A Randomly Perturbed Iterative Proper Orthogonal Decomposition (RI-POD) Technique for Filtering Distributed Parameter Systems

D. Yu, S. Chakravorty

Abstract—In this paper, we consider the filtering of distributed parameter systems (DPS), i.e., systems governed by partial differential equations (PDE). We adopt a reduced order model (ROM) based strategy to solve the problem. We propose a randomly perturbed iterative version of the snapshot proper orthogonal decomposition (POD) technique, termed RI-POD, to construct ROMs for DPS that is capable of capturing their global behaviour. Further, the technique is entirely data based, and is applicable to forced as well as unforced systems. We apply the ROM generated using the RI-POD technique to construct reduced order Kalman filters to solve the DPS filtering problem. The methodology is tested on the 1-dimensional heat equation.

Keywords: Proper Ortrhogonal Decomposition (POD), Filtering/ Data Assimilation, Distributed Parameter Systems.

I. INTRODUCTION

In this paper, we are interested in the filtering/ data assimilation of distributed parameter sytems (DPS), in particular, systems that are governed by partial differential equations (PDE). We take a reduced order model (ROM) based approach to the problem. We propose a randomly perturbed iterative version of the snapshot proper orthogonal decomposition (POD) technique that allows us to form an ROM of the DPS of interest in terms of the eigenfunctions of the PDE operator. We then apply this ROM, along with the Kalman filtering technqiue, to form a reduced order filter for the DPS. The filter is constructed in an offline-online fashion where the expensive computations for the ROM construction is accomplished offline, while the online part consists of the reduced order Kalman filter which is much more computational tractable than the full problem. The technique is applied to the 1-dimensional heat equation.

The problem of estimating dynamic spatially distributed processes is typically solved using the Ensemble Kalman Filter (EnKF) and has been used extensively in the Geophysics litertaure [1], [2] and more recently, in Dynamic Data Driven Application Systems (DDDAS) and traffic flow problems [3]–[9]. The EnKF is a particle based Kalman filter that maintains an ensemble of possible realizations of the dynamic map. The Kalman prediction and measurement update steps are performed using ensemble operations instead of the traditional matrix operations. A primary issue with the EnKF

D. Yu is a Graduate Student Researcher, Department of Aerospace Engineering, Texas A&M University, College Station
S. Chakravorty is an Associate Professor of Aerospace Engineering, Texas A&M University, College Station

is the choice of the ensemble realizations and their number. This is almost always done in a heuristic fashion. Also, the prediction stage requires expensive forward simulations of the realizations using a solver which can take a significant amount of time. Thus, real time operation is an issue. In contrast, all the expensive computations for our ROM based technique, namely POD basis and ROM generation, are done offline and hence, real time operation is never an issue given the offline computations. Thus, we may think of our approach as a computationally tractable alternative to the EnKF algorithm. Historically, there has been a lot of theoretical research in the Control Systems community on the estimation and control of systems driven by PDEs [10]–[19]. In fact, it is well known that for linear PDEs, an infinite dimensional version of the Kalman filter exists which involves the solution of an operator Ricatti equation [20]. This can be very computationally intensive and may be unsuitable for online implementation. In contrast, the major computational complexity of the ROM based technqiue that we propose is offline and the online computations are essentially trivial thereby making the technique very suitable for online implementation.

As has been noted before, we take a ROM based approach to solving the problem of filtering in DPS. In particular, we rely on the so-called proper orthogonal decomposition (POD), more precisely, the snapshot POD technique, to construct ROMs for the DPS of interest. The POD has been used extensively in the Fluids community to produce ROMs of fluid physics phenomenon such as turbulence and fluid structure interaction [21]–[24]. There has also been work recently to produce so-called balanced POD models to better approximate outputs of interest through an amalgam of the snapshot POD and the balanced model reduction paradigm of control theory [25] to produce computationally efficient balanced POD models of the physical phenomenon of interest [26], [27]. More recently, there has been work on obtaining information regarding the eigenfunctions of a system based on the snapshot POD, called the dynamic mode decomposition (DMD) [28], [29]. However, a couple of issues are central to the construction of the snapshot POD technique: 1) at what times do we take snapshots of the process, and 2) the snapshot POD essentially provides a reduced basis approximation of the localized behaviour of a system, is there a constructive way to infer the global behaviour of a system from the snapshot POD? We propose a randomly

perturbed iterative approach to the snapshot POD (RI-POD) which iteratively allows us to sample the process of interest at various different time scales thereby answering question 1 above. Further, we show that this process allows us to theoretically reconstruct all the eigenfunctions of the original system using either data from experiments or from numerical simulations (similar to the DMD approach) thereby allowing us to infer global behaviour of the system. Moreover, to the best of the knowledge of the authors, except for some recent work in Fuid mechanics [30], the use of POD based ROMs for filtering DPS is relatively non-existent. We apply the RI-POD based ROMs to the filtering problem and illustrate our technique on the 1-dimensional heat and wave equations as well as the 1-dimensional traffic problem.

The paper is organized as follows. In section II, we introduce the DPS filtering problem. In section III, we present the offline RI-POD procedure used to construct the ROM of the DPS. In section IV, we outline the online portion of the reduced order Kalman filter constructed for the DPS filtering along with error bounds for the resulting approximations. Further,we also apply the offline-online procedure problem to the solution of 1-dimensional DPS filtering problem for the heat equation.

