An autoregressive (AR) model based stochastic unknown input realization and filtering technique

Dan Yu, Suman Chakravorty

Abstract—This paper studies the state estimation problem of linear discrete-time systems with stochastic unknown inputs. The unknown input is a wide-sense stationary process while no other prior information needs to be known. We propose an autoregressive (AR) model based unknown input realization technique which allows us to recover the input statistics from the output data by solving an appropriate least squares problem, then fit an AR model to the recovered input statistics and construct an innovations model of the unknown inputs using the eigensystem realization algorithm (ERA). An augmented state system is constructed and the standard Kalman filter is applied for state estimation. A reduced order model (ROM) filter is also introduced to reduce the computational cost of the Kalman filter. Two numerical examples are given to illustrate the procedure.

I. INTRODUCTION

In this paper, we consider the state estimation problem for systems with unknown stochastic inputs. The main contribution of our work is that when no prior information of the unknown inputs is known, we recover the statistics of the unknown inputs from the measurements, and then construct an innovations model of the unknown inputs from the recovered statistics such that the standard Kalman filter can be applied for state estimation. The innovations model is constructed by fitting an autoregressive (AR) model to the recovered input correlation data from which a state space model is constructed using the balanced realization technique. The method is tested on stochastically perturbed heat and laminar flow problems.

The problem of state estimation of systems with unknown inputs has received considerable attention over the past few decades. The unknown input observer (UIO) has been well established for deterministic systems [1]–[3]. Various methods of building full-order or reduced-order observers have been developed, such as [4]–[6]. Recently, sliding mode observers have been proposed for systems with unknown inputs [7]. The design parameters and matrices need to be well chosen to satisfy certain conditions in order for the observers to perform well. For systems without the “observer matching” condition being satisfied, a high-gain approach is proposed [8]. The high-gain observers are used as approximate differentiators to obtain the estimates of the auxiliary outputs. In the presence of measurement noise, the high-gain observer amplifies the noise, and extra care needs to be taken when designing the gain matrix.

For stochastic systems, the problem of state estimation is known as unknown input filtering (UIF), and many UIF approaches are based on the Kalman filter [9]–[11]. When the dynamics of the unknown inputs is available, for example, if it can be assumed to be a wide-sense stationary process with known mean and covariance, one common approach called Augmented State Kalman Filter (ASKF) is used, where the states are augmented with the unknown inputs [12]. To reduce the computational complexity of ASKF, optimal two-stage and three-stage Kalman filters have been developed to decouple the augmented filter into two parallel reduced-order filters by applying a U-V transformation [13]–[15]. When no prior information about the unknown input is available, an unbiased minimum-variance (UMV) filtering technique has been developed [16], [17]. The problem is transformed into finding a gain matrix such that the trace of the estimation error matrix is minimized. Certain algebraic constraints must be satisfied for the unbiased estimator to exist. In both the approaches above, the process noise is assumed to be white noise with known covariance.

In practice, there are many applications where the unknown inputs can be modeled as a stochastic process. For example, the state estimation of perturbed laminar flows is considered in [18]. It shows that the external disturbances (as well as the sensor noise and initial conditions) can be modeled as unknown stochastic inputs which perturb the linearized Navier-Stoke equations. Thus, the state estimation problem of such system is transformed into the unknown input filtering problem with stochastic unknown inputs. Also, our work can be applied to identify the statistics of colored process noise. There is some research that considers the Kalman filtering with unknown noise covariances [19], [20]. The process noise is assumed to be white noise with unknown covariance, while in our approach, the process noise can be colored in time as well. There are also applications of our technique in signal processing, such as the wideband power spectrum estimation [21], where the problem is to recover the unknown power spectrum of a wide-sense stationary signal from the obtained sub-Nyquist rate samples.

In this paper, we address the state estimation problem of systems with stochastic unknown inputs. The unknown inputs are assumed to be wide sense stationary, while no other information about the unknown inputs is known. We propose a new unknown input filtering approach based on system realization techniques. Instead of constructing the gain matrix which needs to satisfy certain constraints, we apply the standard Kalman filtering using the following procedure: 1) recover the statistics of the unknown inputs from...
the measurements by solving an appropriate least squares problem, 2) find a spectral factorization of unknown input process by fitting an autoregressive (AR) model, 3) construct an innovations model of the unknown inputs via the eigen-system realization algorithm (ERA) [22] to the recovered input correlation data, and 4) apply the Augmented State Kalman Filter for state estimation. Different from existing methods, we construct a stochastic unknown input model from sensor data, which can be colored in time. To reduce the computational cost of the ASKF, we apply the Balanced Proper Orthogonal Decomposition (BPOD) technique [23] to construct a reduced order model (ROM) for filtering.

