consider the stable linear system (1), we randomly choose r columns from B according to the uniform distribution, denoted as \hat{B} , and randomly choose s rows from C with uniform distribution, denoted as \hat{C} . Denote $(\cdot)_{(:,i)}$ as the i th column of (\cdot) , and $(\cdot)_{(j,:)}$ as the j th row of (\cdot) .

The original Hankel matrix H was previously defined in (4). The reduced order Hankel matrix \hat{H} is then constructed using \hat{B} , \hat{C} and randomly choose m_1 and m_2 snapshots from the primal and adjoint snapshots (t_1, \dots, t_{M_1}) and $(\hat{t}_1, \dots, \hat{t}_{M_2})$ with uniform distribution respectively. Thus, the RPOD technique can be seen as randomly choosing $\hat{p}m_1$ columns from the H matrix to form the \hat{H} matrix, and then randomly choosing $\hat{q}m_2$ rows from the \hat{H} matrix to form \hat{H} . Alternatively, it essentially is equivalent to choosing a suitable random subset of the columns of the primal/ adding responses, namely \hat{X} and \hat{Y} to generate the sub-Hankel matrix $\hat{H} = \hat{Y}'\hat{X}$. The RPOD procedure is summarized in Algorithm 1.

First, we provide a general result regarding randomly choosing a rank “ l ” sub-matrix from a large rank “ l ” matrix. Suppose $W \in \mathfrak{R}^{N \times a}$ is a rank l matrix, and suppose that W is spanned by the vectors $\{v_1, v_2, \dots, v_l\}$, where $v_i \in \mathfrak{R}^N$, $l \ll N, a$. Let $W^{(i)}$ denote the set of columns of W that contain the vector v_i . Let

$$\epsilon_i = \frac{\text{no.}(W^{(i)})}{N}, \quad (13)$$

denote the fraction of the columns in W in which vector v_i is present. Further let

$$\bar{\epsilon} = \min_i \epsilon_i, \quad (14)$$

and note that $\bar{\epsilon} > 0$.

Proposition 2: Let M columns be sampled uniformly from among the columns of the matrix W without replacement, and denote the sampled sub-matrix by \hat{W} . Let $(\Omega, \mathcal{F}, P_f)$ denote the underlying probability space for the experiment. Given any $\beta > 0$, if the number M is chosen such that

$$M > \max(l, \frac{1}{\bar{\epsilon}} \log(\frac{l}{\beta})), \quad (15)$$

then $P_f(\rho(\hat{W}) < l) < \beta$, where $\rho(\hat{W})$ denotes the rank of the sampled matrix \hat{W} .

Proof: Let $\hat{W}(\omega) = \{W_1(\omega), \dots, W_M(\omega)\}$ denote a random M -choice from the columns of W . If the ensemble

Algorithm 1 RPOD Algorithm

- 1) Pick $c_i \in \{1, \dots, p\}$ with probability $P[c_i = k] = \frac{1}{p}, k = 1, \dots, p, i = 1, \dots, \hat{p}$
 - 2) Pick $r_j \in \{1, \dots, q\}$ with probability $P[r_j = k] = \frac{1}{q}, k = 1, \dots, q, j = 1, \dots, \hat{q}$
 - 3) Set $\hat{B}_{(:,i)} = B_{(:,c_i)}, \hat{C}_{(j,:)} = C_{(r_j,:)}$
 - 4) Randomly choose m_1 snapshots from $(t_1, t_2, \dots, t_{M_1})$ with uniform distribution, denoted as $t = \tilde{t}_1, \dots, \tilde{t}_{m_1}$. Use $\hat{B}_{(:,i)}, i = 1, \dots, \hat{p}$ as the initial conditions for the primal simulation, collect the snapshots at $t = \tilde{t}_1, \dots, \tilde{t}_{m_1}$, denoted as \hat{X}
 - 5) Randomly choose m_2 snapshots from $(\hat{t}_1, \dots, \hat{t}_{M_2})$ with uniform distribution, denoted as $t = \tilde{t}_1, \dots, \tilde{t}_{m_2}$. Use $\hat{C}'_{(j,:)}, j = 1, \dots, \hat{q}$ as the initial conditions for the adjoint simulation, collect the snapshots at $t = \tilde{t}_1, \dots, \tilde{t}_{m_2}$, denoted as \hat{Y}
 - 6) Construct the reduced order Hankel matrix $\hat{H} = \hat{Y}'\hat{X}$
 - 7) Solve the SVD problem of $\hat{H} = \hat{U}_p \hat{\Sigma}_p \hat{V}_p'$
 - 8) Construct the POD basis: $\hat{T}_r = \hat{X} \hat{V}_p \hat{\Sigma}_p^{-1/2}, \hat{T}_l = \hat{Y}' \hat{U}_p' \hat{\Sigma}_p^{-1/2}$
 - 9) Construct the matrix: $\tilde{A} = \hat{T}_l' \hat{A} \hat{T}_r$, and $(\hat{\Lambda}, \hat{P})$ are the eigenvalues and eigenvectors of \tilde{A}
 - 10) Construct new POD basis: $\hat{\Phi}'_{ij} = \hat{P}^{-1} \hat{T}_l'$ and $\hat{\Psi}_{ij} = \hat{T}_r \hat{P}$
 - 11) The ROM is: $\hat{A}_r = \hat{\Phi}'_{ij} \hat{A} \hat{\Psi}_{ij}, \hat{B}_r = \hat{\Phi}'_{ij} B, \hat{C}_r = C \hat{\Psi}_{ij}$
-