II. Preliminaries

In this work,we are interested in the filtering of distributed parameter systems, systems whose evolution is governed by a partial differential equation (PDE), given sparse measurements of the spatio-temporal field variable. Mathematically, we are interested in estimating the state of the field variable $X \in H$, for some suitable Hilbert space H. The state X is governed by the operator equation

$$\dot{X} = \mathcal{A}X + \mathcal{W}, \tag{1}$$

where $\mathcal{W}$ is a spatially distributed Gaussian white noise process perturbing the motion of the system. We assume that the boundary conditions for the PDE are known. We do not have access to measurement of the entire state, instead we only have access to measurements of the field at some sparse set of spatial locations in the domain of the process given by

$$Y(t_k, x_j) = \mathcal{C}X(t_k) + \mathcal{V}_k^{(j)}, \tag{2}$$

where $X(t_k)$ represents the state at the discrete time instant t_k, and $Y(t_k, x_j)$ represents a localized measurement of the state variable at the sparse set of locations given by x_j, $j = 1, \cdots m$, and $\mathcal{V}_k^{(j)}$ is a discrete time white noise process corrupting the measurements at the spatial location x_j. We assume that the the differential operator $\mathcal{A}$ is self adjoint with a compact resolvent, and thus, A has a discrete spectrum with a full set of eigenvectors that forms an orthonormal basis for the Hilbert space H. Further, we assume that the operator $\mathcal{A}$ generates a stable semigroup.

Given the above assumptions,we can discretize the PDE above using computational techniques such as finite elements(FE)/ finite difference (FD) to obtain a discretized ver-

sion of the operator equations in Euclidean space $\Re^N$ given by the following:

$$\dot{x} = Ax + w, \tag{3}$$
$$y_k = Cx_k + v_k, \tag{4}$$

where x is a discretized version of the state X that resides in the high dimensional Euclidean space $\Re^N$, while w and v_k are discretized versions of the white noise process corrupting the state and measurement equations respectively. Given that the operator $\mathcal{A}$ is self-adjoint, the discretized operator A is a symmetric matrix with a full set of eigenvectors and real eigenvalues whose eigenvectors form an orthonormal basis of $\Re^N$ and for suitable large N, are arbitrarily good approximations of the true eigenfunctions of the original operator $\mathcal{A}$. We shall assume throughout this paper that a fine enough discretization is given to us and thus, the behavior of the original system is captured sufficiently well by the discretized version of the system. Thus, in the rest of the paper, we only consider the discretized version of the problem. Given the above discretization, a naive approach to the solution of the filtering problems is to use a standard Kalman filter to solve the problem. However, due to the very high dimensionality of $\Re^N$, since N can easily run into millions of degrees of freedom (DOF) for a general finely discretized PDE, the Kalman filtering equations are computationally intractable for such high DOF systems. Thus, first we need to suitably reduce the order of the system before we can hope to apply Kalman filtering techniques to the above problem.

III. A Randomly Perturbed Iterative Approach to Proper Orthogonal Decomposition (RI-POD)

Consider the following linear system:

$$\dot{x} = Ax, \text{ given } x(0), \tag{5}$$

where $x \in \Re^N$, N is very large and A is a symmetric matrix. Recall that the above high dimensional linear system results from the discretization of a self adjoint linear operator with a compact resolvent.

A 1. *We assume that there is a unique null vector corresponding to A and that the matrix A is Hurwitz, i.e., the system is stable.*

Suppose that we choose some arbitrary initial condition $x(0)$ and take M snapshots of the system's trajectory at the time instants $t_1 < t_2 < \cdots < t_M$, where these snapshots need not be equi-spaced. Let us denote the data matrix of the stacked snapshots by X, i.e.,

$$X = [x_1, x_2, \cdots, x_M],$$

where $x_i = x(t_i)$. Suppose now that the number of snapshots is much smaller than the dimension of the system, i.e., $M << N$. Then, using the snapshot POD technique, we can get the POD basis T of the trajectory encoded in the snapshot ensemble X as follows:

$$T = XV_p\Sigma_p^{-1/2}, \tag{6}$$

where V_p and Σ_p are the eigenvector-eigenvalue pair corresponding to the correlation matrix $X'X$, i.e.,

$$(X'X)V_p = V_p\Sigma_p. \tag{7}$$

Note that the MxM eigenvalue problem to be solved for the POD eigenfunctions is much easier than the high dimensional NxN eigenvalue problem that needs to be solved if we were interested in solving for the eigenvalues and eigenvectors of A. Given the snapshot POD eigenfunctions, we can obtain a reduced order approximation of the system in Eq. 5 as follows:

$$\dot{\psi} = (T'AT)\psi \equiv \tilde{A}\psi, \tag{8}$$

where ψ represents the projection of the system state onto the POD eigenfunctions and $\tilde{A}$ represents the reduced order MxM system matrix.