The paper is organized as follows. In Section II, the problem is formulated, and general assumptions are made about the system and the unknown inputs. In Section III, the AR based unknown input realization approach is proposed. The unknown input statistics are recovered from the measurements, then a linear model is constructed using an AR model and the ERA is used to generate a balanced minimal realization of the unknown inputs. After an innovations model of the unknown inputs is constructed, the ASKF is applied for state estimation in Section IV. Also, a ROM constructed using the BPOD is introduced to reduce the computational cost of Kalman filter. Section V presents two numerical examples that utilize the proposed technique.

II. PROBLEM FORMULATION

Consider a complex valued linear time-invariant discrete time system:

\[ x_k = Ax_{k-1} + Bu_{k-1}, \]
\[ y_k = Cx_k + v_k, \]

where \( x_k \in \mathbb{C}^n, y_k \in \mathbb{C}^q, v_k \in \mathbb{C}^q, u_k \in \mathbb{C}^p \) are the state vector, the measurement vector, the measurement white noise with known covariance, and the unknown stochastic inputs respectively. The process \( u_k \) is used to model the presence of the external disturbances, process noise, and unmodelled terms. Here, \( A \in \mathbb{C}^{n \times n}, B \in \mathbb{C}^{n \times p}, C \in \mathbb{C}^{q \times n} \) are known.

The following assumptions are made about the system (1):

- A1. \( A \) is a stable matrix.
- A2. \( \text{rank}(B) = p, \text{rank}(C) = q, \text{rank}(CB) = \text{rank}(B) \) which implies that \( p \leq q \).
- A3. \( u_k \) and \( v_k \) are uncorrelated.
- A4. We further assume that the unknown input \( u_k \) is generated by a linear stochastic system:

\[ \xi_k = A_c \xi_{k-1} + B_c v_{k-1}, \]
\[ u_k = C_c \xi_k + \mu_k, \]

where \( \nu_k, \mu_k \) are uncorrelated white noise processes.

Remark 1: A2 is a general assumption in unknown input observer/filtering, the so-called “observer matching” condition.

A4 implies that \( u_k \) is a wide-sense stationary (WSS) process with a rational power spectrum.

In this paper, we consider the state estimation problem when the system (2), i.e., \( (A_c, B_c, C_c) \) are unknown. Given the output data \( y_k \), we want to construct an innovations model for the unknown stochastic input \( u_k \), such that the output statistics of the innovations model and system (2) are the same. Given such a realization of the unknown input, we apply the standard Kalman filter for state estimation, augmented with the unknown input states.

III. AR BASED UNKNOWN INPUT REALIZATION TECHNIQUE

In this section, we propose an AR based unknown input realization technique which can construct an innovations model of the unknown inputs such that the ASKF can be applied for state estimation. First, a least squares problem is formulated based on the relationship between the inputs and outputs to recover the statistics of the unknown inputs. Then an AR model is constructed using the recovered input statistics, and a balanced realization model is then constructed using the ERA.

A. Extraction of Input Autocorrelations via a Least Squares Problem

Consider system (1) with zero initial conditions, the output \( y_k \) can be written as:

\[ y_k = \sum_{i=1}^{\infty} h_i u_{k-i} + v_k, \] (3)

where \( h_i \) are denoted as the Markov parameters of system (1):

\[ h_i = CA^{i-1}B, i = 1, 2, \ldots. \] (4)

For a linear time-invariant (LTI) system, under assumption A1 that \( A \) is stable, the output \( \{y_k\} \) is a wide-sense stationary process when \( \{u_k\} \) is wide-sense stationary. From the definition of the autocorrelation function of a WSS process, the output autocorrelation can be written as:

\[ R_{yy}(m) = E[y_k y_{k+m}^*] \]
\[ = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_i u_{k-i} u_{k+m-j}^* h_j^* + R_{vv}(m) \]
\[ = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_i R_{uu}(m+i-j) h_j^* + R_{vv}(m), \] (5)

where \( m = 0, \pm 1, \pm 2, \ldots \) is the time-lag between \( y_k \) and \( y_{k+m} \). Here, assumption A3 is used. Notice that \( R_{yy}(-m) \neq R_{yy}(m) \) when \( \{y_k\} \) is a sequence of complex valued vectors. We use \( x^* \) to denote the complex conjugate transpose of \( x \), and \( x^T \) to denote the transpose of \( x \).