\hat{W} has rank less than l then note that at least one of the vectors v_i has to be absent from the ensemble. Define the events

$$G = \{\omega \in \Omega : \rho(\hat{W}(\omega)) < l\}, \text{ and} \quad (16)$$

$$G_i = \{\omega \in \Omega : W_k(\omega) \in \tilde{W}^{(i)}, \forall k\}, \quad (17)$$

where $\tilde{W}^{(i)}$ denotes the complement set of columns in W to the set $W^{(i)}$. Due to the fact that the ensemble \hat{W} is rank deficient if all of the columns of \hat{W} are sampled from at least one of the sets $\tilde{W}^{(i)}$, and the fact that if \hat{W} is rank deficient, all the columns of \hat{W} have to be sampled from at least one of the sets \tilde{W}_i , it follows that:

$$G = \bigcup_i G_i. \quad (18)$$

If we sample the M columns with replacement, $P_f(G_i) = (1 - \epsilon_i)^M$, and $P_f(G) \leq (1 - \epsilon_i)^M$ if we sample the M columns without replacement. Thus, it follows that

$$P_f(G) \leq \sum_{i=1}^l P_f(G_i) = \sum_{i=1}^l (1 - \epsilon_i)^M \leq l(1 - \bar{\epsilon})^M. \quad (19)$$

Hence, it follows that $P_f(\rho(\hat{W}) < l) \leq l(1 - \bar{\epsilon})^M$. If we require this probability to be less than some given $\beta > 0$, then, it can be shown by taking log on both sides of the above expression that M should satisfy

$$M > \frac{1}{\bar{\epsilon}} \log(\frac{l}{\beta}). \quad (20)$$

Noting that \hat{W} is rank deficient unless $M \geq l$, the result follows. ■

Remark 2: Effect of $l, \bar{\epsilon}$ on the bound M : It can be seen that the number of choices M is influenced primarily by $\bar{\epsilon}$ and not significantly by the number of active modes/ rank of the ensemble l , since l appears in the bound under the logarithm. Thus, the difficulty of choosing a sub-ensemble that is rank l is essentially decided by the fraction $\bar{\epsilon}_i$ of the ensemble in which the rarest vector v_i is present. Moreover, note that as the number l increases, we need only sample $O(l)$ columns to have a rank “ l ” sub-ensemble.

Remark 3: Effect of Sampling non-uniformly: In certain instances, for instance, when we have a priori knowledge, we may choose to sample the columns of W non-uniformly. Define

$$\epsilon_i^\Pi = \sum_{j=1}^N 1_i(W_j) \pi_j, \quad (21)$$

where π_j is the probability of sampling column W_j from the ensemble W , and $1_i(W_j)$ represents the indicator function for vector v_i in column W_j , i.e, it is one if v_i is present in W_j and 0 otherwise. Note that ϵ_i as defined before is the above quantity with the uniform sampling distribution $\pi_j = \frac{1}{N}$ for all j . It is reasonably straightforward to show that Proposition 2 holds with $\bar{\epsilon}^\Pi = \min_i \bar{\epsilon}_i^\Pi$ for any sampling distribution Π (we replace ϵ_i in (19) with $\bar{\epsilon}_i^\Pi$). The effect of a good sampling distribution is to lower the bound M by raising the number $\bar{\epsilon}^\Pi$ over that of a uniform distribution. This may be an intelligent option when otherwise the bound on M with uniform sampling can be very high, for instance when one of the vectors v_i is present in only a very small fraction of the ensemble W . However, we might have some a priori information regarding the columns where v_i may be present and thus, bias the sampling towards that sub-ensemble.