Consider the reduced order system matrix $\tilde{A}$. We know that $\tilde{A}$ is symmetric and thus, has a full eigenvalue decomposition. Let (Λ, P) represent the eigenvalue-eigenvector pair for $\tilde{A}$, i.e.,

$$\tilde{A}P = P\Lambda. \tag{9}$$

Noting that $\tilde{A} = P\Lambda P'$, the ROM matrix $\tilde{A}$ transformed to the orthonormal co-ordinates specified by P, can be represented in the modal co-ordinates ϕ as:

$$\dot{\phi} = \Lambda\phi. \tag{10}$$

Thus it follows that

$$\Lambda = (P'T')A(TP), \tag{11}$$

where T is the POD eigenfunction matrix and P is the ROM eigenfunction matrix. Note that T is NxM and that P is MxM, and hence, TP is NxM. The above equation looks suspiciously like an eigen decomposition of the matrix A except that $M << N$ and thus, this is not necessary. Note that the transformation TP denotes the transformation from the original state space to the POD eigenfunction space to the ROM eigenfunction space. In the following, we relate the eigenvalues and eigenvectors of A to the diagonal form Λ and the transformation TP.

A 2. *Assume that "a" eigenvectors of the matrix A are active in the snapshot ensemble X, i.e.,*

$$x_i = \sum_{j=1}^{a} \alpha_j^i v_j,$$

where $a \leq M$ and without loss of generality, it is assumed that the active eigenvectors consist of the first "a" eigenvectors. This assumption essentially implies that the number of modes active within the snapshots is less than the number of snapshots in the ensemble.

The following result is then true.

Proposition 1. *The columns of the transformation TP are the eigenvectors of A with corresponding eigenvalues encoded in the diagonal matrix Λ, i.e.,*

$$A(TP) = \Lambda(TP). \tag{12}$$

Proof:
Recall that $T = XV_p\Sigma_p^{-1/2}$. We have

$$X = V\alpha = [v_1, v_2, \cdots v_a]\begin{bmatrix} \alpha_1^1 & .. & \alpha_1^M \\ . & .. & . \\ \alpha_a^1 & .. & \alpha_a^M \end{bmatrix},$$

where note that V is an Nxa and α is an axM matrix. For simplicity, let $a = M$. Then, it follows that

$$\tilde{A} = T'AT = \Sigma_p^{-1/2}V_p'X'AXV_p\Sigma_p^{-1/2},$$
$$= \underbrace{(\Sigma_p^{-1/2}V_p'\alpha')}_{MxM}\underbrace{(V'AV)}_{MxM}\underbrace{(\alpha V_p\Sigma_p^{-1/2})}_{MxM}.$$

Comparing this to the fact that $\tilde{A} = P\Lambda P'$, it follows that

$$P = \Sigma_p^{-1/2}V_p'\alpha', \tag{13}$$

if $(\alpha V_p\Sigma_p^{-1/2})(\Sigma_p^{-1/2}V_p'\alpha') = I$. To show this, note that

$$V_p\Sigma_p^{-1}V_p' = (X'X)^{-1} = (\alpha'V'V\alpha)^{-1} = \alpha^{-1}(\alpha')^{-1}. \tag{14}$$

The inverse of α exists because its full rank (rank M). The above implies that

$$(\alpha V_p\Sigma_p^{-1/2})(\Sigma_p^{-1/2}V_p'\alpha') = \alpha\alpha^{-1}(\alpha')^{-1}(\alpha') = I. \tag{15}$$

Hence, it follows that

$$TP = (XV_p\Sigma_p^{-1/2})(\Sigma_p^{-1/2}V_p'\alpha'),$$
$$= V(\alpha V_p\Sigma_p^{-1/2})(\Sigma_p^{-1/2}V_p'\alpha') = V, \tag{16}$$

i.e., the columns of TP are indeed eigenvectors of A. Moreover, it also follows that owing to the uniqueness of the similarity transformation of $\tilde{A}$ that the eigenvalues corresponding to the eigenvectors in TP are in the diagonal form Λ. Hence, this proves our assertion for the case when $a = M$.

When $a < M$, the rank of the snapshot ensemble X is $a < M$ and hence, the rank of $X'X$ is $a < M$. Thus, it follows that the POD eigenvalues will be non-zero for only a POD eigenfunctions. Therefore, the transformation into the POD basis $T = XV_p\Sigma_p^{-1/2}$ should only include only the POD eigenvectors corresponding to the non-zero eigenvalues. Once this is understood, then we can see that TP is going to be axa and we are back to the situation considered previously. In this fashion, the assertion is also proved for the case when $a < M$ and thereby, completes the proof of the proposition. ∎

At this point, we make the following remark.

Remark 1. *Suppose that $a > M$, i.e., the number of active eigenvectors are more than the number of snapshots. WLOG, let $a = N$. Then*

$$\tilde{A} = T'AT = \Sigma_p^{-1/2}V_p'X'AXV_p\Sigma_p^{-1/2}$$
$$= (\Sigma_p^{-1/2}V_p'\alpha')V'AV(\alpha V_p\Sigma_p^{-1/2})$$
$$= \underbrace{(\Sigma_p^{-1/2}V_p'\alpha')}_{MxN}\underbrace{\Lambda}_{NxN}\underbrace{(\alpha V_p\Sigma_p^{-1/2})}_{NxM}$$
$$= \underbrace{\beta'}_{MxM}\underbrace{\Gamma}_{MxM}\underbrace{\beta}_{MxM},$$

where (β, Γ) represents the eigenvalue decomposition of the

ROM matrix $\tilde{A}$. Note that now owing to the fact that $N > M$, we can no longer use the uniqueness of the similarity transformation of $\tilde{A}$ to conclude that the transformation $T\beta$ contains the eigenvectors of A. In fact, some of them might be the same as the eigenvectors of A, however, it is not necessary. In particular, theoretically, we cannot conclude anything regarding the relationship of $T\beta$ to the eigenfunctions of A.