Since \( v_k \sim N(0, \Omega) \), \( R_{vv}(m) = \Omega \) for \( m = 0 \), and \( R_{vv}(m) = 0 \), otherwise. We denote \( \hat{R}_{yy}(m) = R_{yy}(m) - R_{vv}(m) \), and hence, the relationship between input and output autocorrelation function is given by:

\[ \hat{R}_{yy}(m) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_i R_{uu}(m+i-j) h_j^*. \] (6)
For multiple input multiple output (MIMO) systems, \( h_i, R_{yy}(m), R_{uu}(m) \) are matrices. To solve for the unknown input autocorrelations \( R_{uu}(m) \), first we need to use a theorem from linear matrix equations [24], [25].

**Theorem 1:** Consider the matrix equation

\[
AXB = C, 
\]

where \( A, B, C, X \) are all matrices. If \( A \in \mathbb{C}^{m \times n} = (a_1, a_2, \ldots, a_n) \), where \( a_i \) are the columns of \( A \), then define

\[
\text{vec}(A) \in \mathbb{C}^{m \times n \times 1} \text{ as: vec}(A) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.
\]

The matrix equation (7) can be transformed into one vector equation:

\[
(B^T \otimes A)\text{vec}(X) = \text{vec}(C),
\]

where \( B^T \otimes A \) is the Kronecker product of \( B^T \) and \( A \). If \( A \) is an \( m \times n \) matrix and \( B \) is a \( p \times q \) matrix, then the Kronecker product \( A \otimes B \) is the \( np \times nq \) matrix block matrix:

\[
A \otimes B = \begin{pmatrix}
    a_{11}B & a_{12}B & \cdots & a_{1n}B \\
    a_{21}B & a_{22}B & \cdots & a_{2n}B \\
    \vdots & \vdots & \cdots & \vdots \\
    a_{m1}B & a_{m2}B & \cdots & a_{mn}B
\end{pmatrix}.
\]

By applying [24] [25], we can write this as:

\[
\text{vec}(\hat{R}_{yy}(m)) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \vec{h}_j \otimes \vec{h}_i \text{vec}(R_{uu}(m + i - j)),
\]

where \( \vec{h}_i \) denotes the matrix \( h_i \) with complex conjugated entries, and \( h_i^* = (h_i)^T \).

Now, we estimate the unknown input autocorrelations by the following procedure.

1) **Choose design parameter M:** For a stable system, we make the following assumption.

- **A5.** Assume that there exists a finite number \( M \) such that the Markov parameters of the system \( ||h_i|| \leq \delta, i > M \), where \( \delta \) is small enough.

Here, \( ||A|| = (\sum_{i,j=1}^{n} |a_{i,j}|^2)^{1/2} \) denotes the Frobenius norm of matrix \( A \), and \( ||x||_2 = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{1/2} \) denotes the Euclidean norm of vector \( x \). \( M \) is a design parameter that varies with different systems and can be chosen as large as desired. Thus, (10) can be written as:

\[
\text{vec}(\hat{R}_{yy}(m)) = \sum_{i=1}^{M} \sum_{j=1}^{M} \vec{h}_j \otimes \vec{h}_i \text{vec}(R_{uu}(m + i - j)).
\]

2) **Choose design parameters \( N_o, N_i \):** If \( \{u_k\} \) is WSS, then we make the following assumption.

- **A6.** \( \hat{R}_{yy}(m) \) only has significant values within a range \(-N_o \leq m \leq N_o\), and negligible values outside this range. Also, we assume the support of \( R_{uu}(m) \) is limited to \(-N_i \leq m \leq N_i\).

This is a rather standard assumption when computing a power spectrum from an autocorrelation function. The numbers \( N_o \) and \( N_i \) depend on the dynamic system and unknown inputs, and are design parameters that can be chosen as large as required.

Under assumption A6, we have the following proposition.