Next, it can be seen how the RPOD procedure extends the above result to the Balanced POD scenario where we consider the Hankel matrix $H = Y'X$, where $H \in \mathbb{R}^{qM_2 \times pM_1}$. The RPOD chooses a small number of inputs/outputs, namely \hat{p}/\hat{q} respectively, and then chooses a small number of times, m_1 for the input and m_2 for the outputs, at which to sample the input/output trajectories, and form the sub-Hankel matrix $\hat{H} \in \mathbb{R}^{\hat{q}m_2 \times \hat{p}m_1}$. This is equivalent to a uniform sampling of the columns of the input and output ensembles X and Y respectively to form $\hat{H} = \hat{Y}'\hat{X}$.

Under Assumption 1, the output and input ensembles Y and X are spanned by l left eigenvectors U_S and right eigenvectors V_S respectively. Define:

$$\bar{\epsilon}_X = \min_i \epsilon_{X,i}, \bar{\epsilon}_Y = \min_j \epsilon_{Y,j}, \quad (22)$$

where $\epsilon_{X,i}$ is the fraction of columns in X in which the right eigenvector v_i is present, and $\epsilon_{Y,j}$ is the fraction of the columns in Y in which the left eigenvector u_j is present.

Note that due to Proposition 2, given any $\beta > 0$, if we choose $\hat{p}m_1$ and $\hat{q}m_2$ satisfy the bounds:

$$\hat{p}m_1 > \max(l, \frac{1}{\bar{\epsilon}_X} \log(\frac{l}{\beta})),$$

$$\hat{q}m_2 > \max(l, \frac{1}{\bar{\epsilon}_Y} \log(\frac{l}{\beta})), \quad (23)$$

then the probability of \hat{H} having rank less than l is less than $\gamma = 1 - (1 - \beta)^2$, since then the probability that the ranks of the sampled input and output ensembles are less than l , is less than β . Thus, if we repeatedly choose K such ensembles with replacement, the probability of having a sub-Hankel matrix \hat{H} that is still less than rank l after the K picks, has to be less than γ^K . Thus, the probability of choosing a rank l sub-Hankel matrix \hat{H} exponentially approaches unity with the number of trials. Again, noting that the value of β does not have a significant influence on the bounds above, it follows that β can be chosen to be quite small without significantly affecting the number of columns that need to be chosen to satisfy the confidence level of β , and thus, the probability of choosing a rank l sub-Hankel matrix can be made arbitrarily high by judiciously choosing the number of columns in the input/ output ensembles according to the bounds in (23).

We summarize the development above in the following proposition.

Proposition 3: Let Hankel matrix $H = Y'X \in \mathbb{R}^{qM_2 \times pM_1}$ with p inputs, q outputs, M_1, M_2 time snapshots in every input and output trajectory respectively. Let the left/ right eigenvectors $U_S = \{u_1, \dots, u_l\}$, and $V_S = \{v_1, \dots, v_l\}$ denote the eigenvectors spanning the input and output ensembles X and Y respectively. Let $\bar{\epsilon}_X, \bar{\epsilon}_Y$ be as defined in (22) and $\beta > 0$ be given. Suppose we construct a sub-Hankel matrix \hat{H} according to the RPOD procedure: by uniformly sampling \hat{p} inputs with m_1 time snapshots, and \hat{q} outputs with m_2 snapshots, and that $\hat{p}m_1$ and $\hat{q}m_2$ are chosen as in (23), then the probability that the sub-Hankel matrix has rank less than l is less than $\gamma = 1 - (1 - \beta)^2$. Moreover, the probability that after K RPOD choices, with replacement, the probability that the sub-Hankel matrix is less than rank l is less than γ^K .

The following corollary immediately follows due to the developments in section II.

Corollary 1: Let (Λ_S, U_S, V_S) be the eigenvalues, left and right eigenvectors underlying the data in the full Hankel matrix. Given any $\beta > 0$, and that a sub-Hankel matrix \hat{H} is chosen as in Proposition 3, the same (Λ_S, U_S, V_S) triple can be extracted from the sub-Hankel matrix \hat{H} with probability at least $(1 - \beta)^2$, and hence, with probability $(1 - \beta)^2$, the information contained in H and \hat{H} is identical in terms of the (Λ_S, U_S, V_S) triple.