The above proposition and the remark above suggest a technique through which eigenvectors of the system matriX A can be extracted upto any time-scale. First,we make the following assumption.

A 3. *We assume that there are K characteristic timescales embedded in the matrix A, namely $T_1, \cdots T_K$. Let the eigenvalues corresponding to timescale T_j be $\{\lambda_1^{(j)}, \cdots \lambda_{M_j}^{(j)}\}$ and let the corresponding eigenvectors be $[v_1^{(j)}, \cdots v_{M_j}^{(j)}] \equiv V^{(j)}$. Further, we assume that the timescales are well-separated, i.e., if for some t, $e^{\lambda_k^{(j)}t} \neq 0$, then $e^{\lambda_k^{(i)}t} \approx 0$ for all $i < j$. The above assumption essentially implies all the eigenvectors corresponding to timescales below a given timescale decay well before the eigenvectors at the given timescale decay.*

At this point, we also need to make sure that all possible eigenfunctions corresponding to any timescale are excited. The following result assures us of this:

Proposition 2. *Let the initial condition to the linear system in Eq. 5 be chosen according to a Gaussian distribution $N(0, \sigma^2 I)$. Let the jth such trajectory be denoted by $X^{(j)}$. Then, every eigenfunction of A is excited almost surely, i.e., given any eigenfunction, there is atleast one trajectory $X^{(j)}$ such that the eigenfunction is active within the ensemble as $j \to \infty$.*

Proof: Due to the eigenvalue decomposition of A, we may write:

$$x(t) = \sum_{i=1}^{N} e^{\lambda_i t}(x(0), v_i)v_i,$$

where $(.,.)$ denotes the inner product in $\Re^N$. The above implies that $(x(t), v_i) = e^{\lambda_i t}(x(0), v_i)$, and hence

$$E|(x(t), v_i)|^2 = e^{2\lambda_i t}E|(x(0), v_i)|^2$$
$$= e^{2\lambda_i t}v_i' R_0 v_i = \sigma^2 e^{2\lambda_i t}. \tag{17}$$

Thus, the ith component of the system trajectory, i.e., the contribution of the ith eigenvector, is a Gaussian random variable with zero-mean and a variance that exponentially decays in time as shown above. Thus, the ith mode is bound to be active for atleast one among the ensemble of trajectories. In fact, owing to the Gaussian nature of the component, it is true that its absolute value will be above any given threshold, at any given time, with a finite probability. ∎

Given the results above and assumption **A3**, we are in a position to outline a procedure that allows us to isolate all eigenfunctions at any given timescale.

Suppose without loss of generality that $T_1 > T_2 \cdots > T_K$. Suppose now that we are interested in isolating all the eigenfunctions corresponding to the timescale T_1. We choose an initial time $t_0^{(1)}$ and subsequent snapshot times $t_n^{(1)}$, $n = 1 : M$, such that $M > M_1$ and such that the initial time $t_0^{(1)} >> T_2$. Thus, the snapshot timing assures us that all the eigenvectors at the timescales below T_1 will have decayed by the snapshot times of interest, and thus, the only participating modes are the eigenfunctions corresponding to timescale T_1. Then, using Propositions 1 and 2, we know that we can isolate all the eigenfunctions at the timescale T_1 given enough snapshot ensembles. In particular, suppose that $X_j^{(1)}$ is the jth snapshot ensemble at timescale T_1. Due to proposition 2, as $j \to \infty$, we know that every eigenfunction in set $V^{(1)}$ is bound to be excited. Further, due to the fact that $M > M_1$, it follows using Proposition 1 that the eigenfunctions of the ROM are the same as the eigenfunctions of A. Thus, every snapshot ensemble gives us some of the eigenvectors $v \in V^{(1)}$ and as $j \to \infty$, we are assured that all possible $v \in V^{(1)}$ are recovered.

Given that we have recovered all the eigenfunctions $V^{(1)}$ corresponding to the longest timescale T_1, we can now iteratively recover all the eigenfunctions at all the subsequent timescales as follows. Given $V^{(1)}$, we randomly choose an initial condition $x(0)$ and form the snapshot ensemble X at snapshot times $t_0^{(2)}, \cdots t_M^{(2)}$, such that number of snapshots $M > M_2$, and the initial time of the snapshot $t_0^{(2)} >> T_3$, i.e., such that all eigenfunctions at timescales shorter than T_2 are absent in the ensemble. Given the snapshot ensemble X we "clean" the snapshots by subtracting the contributions of the eigenfunctions from $V^{(1)}$, i.e.,

$$\tilde{x}(t_j^{(2)}) = x(t_j^{(2)}) - \sum_{k=1}^{M_1} e^{\lambda_k^{(1)}t_j^{(2)}}(x_0, v_k^{(1)})v_k^{(1)}. \tag{18}$$

Consider the "cleaned" snapshot ensemble $\tilde{X}$ consisting of the cleaned snapshots from above. It follows that $\tilde{X}$, by construction, only contains eigenfunctions from the set $V^{(2)}$ and thus, following the randomly perturbed POD procedure outline previously, we can recover all the eigenfunctions in $V^{(2)}$. Given $V^{(1)}$ and $V^{(2)}$, we can repeat the cleaning, and randomly perturbed POD procedure, to recursively obtain all the sets $V^{(n)}$ upto any desired timescale T_n. The above development can be summarized in the following algorithm:

The development above and the RI-POD algorithm can be summarized in the following result.