**Proposition 1:** The relation \( N_i \leq N_o \) holds, which implies that all significant input autocorrelations can be recovered from the output autocorrelations.

**Proof:** From the assumption that the support of \( \hat{R}_{yy} \) is limited to \((-N_o, N_o)\), we have:

\[
\hat{R}_{yy}(N_o + 1) = 0.
\]

Using (10), \( \vec{R}_{yy}(N_o + 1) \) can be written as:

\[
\vec{R}_{yy}(N_o + 1) = \sum_{i=1}^{\infty} \vec{h}_i \otimes \vec{h}_i \text{vec}(R_{uu}(N_o + 1)) + \cdots.
\]

If \( N_i > N_o \), which means \( R_{uu}(N_o + 1) \neq 0 \), then it follows that \( \hat{R}_{yy}(N_o + 1) \) is also not negligible, which contradicts the assumption. Thus, as a consequence, \( N_i \leq N_o \).

Thus, the following equation is used for computation of the unknown input autocorrelations.

\[
3) \text{Solve the least squares problem: We collect } 2N_o + 1 \text{ output autocorrelations, and from the above assumptions, there are } 2N_i + 1 \text{ unknown input autocorrelations:}
\]

\[
\begin{pmatrix}
    \text{vec}(\hat{R}_{yy}(-N_o)) \\
    \text{vec}(\hat{R}_{yy}(-N_o + 1)) \\
    \vdots \\
    \text{vec}(\hat{R}_{yy}(0)) \\
    \text{vec}(\hat{R}_{yy}(1)) \\
    \vdots \\
    \text{vec}(\hat{R}_{yy}(N_o))
\end{pmatrix} = C_{yu}
\begin{pmatrix}
    \text{vec}(R_{uu}(-N_i)) \\
    \text{vec}(R_{uu}(-N_i + 1)) \\
    \vdots \\
    \text{vec}(R_{uu}(0)) \\
    \text{vec}(R_{uu}(1)) \\
    \vdots \\
    \text{vec}(R_{uu}(N_i))
\end{pmatrix},
\]

where \( C_{yu} \) is the coefficient matrix and can be calculated from (14).

Under assumption A2 and A6, we have the following proposition.

**Proposition 2:** Equation (15) has a unique least squares solution \( \hat{R}_{uu}(m), m = \pm 1, \pm 2, \cdots, \pm N_i \).

**Proof:** First, we prove that A2 is a necessary condition. The \((i^{th}, j^{th})\) term in \( C_{yu} \) is:

\[
C_{yu}(i, j) = \sum_{s} \hat{h}_i \otimes \vec{h}_s, \text{ for some } s, t
\]

where \( i = 1, 2, \cdots, 2N_o + 1, j = 1, 2, \cdots, 2N_i + 1 \).

Since \( h_i = CA^{-1}B(\in C^{q \times p}), \text{ rank}(h_i) \leq \min(p, q) = k \).
Algorithm 1 Conjugate gradient algorithm

1) For a least squares problem $C_s\bar{x} = L_s$, where $C_s = C^*_s$, $\bar{x}$ is unknown.
2) Start with a randomly initial solution $\bar{x}_0$.
3) $r_0 = L_s - C_s\bar{x}_0$, $p_0 = r_0$.
4) for $k = 0$, repeat
5) $\alpha_k = \frac{r_k^T r_k}{p_k^T C_s p_k}$,
   $\bar{x}_{k+1} = \bar{x}_k + \alpha_k p_k$,
   $r_{k+1} = r_k - \alpha_k C_s p_k$,
   if $r_{k+1}$ is sufficiently small then exit loop.
   $\beta_k = \frac{r_{k+1}^T r_k}{r_k^T r_k}$,
   $p_{k+1} = r_{k+1} + \beta_k p_k$,
   $k = k + 1$.
end repeat.
6) The optimal estimation is $x_{k+1}$.

From the property that rank$(A \otimes B) = \text{rank}(A)\text{rank}(B)$, we have rank$(h_t \otimes h_a) = \text{rank}(h_t)\text{rank}(h_a) = k^2$, which means $C_{yu}(i,j)$ has at most $k^2$ independent columns. It follows that $C_{yu}$ has at most $k^2(2N_i + 1)$ independent columns, i.e., rank$(C_{yu}) \leq k^2(2N_i + 1)$. For $C_{yu} \in \mathbb{C}^{T(2N_i+1) \times p(2N_i+1)}$ to have a left inverse, $C_{yu}$ should have full column rank, i.e., rank$(C_{yu}) = p^2(2N_i + 1)$. Thus, $C_{yu}$ has a left inverse if and only if $k = p$, which implies $\min(p,q) = p$, and hence, $p \leq q$.