Remark 4: Several remarks are made below about the above results.

- 1) The fractions $\bar{\epsilon}_X$ and $\bar{\epsilon}_Y$ are metrics of the “difficulty” of the problem. For instance, if all the relevant modes were controllable/observable from every input/output, then these fractions are unity, and any RPOD choice would have rank l . The lower these fractions are, the higher the number of rows and columns $\hat{q}m_2$ and $\hat{p}m_1$ need to be chosen such that Proposition 3 holds for the sampled sub-Hankel matrix. This corresponds to a mode, or set of modes, being controllable/ observable

only from a very sparse set of actuator/ sensor locations respectively.

- 2) We do not know $\bar{\epsilon}_X, \bar{\epsilon}_Y$ a priori, and thus, we cannot directly apply Proposition 3. In practice, we repeatedly sample sub-Hankel matrices, and check the underlying eigenmodes from each choice. If the underlying modes from different choices are identical, then we can give a guarantee that the Hankel matrix is actually rank l , given a difficulty level $\bar{\epsilon}$. Thus, we are able to quantify the confidence in our ROMs for different values of the difficulty level $\bar{\epsilon}$. Typically, we have seen that if the number of rows/ columns sampled are large enough, we are able to extract all the relevant modes.
- 3) We can also vary the size of the sampled sub-Hankel matrices which in turn raises the probability of sampling a random choice with rank equal to that of the full Hankel matrix.
- 4) If we have a priori knowledge of the system, we can sample the sub-Hankel matrix using some sampling distribution other than the uniform distribution function, which as mentioned previously, has the effect of raising the fractions $\bar{\epsilon}_X, \bar{\epsilon}_Y$, and thus, lower the required size of the sub-Hankel matrix.
- 5) In reality, the Hankel matrix is not exactly rank l but approximately rank l . In such a case, we can appeal to Proposition 1 to show that the errors incurred due to this fact is small if the contribution from the modes other than the dominant l modes are small.

Remark 5: Note that Proposition 3 and Corollary 1 assume that assumption 1 is satisfied. If it is not, then, the primal snapshots X are spanned by $\{V_S, V_{c\bar{o}}\}$ and the adjoint snapshots Y are spanned by $\{U_S, U_{\bar{c}o}\}$, where $V_S = \{v_1, \dots, v_l\}$, $U_S = \{u_1, \dots, u_l\}$, $V_{c\bar{o}}$ are the unobservable modes and $U_{\bar{c}o}$ are the uncontrollable modes. The extracted eigenvalues $\Lambda_{ij} = \Lambda_s = \text{diag}\{\lambda_1, \dots, \lambda_l\}$ still correspond to the most controllable and observable modes. However, the projection vectors (Φ_{ij}, Ψ_{ij}) extracted from H , and $(\hat{\Phi}_{ij}, \hat{\Psi}_{ij})$ from \hat{H} , are no longer the most observable and controllable modes (U_S, V_S) , but contain spillover from the unobservable and uncontrollable modes $V_{c\bar{o}}$ and $U_{\bar{c}o}$ (See Remark 1). However, due to Remark 1, we see that the impulse response of the system obtained from H is the same as that obtained from \hat{H} since: $C\Psi_{ij}\Lambda_{ij}^k\Phi'_{ij}B \approx CV_S\Lambda_S^kU'_S B \approx C\hat{\Psi}_{ij}\Lambda_{ij}^k\hat{\Phi}'_{ij}B$, because of the uncontrollability and unobservability of the modes $U_{\bar{c}o}$ and $V_{c\bar{o}}$ respectively. Hence, in terms of the impulse response, the information extracted from \hat{H} is still the same as that from H , albeit the projections (Φ_{ij}, Ψ_{ij}) and $(\hat{\Phi}_{ij}, \hat{\Psi}_{ij})$ extracted are, in general, different.

IV. COMPUTATIONAL RESULTS

In the following, we will show the comparison of RPOD with Balanced POD for a linearized channel flow problem.