Proposition 3. *Under assumptions A1-A3, the RI-POD algorithm can extract all eigenfunctions $V^{(i)}$ corresponding to any given time scale $T^{(i)}$.*

Remark 2. *The timescales $T_1, \cdots T_k$ are dependent on the Physics and can be inferred from physical insight or simulations. The number of snapshots that are required to extract the eigenfunctions have to be "large enough". Of course, it might not be possible to know a priori when M is large enough. However, some amount of trial and error can tell us as to what is a suitable number for M. In fact, a good heuristic*

Algorithm 1 Algorithm RI-POD
---
1) Given timescales $T_1, \cdots T_K$
2) Set $i = 1$, $V^{(0)} = \phi$
3) WHILE $i \leq K$
   DO
      a) Choose snapshot times $t_0^{(i)}, \cdots t_M^{(i)}$, such that $t_0^{(i)} >> T_{i+1}$ and $M > M_i$
      b) Set $j = 1$
         i) Choose $x_{0,j}^{(i)}$, the initial condition of the jth snapshot ensemble at time scale T_i, from $N(0, \sigma^2 I)$ and generate the jth snapshot ensemble $X_j^{(i)}$
         ii) Clean all the slower eigenfunctions from the snapshot ensemble using Eq. 18, and the previously extracted eigenfunctions from the sets $V^{(1)}, V^{(2)}, \cdots V^{(i-1)}$
         iii) Isolate the eigenfunctions at timescale T_i, $V^{(i)}$, using the snapshot POD. Set $j = j + 1$
         iv) If all eigenfunctions in $V^{(i)}$ have been obtained, go to step (c), else go to step (i)
      c) Set $i = i + 1$
4) Output the eigenfunctions in sets $V^{(1)}, \cdots V^{(K)}$
---

measure is to increase the initial time of the snapshots till we have lesser number of modes participating than the number of snapshots. This can easily be construed from the eigenvalue decomposition of the snapshot ensemble by checking for zero eigenvalues and eigenvectors.

The RI-POD technique is a completely data based technique and does not need knowledge of the system matrix A. Note that ultimately, the ROM $\tilde{A} = T'AT$, contains all the information regarding the eigenfunctions of the operator A under the assumptions above. Again note that $T = XV_p\Sigma_p^{-1/2}$, and thus, it follows that $\tilde{A} = T'AXV_p\Sigma_p^{-1/2} = T'\tilde{X}V_p\Sigma_p^{-1/2}$, where $\tilde{X}$ is the one time step advanced version of the snapshot ensemble X (in the discrete time case), and can be obtained directly from simulation or experimental data. In the continuous time case, $\tilde{X}$ may be obtained as follows: advance the snapshots by a very short time δT, to obtain $\delta X = X' - X$, where X' is the short time advanced ensemble, and then obtain $\tilde{X} = AX \approx \frac{\delta X}{\delta T}$. Hence, the RI-POD technique is truly data based (this is similar to the DMD technique [28]).

It should also be noted that Proposition 1 does not distinguish between forced systems and unforced systems since Assumption 2 under which the result is valid only asks for certain sufficient conditions on the active eigenfunctions of the system in the snapshot ensemble. Since the forced response of a linear system is also expressed in terms of the eigenfunctions, the RI-POD procedure is valid for forced systems as well as long as Assumption 2 is valid. Hence, the procedure can be used on experimental data, where the system response may be forced. Of course, the issue is that Assumption 2 underlying Proposition 1 may not be satisfied for forced

systems. However, in our experiments we do see that this assumption is indeed satisfied and that we can actually extract the eigenfunctions of the forced system using the RI-POD procedure. Representative results from our experiments are shown in Figures 1-3. In Figs. 1 and 2, we compare the actual eigenvalues of a randomly generated 100 x 100 system with those obtained by the RI-POD procedure, for an unforced as well as a forced (constant forcing) system. The results show that the RI-POD eigenvalues agree very well with the actual eigenvalues. In Fig.3, we show the comparison for a white noise forced 20 x 20 system (a discrete time case). Again, this case also shows very good agreement between the RI-POD eigenvalues and the actual system eigenvalues.
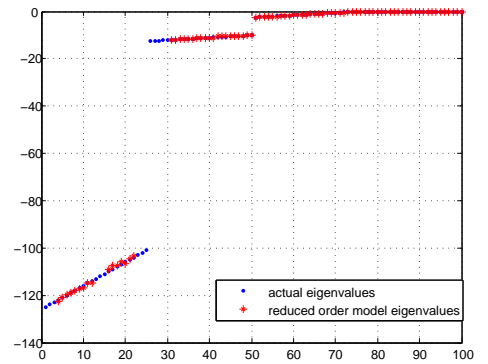

Fig. 1. Comparison of actual eigenvalues with those obtained using RI-POD for an unforced 100 x 100 system