There are $q^2(2N_q + 1)$ equations with $p^2(2N_i + 1)$ unknowns, from the assumptions $p \leq q$ and $N_i \leq N_q$, there is an unique solution $\hat{R}_{yu}(m), m = 0, \pm 1, \ldots, \pm N_i$ in the least square sense, and $R_{yu}(m)$ is the input autocorrelations we extract from the output autocorrelations. □

Remark 2: The size of $C_{yu}$ is $q^2(2N_q+1) \times p^2(2N_i+1)$ and it would be large when $p$ and $q$ increase, and hence, large scale least squares problem need to be solved for systems with large number of inputs/outputs. For example, a modified conjugate gradients method [26] could be used as follows.

The least squares problem need to be solved is:

$$\text{vec}(\hat{R}_{yy}) = C_{yu}\text{vec}(R_{yu}),$$

and multiply $C_{yu}$ on both sides:

$$C^*_{yu}\text{vec}(\hat{R}_{yy}) = C^*_{yu}C_{yu}\text{vec}(R_{yu}).$$

If we denote $L_s = C^*_{yu}\text{vec}(\hat{R}_{yy}), \bar{x} = \text{vec}(R_{yu})$, and $C_s = C^*_{yu}C_{yu}$, then $C_s = C^*_s$, and the problem is equivalent to solve the least squares problem for $\bar{x}$:

$$C_s\bar{x} = L_s,$$

and a conjugate gradient method to solve this problem is summarized in Algorithm 1.

Denote $R_{yu}(m)$ as the “true” input autocorrelations, and $\Delta(m) = R_{yu}(m) - \hat{R}_{yu}(m)$ as the error of the input autocorrelations we extract, $\Delta(m)$ results from two design parameters: the choice of $M$ and $N_i$. We analyze the errors seperately, in the following.

Proposition 3: Denote $R_{yu}^M(m)$ as the input autocorrelations we extract by using $M$ Markov parameters of the dynamic system. The errors of input autocorrelations resulting from assumption A5 is: $||\Delta(m)|| \leq k_M \delta$, where $k_M$ is some constant, $\delta$ is defined in assumption A5.

The Perturbation theory [27] is used to prove the above result, and the proof is shown in Appendix 1.

Remark 3: Error analysis in the Fourier domain.

The power spectral density is defined as:

$$S_{yy}(\omega) = \sum_{k=-\infty}^{\infty} \tilde{R}_{yy}(k)e^{-jk\omega},$$

$$S_{yy}(\omega) = \sum_{k=-\infty}^{\infty} \tilde{R}_{yu}(k)e^{-jk\omega},$$

Thus, by substituting (8), the relationship between the output power spectral density and input power spectral density is:

$$S_{yy}(\omega) = \sum_{k=-\infty}^{\infty} \sum_{i=1}^{M} \sum_{t=1}^{\infty} \tilde{h}_iR_{yu}(k + i - t)h^*_t e^{-jk\omega} = S^M_{yy}(\omega) + \Delta S_M(\omega),$$

Thus, $||\Delta S_M(\omega)|| \leq k_1 \delta$, where $k_1$ is some constant. Hence, the truncation error by using $M$ Markov parameters can be seen to be a small perturbation in the frequency domain.

Proposition 4: Denote $R_{uu}^N(m)$ as the input autocorrelations we extract under assumption A6. The errors resulting from this assumption is $||\Delta_N(m)|| \leq k_N \delta$, where $k_N$ is some constant, $\delta$ is defined in assumption A6.

The proof is shown in Appendix 1.

Remark 4: Error analysis in frequency domain.

$$S_{yy}(\omega) = \sum_{k=-\infty}^{\infty} \sum_{i=1}^{\infty} \sum_{t=1}^{\infty} \tilde{h}_iR_{yu}(k + i - t)h^*_t e^{-jk\omega}$$

$$+ \sum_{k=-\infty}^{\infty} \sum_{i=1}^{\infty} \sum_{t=1}^{\infty} \tilde{h}_iR_{yu}(k + i - t)h^*_t e^{-jk\omega}$$

$$= S^N_{yy}(\omega) + \Delta S_N(\omega),$$

where $||\Delta S_N(\omega)|| \leq k_2 \delta$, where $k_2$ is some constant.
Under the assumptions A5 and A6, the following proposition considers the total errors of input autocorrelations we recover.