Consider the problem of the fluid flow in a plane channel. We focus on the linearized case when there are small perturbations about a steady laminar flow. The flow is perturbed by body force $B(y, z)f(t)$, which means the force is acting in the wall-normal direction. There is no-slip boundary

condition at the walls $y = \pm 1$ and the flow is assumed to be periodic in the x and z direction. Assume there is no variations in the x direction, then the linearized equation of the wall-normal velocity v and the wall-normal vorticity η are given by:

$$\begin{aligned} \frac{\partial v}{\partial t} &= \frac{1}{R}\nabla^2 v + Bf, \\ \frac{\partial \eta}{\partial t} &= \frac{1}{R}\nabla^2 \eta - U' \frac{\partial v}{\partial z}, \end{aligned} \quad (24)$$

where $R = 100$ is the Reynolds number and $U(y) = 1 - y^2$ is the steady state velocity. The domain $z \in [0, 2\pi]$. We discretize the system using the finite difference method, where both the y direction and z direction are discretized into 21 nodes. Thus, the size of the system is 882×882 . There are 2 constant body forces on $y = 0$, and the measurements are taken on all the nodes on boundaries. For BPOD, we use 80 measurements on the boundaries, and take 1000 snapshots from $t \in [0, 1000s]$ for the primal simulation, 50 snapshots from $t \in [0, 500s]$ for the adjoint simulation, which leads to a 8000×2000 SVD problem. For RPOD, we randomly choose 50 measurements from the 80 measurements on the boundaries, take 200 snapshots from $t \in [0, 200s]$ for the primal simulation, and take 20 snapshots from $t \in [0, 200s]$ for the adjoint simulation. Thus, we need to solve a 2000×400 SVD problem for RPOD. The actual velocity and vorticity at $t = 1000s$ are shown in Fig. 1.

In Fig. 2, we compare the velocity modes of the system using RPOD with the actual velocity modes. The comparison of the vorticity modes are omitted here due to the page limit.

Here, we should note that the sign and the modulus of the ROM velocity modes are not the same as the actual modes, however, if needed, we can rescale the ROM modes to make them match. For both methods, we extract 40 modes, the first 30 extracted eigenvalues are compared in Fig. 3.

The comparison of the state errors and output errors are shown in Fig. 4. To test the ROM, we use 20 different white noise forcings and take the average output/state error over these 20 simulation. We can see that the eigenvalues extracted by RPOD and BPOD are almost the same. In this simulation, we notice that at first, the state error and output error using BPOD are slightly better than using RPOD, but after some time, the errors are almost the same. The output errors using both methods are less than 0.1%, and the state

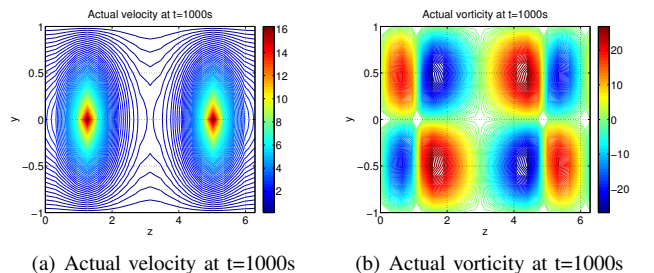


Fig. 1. Actual velocity and vorticity of the channel flow problem

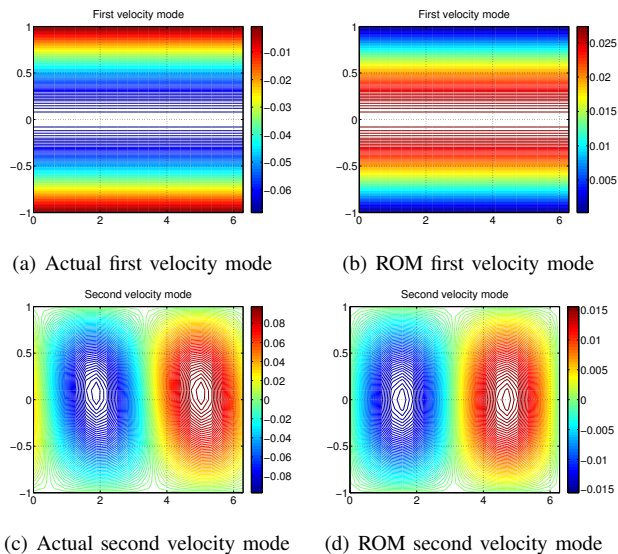


Fig. 2. Comparison between ROM and actual velocity modes

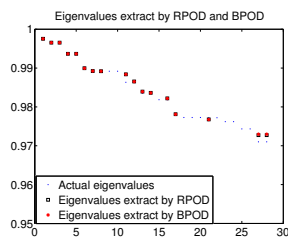


Fig. 3. Comparison of eigenvalues extracted by RPOD and BPOD for channel flow problem

errors using both methods are around 5%. Thus, we can conclude that RPOD is comparable to BPOD but requires far less computation.