IV. APPLICATION OF RI-POD TO FILTERING OF DISTRIBUTED PARAMETER SYSTEMS

Consider now the continuous -discrete filtering of the distributed parameter system in the high dimensional discretized setting:

$$\dot{x} = Ax + w, \tag{19}$$
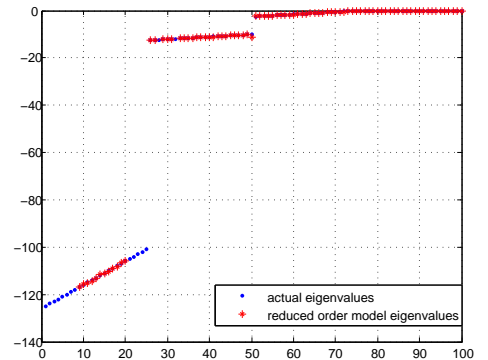
$$y(t_k) = Cx(t_k) + v(t_k), \tag{20}$$

Fig. 2. Comparison of actual eigenvalues with those obtained using RI-POD for a forced 100 x 100 system

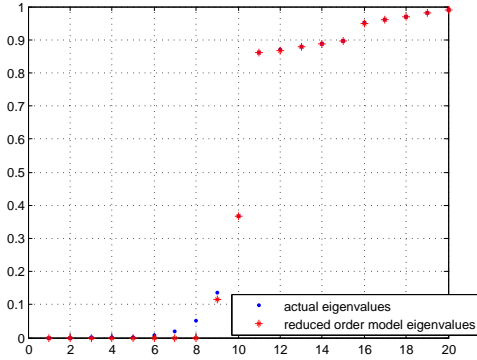Fig. 3. Comparison of actual eigenvalues with those obtained using RI-POD for a white noise forced 20 x 20 system

where recall that $y(t_k) \in \Re^p$ is the measurement at time t_k, w is a continuous white noise process perturbing the systems while v_k is a measurement white noise process corrupting the measurement at time t_k. Typically, the measurements are very sparse, i.e., $p << N$, and N is very large. Hence, using standard estimation theoretic techniques such as the Kalman filter for filtering the above system is out of question owing to the high dimensionality of the resulting problem (the Kalman filter requires $O(N^2)$ operations at every update step. Thus, it is vitally important that suitable ROMs be devised to alleviate the computational intractability of the problem above. Since we are considering the discrete setting for filtering, let us assume that the measurements are taken time T apart.

In order to form a suitable ROM of the above system, suppose that we keep only N_r of the eigenfunctions of A as modes of the ROM. The expected value of the error between the true system and the ROM at any time is given by the following result.

Proposition 4. *The expected value of the squared error in keeping only N_r modes in the ROM is given by*

$$E||e(t)||^2 = \sum_{i=N_r+1}^{N} e^{2\lambda_i t} E|(x(0), v_i)|^2 +$$
$$\sum_{i=N_r+1}^{N} \sigma_i^2 \left(\frac{e^{2\lambda_i t} - 1}{2\lambda_i} \right),$$
$$\sigma_i^2 = v_i' R_W v_i, \quad (21)$$

where R_w represents the covariance of the white noise process w. The first term in the above expression is due to the initial conditions while the second term is due to the random perturbation w.

Proof:
The error incurred in keeping only N_r modes in the ROM is given by:

$$e(t) = \sum_{i=N_r+1}^{N} e^{\lambda_i t}(x(0), v_i)v_i +$$

$$\sum_{i=N_r+1}^{N} \Delta_i^w(t)v_i,$$

where

$$\Delta_i^w(t) = \int_0^t e^{\lambda_i(t-\tau)} c_i^w(\tau) d\tau,$$
$$c_i^w(\tau) = (w(\tau), v_i).$$

Then , it follows that:

$$E||e(t)||^2 = \sum_i e^{2\lambda_i t} E|(x(0), v_i)|^2 + \sum_i E|\Delta_i^w(t)|^2, \quad (22)$$

where for notational ease the subscript i is used to denote the summation from $N_r + 1$ to N. Then,

$$|\Delta_i^w(t)|^2 = \int_0^t \int_0^t e^{\lambda_i(t-\tau)} e^{\lambda_i(t-s)} c_i^w(\tau) c_i^w(s) d\tau ds. \quad (23)$$

Noting that

$$c_i^w(\tau) c_i^w(s) = (w(\tau), v_i)(w(s), v_i) = \sum_{j,k} w_j(\tau) w_k(s) v_{ij} v_{ik},$$

where v_{ij} denotes the jth component of v_i, it follows that

$$E[c_i^w(\tau) c_i^w(s)] = \sum_{j,k} v_{ij} v_{ik} E[w_j(\tau) w_k(s)]$$
$$= v_i' R_w \delta(t-s) v_i, \quad (24)$$

where $\delta(.)$ denotes the Dirac delta function. Then, substituting the above equation back into Eq. 23, and using the result in Eq. 22, while using the sampling property of the dirac delta function under an integral, the result follows. ∎

Thus,we have:

$$E||e(t)||^2 \leq \underbrace{\sum_{i=N_r+1}^{N} e^{2\lambda_i t} E|(x(0), v_i)|^2}_{\text{initial condition error}} +$$
$$\underbrace{\sum_{i=N_r+1}^{N} \sigma_i^2 \left(\frac{-1}{2\lambda_i} \right)}_{\text{random perturbation error}}. \quad (25)$$

The above expression gives an estimate of the error made in keeping only N_r modes in the solution. Note that the measurement equations are immaterial in these error estimates since they do not alter the system equations in any fashion. In fact, the above is an a priori estimate that is averaged over all possible future observations.