**Proposition 5:** Denote $\hat{R}_{uu}(m)$ as the input autocorrelation function we estimate from the output autocorrelations, and let $\Delta(m) = R_{uu}(m) - \hat{R}_{uu}(m)$ be the error between the estimated input autocorrelation and the “true” input autocorrelation. Then $\|\Delta(m)\| \leq k\delta$, where $k$ is some constant.

Proposition 3 and 4 are used for the proof, and the proof is shown in Appendix III. The results above show that if $M$, $N_i$, $N_o$ are chosen large enough, the errors in estimating the input autocorrelations can be made arbitrarily small.

**B. Construction of the AR Based Innovations Model**

After we extract the input autocorrelations from the output autocorrelations, we want to construct a system which will generate the same statistics as the ones we recovered in Section III-A. If assumption A4 is satisfied, i.e., $\{u_k\}$ is WSS with a rational power spectrum, the power spectrum of $u_k$ is continuous, and can be factored [28]. Such system can be constructed by using an Autoregressive (AR) model. In an AR model, the time series can be expressed as a linear function of its past values, i.e.,

$$u(k) = \sum_{i=1}^{M_i} a_i u(k-i) + \epsilon(k),$$

(26)

where $\epsilon(k)$ is white noise with distribution $N(0, \Omega_r)$. $M_i$ is the order of the AR model, and $a_i$, $i = 1, 2, \cdots, M_i$ are the coefficient matrices. For a vector autoregressive model with complex values, the Yule-Walker equation [29] which is used to solve for the coefficients needs to be modified. The modified Yule-Walker equation can be written as:

$$
\begin{pmatrix}
R_{uu}(1) & R_{uu}(2) & \cdots & R_{uu}(M_i) \\
R_{uu}(0) & R_{uu}(1) & \cdots & R_{uu}(M_i-1) \\
R_{uu}(1) & R_{uu}(0) & \cdots & R_{uu}(2-M_i) \\
\vdots & \vdots & \ddots & \vdots \\
R_{uu}(M_i-1) & R_{uu}(M_i-2) & \cdots & R_{uu}(0)
\end{pmatrix}
\times
\begin{pmatrix}
a_1^* \\
a_2^* \\
\vdots \\
a_{M_i}^*
\end{pmatrix}
$$

(27)

Equation (27) is used to solve for the coefficient matrices $a_i$, $i = 1, 2, \cdots, M_i$. The covariance of the residual white noise $\epsilon(k)$ can be solved using the following equation:

$$R_{\epsilon\epsilon}(m) = R_{uu}(m) - \sum_{i=1}^{M_i} \sum_{j=1}^{M_i} a_i R_{uu}(m+i-j) a_j^*,$$

(28)

where $\Omega_r = R_r(0)$. The balanced minimal realization for the AR model (26) can be expressed as:

$$
\begin{align*}
\eta_k &= A_n \eta_{k-1} + B_n u_{k-1}, \\
u_k &= C_n \eta_k + \epsilon_k
\end{align*}
$$

(29)

where $A_n, B_n, C_n$ are solved by using the ERA technique [22] with $a_i, i = 1, \cdots, M_i$ as the Markov parameters of the system. A brief description of the ERA is given in Appendix IV.

Equation (29) is equivalent to:

$$u_k = C_n \eta_k + \epsilon_k,$$

(30)

where $\epsilon_k$ is white noise with covariance $\Omega_r$. By using the Cholesky Decomposition, we can find a unique lower triangular matrix $P$ such that:

$$\Omega_r = PP^*.$$

(31)

If $w_k$ is white noise with distribution $N(0, 1)$, then $Pw_k$ would be white noise with distribution $N(0, \Omega_r)$. Thus, the innovation model we construct that has the same statistics as the unknown input system (2) is:

$$u_k = C_n \eta_k + Pw_k,$$

(32)

where $w_k$ is a randomly white noise with standard normal distribution.

Under assumption A4, we have the following proposition.