Moreover, sometimes, it may be impossible to solve the SVD problem resulting from BPOD. For example, in the channel flow problem, if we use the full state measurements (882 measurements) and we take 20 snapshots for the adjoint simulation, there are 80 sources on the boundary and we take 1000 snapshots for the primal simulation, then we need to solve a 17640×80000 SVD problem for BPOD, which is not solvable in Matlab. For RPOD, we randomly choose 50 sources on the boundaries, and randomly choose 400 measurements. If we take 100 snapshots for the primal

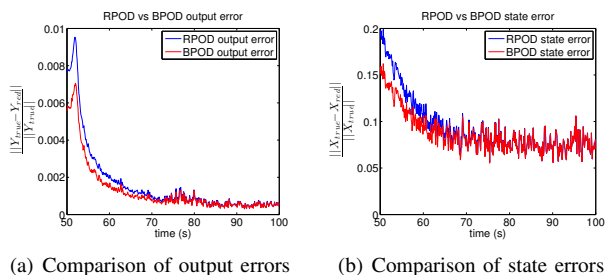


Fig. 4. Comparison between RPOD and BPOD for channel flow problem

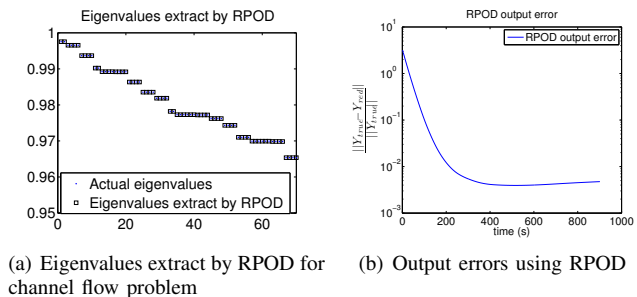


Fig. 5. Simulation results using RPOD for channel flow problem

simulation, and 20 snapshots for the adjoint simulation, then it leads to a 8000×5000 SVD problem, which is a relatively small problem. We compare the first 70 extracted eigenvalues with the actual eigenvalues and the output errors in Fig. 5. Thus, in problems where there are a large numbers of actuators/sensors, the savings can be very significant. In terms of an experiment, this observation may have added implications as it implies that we can reduce the scale of the instrumentation required to get the data required to form an ROM by orders of magnitude without losing much information that can be extracted from the resulting data, which can result in significant cost savings.

V. CONCLUSION

In this paper, we have introduced a randomized POD (RPOD) procedure for the extraction of ROMs for large scale systems such as those governed by PDEs. The RPOD procedure extracts almost the same information from a randomly chosen sub-Hankel matrix extracted from the full order Hankel matrix as is obtained by the BPOD procedure from the full order Hankel matrix without sacrificing too much accuracy. This leads to an orders of magnitude reduction in the computation required for constructing ROMs for large scale systems with a large number of inputs/ outputs over the BPOD procedure. The computational results shown for a set of moderately high dimensional advection diffusion equations seem to reach the same conclusion. The next step in this process would require us to consider more realistic, high dimensional, and nonlinear PDEs arising in problems such as fluid flows and aeroelasticity.

APPENDIX PROOF OF PROPOSITION 1

Here, we establish bounds on the eigenfunction reconstruction errors using the cross correlation matrix $Y^T X$.

We denote $X = V_S \alpha_S + \delta \alpha_D V_D$, and $Y = U_S \beta_S + U_D \delta \beta_D$, where U_S, V_S are the active left and right eigenvectors corresponding to the same eigenvalues Λ_S in the snapshots, and U_D, V_D are rest of the left and right eigenvectors. Under assumption 1, $\|\delta \alpha_D\| \propto O(\epsilon)$, and $\|\delta \beta_D\| \propto O(\epsilon)$, where ϵ is sufficient small. First, we need to solve the SVD problem of $Y^T X$.

$$\begin{aligned} Y^T X &= (\beta_S^T U_S^T + \delta \beta_D^T U_D^T)(V_S \alpha_S + V_D \delta \alpha_D) \\ &= \beta_S^T \alpha_S + \delta \beta_D^T \delta \alpha_D = \beta_S^T \alpha_S + \Delta_1, \end{aligned} \quad (25)$$

where $\|\Delta_1\| \propto O(\epsilon^2)$, and thus $\|Y'X - \beta'_S \alpha_S\| \leq c_1 \epsilon^2$. If (U_p, Σ_p, V_p) are the left singular vectors, non-zero singular values and right singular vectors of $Y'X$, i.e.