Given the measurement time interval T, and the probability density function of the initial state x_0, we can neglect those modes such that $e^{2\lambda_i T} \approx 0$ and thus, the perturbation due to the initial condition is negligible. Of course, the error due to the stochastic perturbations remains, however, theoretically, we can get all λ_i since that is assured us by Proposition 3 and RI-POD, and hence, we can make an a priori error estimate regarding the error made in keeping only N_r modes of the system.

Given that we have sufficient number of modes in our ROM such that the error in the reduced solution is within some pre-specified bounds, the filtering of the DPS system proceeds as follows. We choose those eigenfunctions that need to be kept, given that they have already been extracted using the RI-POD procedure, and form the ROM for the filtering problem as follows. Define the transform $V_r = [v_1, \cdots v_{N_r}]$ consisting of the retained eigenmodes. The filtering ROM is the following:

$$\dot{\psi} = (V_r'AV_r)\psi + V_r'w, \ \psi_i(0) = (x(0), v_i),$$
$$y(t_k) = (CV_r)\psi + v_k. \tag{26}$$

In the above equation ψ_i represents the ith component of the ROM state ψ. The above system now results in an $N_r \text{x} N_r$ filtering problem with $N_r << N$ and thus, standard estimation theoretic methods such as the Kalman filter can be used to solve the problem. In the following sections, we show the application of the RI-POD based filtering technique to the 1-D Heat equation.

A. Heat Equation

The heat transfer by conduction along a slab is given by the partial differential equation:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} \tag{27}$$

The length of the slab is $1m$ and the continuous spatial domain X is divided into 200 grid cells of equal length. The model is simulated for a period of 10 seconds and the time horizon is discretized into 500 time steps.

The ROM for the heat equation from Section III is given by:

$$\dot{\psi} = (V_r'AV_r)\psi + V_r'w, \ \psi_i(0) = <x(0), v_i>,$$
$$y(t_k) = (CV_r)\psi + v_k. \tag{28}$$

Using RI-POD, the measurements can be collected from a random initial condition. The slab is assumed to be heated to 200 C at its left end and the temperature maintained there, for generating the POD model.

The temperature of the slab is measured at five equispaced points along its length. In the simulation case shown, the left end is actually heated to 300 C. The comparisons between the reduced order filter and full order filter at five different time steps are shown in Figure 4. The red curves are filter results from original model and the blue curves are filter results from the reduced model. In Figure 5, the error, and the 3σ boundary for the error, for the reduced model, at a randomly chosen location is shown.

It can be seen that the RI-POD ROM based Kalman filter provides a good estimate of the temperature profile for the problem.
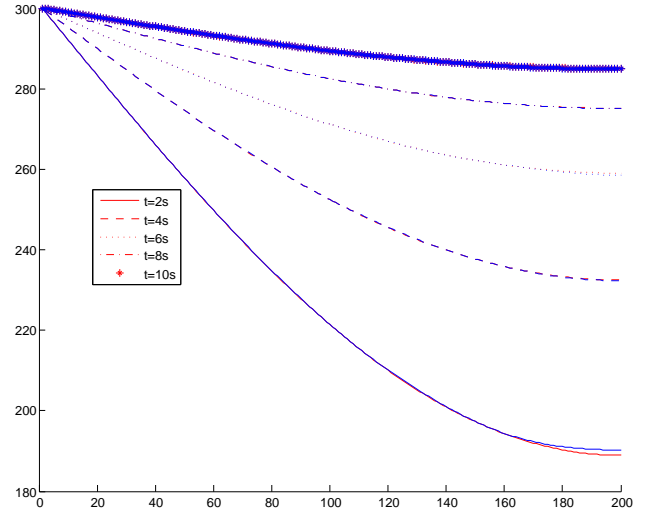

Fig. 4. Comparison of ROM filter with actual temperature profile

V. CONCLUSION

In this paper, we have introduced a randomly perturbed iterative snapshot POD (RI-POD) based approach to form global ROMs for compact self-adjoint linear systems. We have used the RI-POD based ROM to form a reduced order Kalman filter for application to the filtering of linear distributed parameter systems. We have shown the application of the RI-POD ROM based filtering technique to the heat equation. In the future, we will apply our technique to more realistic 2 and 3-dimensional distributed parameter systems that may be encountered in practice such as chemical plumes and wildfires, and more ambitiously, 3-D fluid flow problems. Finally, the RI-POD technique is suitable for symmetric systems, however, there does not seem to be a straightforward way to extend it to non-symmetric systems, and thus, this extension could be another possible avenue for future research.

REFERENCES

[1] L. Bertino, G. Evensen, and H. Wackernagal, "Combining geostatistics and kalman filtering for data assimiltaion in an estuarine system," *Inverse MEthods*, vol. 18, pp. 1–23, 2002.

[2] P. L. Houtekamer and H. L. Mitchell, "A sequential ensemble filter for atmosheric data assimilation," *Mon. weath. Rev.*, vol. 129, pp. 123–137, 1999.