**Proposition 6:** Denote $\hat{R}_{uu}(m)$ as the input autocorrelations recovered from the measurements, then $R_{uu}(m)$ can be reconstructed exactly by using the innovations model (32), i.e., $\hat{R}_{uu}(m) = \hat{R}_{uu}(m)$, where $\hat{R}_{uu}(m)$ is the input autocorrelations of the realization of system (32).

From Proposition 5 and 6 under the same assumptions, the following corollary immediately follows.

**Corollary 1:** Denote $u_k$ as the actual unknown input process, and $R_{uu}(m)$ as the actual input autocorrelation function. Then $\|R_{uu}(m) - \hat{R}_{uu}(m)\| \leq k_n \delta$, where $k_n$ is some constant, when $\delta$ is small enough. System (32) is an innovations model for the unknown input $u_k$.

The procedure of constructing the innovations model is summarized in Algorithm 2.

**Remark 5:** For real valued system, we can save the computation by using the properties of autocorrelation functions:

$$R_{uu_i}(m) = R_{uu_i}(m),$$

$$R_{uu_i}(m) = R_{uu_j}(m), i \neq j$$

(33)

Thus, we only need to collect $N_o + 1$ output autocorrelations and have $p^2(N_o + 1)$ equations with $q^2(N_i + 1)$ unknowns in (15).

**Remark 6:** A generalization to the joint state and unknown input estimation.

When the unknown inputs affect both the states and outputs, i.e.

$$x_{k+1} = Ax_k + Bu_k,$$

$$y_k = Cx_k + Du_k + v_k,$$

(34)

where $u_k$ is the stochastic unknown input, $v_k$ is the mea-
A. Augmented State Kalman Filter

An ROM based filter is also constructed using the BPOD for system with states augmented by the unknown input states. inputs, we apply the standard Kalman filter on the augmented system, where

\[ y_k = \sum_{i=1}^{M} h_i u_{k-i} + D u_k + v_k, \tag{35} \]

and the relationship between output autocorrelations and input autocorrelations is:

\[ R_{yy}(m) = \sum_{i=1}^{M} \sum_{j=1}^{M} h_i R_{uu}(m + i - j) h^*_j + R_{vv}(m) + \]

\[ \sum_{i=1}^{N} h_i R_{uu}(m + i) D^* + \sum_{i=1}^{N} D R_{uu}(m - j) h^*_j + D R_{uu}(m) D^*, \]

which can also be formulated as a least squares problem [15], and an unknown input system may be realized following the same procedure as in Algorithm 2.

IV. AUGMENTED STATE KALMAN FILTER AND MODEL REDUCTION

After we construct an innovations model for the unknown inputs, we apply the standard Kalman filter on the augmented system with states augmented by the unknown input states. A ROM based filter is also constructed using the BPOD for reducing the computational cost of the resulting filter.

A. Augmented State Kalman Filter

The full order system can be represented by augmenting the states of the original system as:

\[
\begin{bmatrix}
    x_{k+1} \\
    \eta_{k+1}
\end{bmatrix} = \begin{bmatrix}
    A & BC_n \\
    0 & A_n + B_n C_n
\end{bmatrix}
\begin{bmatrix}
    x_k \\
    \eta_k
\end{bmatrix} + \begin{bmatrix}
    BP \\
    B_n P
\end{bmatrix} w_k,
\]

\[ y_k = \begin{bmatrix}
    C \\
    0
\end{bmatrix}
\begin{bmatrix}
    x_k \\
    \eta_k
\end{bmatrix} + v_k, \tag{37} \]

where \( w_k \) is white noise with standard normal distribution. \( v_k \) is white noise with known covariance. Thus, we may now use the standard kalman filter for state estimation of the augmented system [37].

B. Unknown Input Estimation Using Model Reduction

For large scale systems, we can use model reduction technique such as Balanced Proper Orthogonal Decomposition (BPOD) to construct a reduced order model (ROM) first, and then extract the input autocorrelations from the reduced order model. We apply the Kalman filter to the ROM to reduce the computational cost. A brief description of BPOD is given in Appendix IV. For a large scale system with a large number of inputs and outputs, we can also use the randomized proper orthogonal decomposition (RPDO) technique [30] for model reduction.