$$\begin{aligned} Y'X &= U_p \Sigma_p V_p', \\ (\beta'_S U'_S)(V_S \alpha_S) &= \beta'_S \alpha_S = \hat{U}_p \hat{\Sigma}_p \hat{V}_p'. \end{aligned} \quad (26)$$

where $(\hat{U}_p, \hat{\Sigma}_p, \hat{V}_p)$ are the left singular vectors, non-zero singular values, and right singular vectors of $\beta'_S \alpha_S$. From the eigenvalue perturbation theory, $\|V_p - \hat{V}_p\| \propto O(\epsilon^2)$, $\|U_p - \hat{U}_p\| \propto O(\epsilon^2)$, $\|\Sigma_p - \hat{\Sigma}_p\| \propto O(\epsilon^2)$. Thus,

$$\begin{aligned} V_p \Sigma_p^{-1/2} &= \hat{V}_p \hat{\Sigma}_p^{-1/2} + \Delta_2, \\ U_p \Sigma_p^{-1/2} &= \hat{U}_p \hat{\Sigma}_p^{-1/2} + \Delta_3, \end{aligned} \quad (27)$$

where $\|\Delta_2\|, \|\Delta_3\| \propto O(\epsilon^2)$. The POD bases are:

$$T_r = X V_p \Sigma_p^{-1/2}, T_l = \Sigma_p^{-1/2} U_p' Y'. \quad (28)$$

We have:

$$\begin{aligned} Y'AX &= (\beta'_S U'_S + \delta \beta'_D U'_D) A (V_S \alpha_S + V_D \delta \alpha_D) \\ &= \beta'_S \Lambda_S \alpha_S + \underbrace{\delta \beta'_D \Lambda_D \delta \alpha_D}_{\Delta_4} = \beta'_S \Lambda_S \alpha_S + \Delta_4, \end{aligned} \quad (29)$$

where $\|\Delta_4\| \propto O(\epsilon^2)$. The reduced order system using this set of POD basis is:

$$\tilde{A} = T_l' A T_r = (\Sigma_p^{-1/2} U_p') (Y'AX) (V_p \Sigma_p^{-1/2}). \quad (30)$$

Substitute (27) and (29) into (30), we have:

$$\begin{aligned} \tilde{A} &= T_l' A T_r \\ &= (\hat{\Sigma}_p^{-1/2} \hat{U}_p' + \Delta_3) (\beta'_S \Lambda_S \alpha_S + \Delta_4) (\hat{V}_p \hat{\Sigma}_p^{-1/2} + \Delta_2) \\ &= \underbrace{(\hat{\Sigma}_p^{-1/2} \hat{U}_p' \beta'_S)}_P \Lambda_S \underbrace{(\alpha_S \hat{V}_p \hat{\Sigma}_p^{-1/2})}_{\hat{P}} + \Delta_5 = \hat{A} + \Delta_5, \end{aligned} \quad (31)$$

where $\|\Delta_5\| \propto O(\epsilon^2)$. First, we need to show $P \hat{P} = I$.

$$P \hat{P} = \hat{\Sigma}_p^{-1/2} \underbrace{\hat{U}_p' \hat{U}_p}_I \hat{\Sigma}_p \underbrace{\hat{V}_p' \hat{V}_p}_I \hat{\Sigma}_p^{-1/2} = \hat{\Sigma}_p^{-1/2} \hat{\Sigma}_p \hat{\Sigma}_p^{-1/2} = I. \quad (32)$$

Since P and \hat{P} are square matrices, thus $\hat{P} = P^{-1}$, $\hat{A} = P \Lambda_S P^{-1}$. From (31),

$$\tilde{A} = \tilde{P} \Lambda_{ij} \tilde{P}^{-1} = \hat{A} + \Delta_5. \quad (33)$$

Using the eigenvalue perturbation theory, $\tilde{P} = P + \Delta_6$, where $\|\Delta_6\| \propto O(\epsilon^2)$, $\|\Lambda_{ij} - \Lambda_S\| \propto O(\epsilon^2)$. where Λ_S are the eigenvalues of the system matrix A . Now, we want to bound the errors between the right and left eigenvectors corresponding to the same eigenvalues.