[3] G. Evensen, "Incerse methods and data assimilation in nonlinear ocean models," *Physica D*, vol. 77, pp. 108–129, 1994.

[4] ——, "Advanced data assimilation for strongly nonlinear dynamics," *Mon. weath. Rev.*, vol. 125, pp. 1342–1354, 1997.

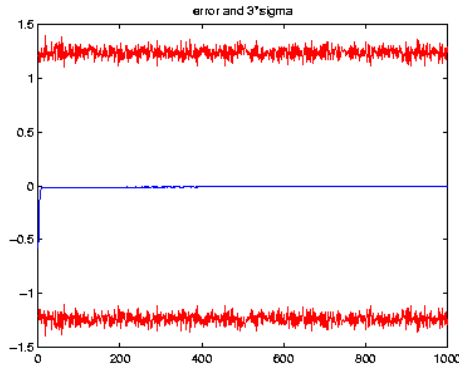[5] G. B. et. al., "Analysis scheme in th ensemble kalman filter," *Mon. Weath. Rev.*, vol. 126, pp. 1719–1724, 1998.

[23] K. C. Hall *et al.*, "Proper orthogonal decomposition technque for transonic unsteady aerodynamic flow," *AIAA Journal*, vol. 38, pp. 1853–1862, 2000.

[24] L. Sirovich, "Turbulence and dynamics of coherent structures. part 1: Coherent structures," *Quarterly of Applied Mathematics*, vol. 45, pp. 561–571, 1987.

[25] B. C. Moore, "Principal component analysis in linear systems: Controllability, observability and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, pp. 17–32, 1981.

[26] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *AIAA Journal*, vol. 40, pp. 2323–2330, 2002.

[27] C. W. Rowley, "Model reduction of fluids using balnced proper orthogonal decomposition," *International Journal fo Bifurcation and Chaos*, vol. 15, pp. 997–1013, 2005.

[28] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid EMchanics*, vol. 656, pp. 5–28, 2010.

[29] C. W. Rowley *et al.*, "Spectral analysis of nonlinear analysis," *Journal of Fluid Mechanics*, vol. 641, pp. 115–127, 2009.

[30] S. Ahuja and C. W. Rowley, "Feedback control of unstable steady states of flow past a flat plate using reduced-order estimators," *Journal of Fluid Mechanics*, vol. 645, pp. 447–478, 2010.

Fig. 5.   Error and 3σ error bounds for a randomly chosen location

[6] G. Evensen and P. J. V. Leeuwen, "An ensemble kalman smoother for nonlinear dynamics," *Mon. Weather Rev.*, vol. 128, pp. 1852–1867, 2000.

[7] G. Evensen, "The ensemble kalman filter: theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, pp. 343–367, 2003.

[8] J.Mandel *et al.*, "A dynamic data driven wildland fire model," in *Springer Leacture Notes on Computer science (LNCS), 4487*, 2007.

[9] D. B. Work *et al.*, "An ensemble kalman filtering approach to highway traffic estimation using gps enabled mobile devices," in *Proc. IEEE Int. Conf. Dec. Cont.*, 2008.

[10] A.Bensoussan *et al.*, *Representation and Control of Infinite Dimensional Systems, Vols. I and II.*   Birkhauser, 1992.

[11] A. Y. Khapalov, "Optimal measurement trajectories for distributed parameter systems," *Systems and Control Letters*, vol. 18, pp. 467–477, 1992.

[12] ——, "Observability of hyperbolic systems with interior moving sensors," *Lecture notes in Control and Info. Sci.*, vol. 185, pp. 489–499, 1993.

[13] ——, "l^∞ exact observability of the heat equation with scanning pointwise sensor," *SIAM J. Control and Optim.*, pp. 1037–1051, 1994.

[14] J. A. burns *et al.*, "A distributed parameter control approach to optimal filtering and smoothing with mobile sensor networks," in *Proc. Mediterrenean Control Conf.*, 2009.

[15] J. A. Burns and B. King, "Optimal sensor location for robust control of distributed parameter systems," in *Proc. IEEE Conf. Dec. Control*, 1993, pp. 3697–3972.

[16] A. E. Jai and A. J. Pritchard, *Sensors and Actuators in Analysis of Distributed Parameter Systems.*   New York, NY: John Wiley, 1988.

[17] Y. Sakawa, "Observability and related problems for partial differential equations of parabolic type," *SIAM J. Control*, vol. 13, pp. 14–27, 1975.

[18] M. Krstic and A. Smyshlaev, *Boundary Controls of PDEs.*   Philadelphia, PA: Advances in Design and Control: SIAM, 2008.

[19] R. Vasquez and M. Krstic, *Control of Turbulent and Magnetohydrodynamic Flows.*   Boston, MA: Systems and Control: Foundations and Applications, Birkhauser, 2007.

[20] H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems.*   Boston, MA: Systems and Control: Foundations and applications, Birkhauser, 1989.

[21] M. Loeve, *Probability Theory.*   New York, NY: Van Nostrand, 1955.

[22] G. Berkooz *et al.*, "The proper orthogonal decomposition in the analysis of turbulent flows," *Ann. Rev.. Fl. Mech.*, vol. 25, pp. 539–575, 1993.