$$\begin{aligned} \Psi_{ij} &= T_r \tilde{P} = X V_p \Sigma_p^{-1/2} (P + \Delta_6) \\ &= (V_S \alpha_S + V_D \delta \alpha_D) (\hat{V}_p \hat{\Sigma}_p^{-1/2} + \Delta_2) (P + \Delta_6) \\ &= (V_S \alpha_S + V_D \delta \alpha_D) (\hat{V}_p \hat{\Sigma}_p^{-1} \hat{U}_p' \beta'_S + \Delta_7) \\ &= V_S \underbrace{\alpha_S \hat{V}_p \hat{\Sigma}_p^{-1} \hat{U}_p' \beta'_S}_{P^{-1}P} + V_D \delta \alpha_D \hat{V}_p \hat{\Sigma}_p^{-1} \hat{U}_p' \beta'_S + \Delta_8 \\ &= V_S + V_D \delta \alpha_D \hat{V}_p \hat{\Sigma}_p^{-1} \hat{U}_p' \beta'_S + \Delta_8, \end{aligned} \quad (34)$$

where, $\|\Delta_7\|, \|\Delta_8\| \propto O(\epsilon^2)$.

Since $\|V_D \delta \alpha_D \hat{V}_p \hat{\Sigma}_p^{-1} \hat{U}_p' \beta'_S\| \propto O(\epsilon)$, then $\|\Psi_{ij} - V_S\| \propto O(\epsilon)$. Similarly, if we denote $\Phi'_{ij} = \tilde{P}^{-1} T_l$, then $\|\Phi_{ij} - U_S\| \propto O(\epsilon)$.

REFERENCES

- [1] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *AIAA Journal*, vol. 40, pp. 2323–2330, 2002.
- [2] C. W. Rowley, "Model reduction for fluids using balanced proper orthogonal decomposition," *International Journal of Bifurcation and Chaos*, vol. 15, pp. 997–1013, 2005.
- [3] B. C. Moore, "Principal component analysis in linear systems: Controllability, observability and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, No.1, pp. 17–32, 1981.
- [4] G. Berkooz, P. Holmes, and J. L. Lumley, "The proper orthogonal decomposition in the analysis of turbulent flows," *Annual review of fluid mechanics*, vol. 25, pp. 539–575, 1993.
- [5] S. Kung, "A new identification method and model reduction algorithm via singular value decomposition," *12th Asilomar Conference on Circuits, Systems and Computers*, pp. 705–714, Nov. 1978.
- [6] J.-N. Juang and R. S. Pappa, "An eigensystem realization algorithm for model parameter identification and model reduction," *Journal of Guidance, Control, and Dynamics*, vol. 8, No.5, pp. 620–627, 1985.
- [7] Z. Ma, S. Ahuja, and C. W. Rowley, "Reduced order models for control of fluids using the eigensystem realization algorithm," *Theoret. Comput. Fluid Dyn.*, vol. 25, pp. 233–247, 2011.
- [8] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010.
- [9] C. W. Rowley, I. Mezic, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of Fluid Mechanics*, vol. 641, pp. 115–127, 2009.
- [10] J. H. Tu and C. W. Rowley, "An improved algorithms for balanced pod through an analytic treatment of impulse response tails," *Journal of Computational Physics*, vol. 231, pp. 5317–5333, June, 2012.
- [11] P. Drineas *et al.*, "Fast monte carlo algorithms for matrices ii: Computing a low rank approximation to a matrix," *SIAM J. Computing*, vol. 36, pp. 158–183, 2006.
- [12] M. W. Mahoney, "Randomized algorithms for matrices and data," arXiv: 1104.5557, Tech. Rep., 2011.
- [13] N. Halko, P. Martinsson, and J. Tropp, "Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions," *SIAM Review*, vol. 53, pp. 217–288, 2011.
- [14] M. Vidyasagar, "Randomized algorithms for robust controller synthesis using statistical learning theory," *Automatica*, vol. 37, pp. 1515–1528, 2001.
- [15] G. Calafiore and M. Campi, "The scenario approach to robust control design," *IEEE Transactions on Automatic Control*, vol. 51, pp. 742–753, 2006.
- [16] D. Yu and S. Chakravorty, "A randomized iterative proper orthogonal decomposition technique with application to filtering of pdes," *Proceedings of American Control Conference*, pp. 4363–4368, 2012.
- [17] —, "An iterative proper orthogonal decomposition (i-pod) technique with application to the filtering of partial differential equations," *Journal of Astronautical Sciences*, vol. Special issue on J. N. Juang's 60th birthday, to appear, 2013.
- [18] T. Kato, *Perturbation Theory for Linear Operators*. New York: Springer-Verlag, 1995.