The ROM system is extracted from the full order system using the BPOD and is denoted by:

\[ x_k = A_r x_{k-1} + B_r u_{k-1}, \]

\[ y_k = C_r x_k + v_k. \tag{38} \]

Let \( \hat{h}_i = C_r A_{r}^{-1} B_r, i = 1, 2, \cdots, M \) be the Markov parameters of the ROM. Then the relationship between input autocorrelations and output autocorrelations can be written as:

\[ \hat{R}_{yy}(m) = \sum_{i=1}^{M} \sum_{j=1}^{M} \hat{h}_i R_{uu}(m + i - j) \hat{h}^*_j, \tag{39} \]

Following the same procedure as in Algorithm 2 we can now recover the input autocorrelations, and construct an innovations model which can generate the same statistics as the unknown inputs. The advantage of using model reduction is that for a large scale system, computing \( \hat{h}_i = C_r A_{r}^{-1} B_r \) much faster than computing \( h_i = C A^{-1} B \) because of the reduction in the size of \( A \). Also, the order of the ROM is much smaller than the order of the full order system, and thus the computational cost of using the Kalman filter is much reduced. Hence, even with the augmented states, the standard Kalman filter remains computationally tractable.

Remark 7: To reduce the computational cost of the augmented states in Kalman filter, we can also use the existing optimal two-stage or three-stage kalman filtering technique [13], [15], which decouple the augmented filter into two parallel reduced order filters. These techniques are preferable when the order of the innovations model is high, while the BPOD based ROM filter is preferable when the order of the dynamic system is high.

V. COMPUTATIONAL RESULTS

We test the method on a one-dimensional heat equation and the perturbed laminar flow equation. We construct the unknown input system by using both the full order system as well as the ROM constructed by BPOD. We check the results by comparing the autocorrelation functions of the inputs, outputs and the states. Also, we show the state estimation using the Kalman filter. We define the relative error as:

\[ R_{relative} = \frac{\|R_{true} - R_{est}\|}{\|R_{true}\|}, \tag{40} \]
\( R_{true} \): actual output/input/state autocorrelation function of the system
\( R_e \): estimated output/input/state autocorrelation function

In the following, we will show simulation results for the stochastically perturbed 1D heat equation and the laminar flow problem.

A. Heat Equation

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

\[
\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2} + f,
\]

(41)

with Neumann boundary condition,

\[-\frac{\partial T}{\partial x}|_{x=0} = -\frac{\partial T}{\partial x}|_{x=L} = 0.5,
\]

(42)

where \( \alpha \) is the thermal diffusivity, and \( f \) is the unknown forcing. The system is discretized by finite difference approach, we assume that there are two unknown forcings and two measurements are taken to satisfy the assumption A2. The measurement noise is white noise with covariance \( 0.1I_{2 \times 2} \). In the simulation, the unknown inputs are colored noise generated by a randomly generated second order system.

The design parameters \( M = 200, N_i = N_o = 100 \) are chosen as follows. \( M \) is chosen so that the Markov parameters \( \| h_i \| \approx 0, i > M \). \( N_i \) and \( N_o \) are chosen by trial and error. First, we randomly choose a suitable \( N_i \) and \( N_o \), where \( N_i \leq N_o \). Then we follow the AR based unknown input realization procedure, and construct the augmented state system [37]. Given the white noise processes \( w_k, v_k \) perturbing the system, we check the output statistics of the augmented state system [37]. If the errors are small enough, we stop, otherwise, we increase the values of \( N_i \) and \( N_o \), and repeat the same procedure until the errors are negligible. Notice that increasing \( M, N_i, N_o \) would increase the accuracy of the input statistics we can recover, but also increases the computational cost.

In the plots below, we show the comparison of the input autocorrelations we recover with the actual input autocorrelations. Since there are two inputs, the cross-correlation function between input 1 and input 2 are also included in the input autocorrelation plots.

Fig. 1 shows the comparison of the output autocorrelations between the estimated output autocorrelations and the actual autocorrelations. The comparison of the state autocorrelations is shown in Fig. 3 for a few randomly chosen states.

From the simulation results above, we can see that the unknown input autocorrelation can be recovered accurately, and the statistics of the states and outputs of the dynamic system perturbed by the innovations model are approximately the same as the actual system. We also construct the innovations model by using a ROM generated by BPOD. The full order system has 50 states, and the ROM has 20 states. We compare the relative error of the input autocorrelations in Fig. 4.
The comparison of the relative error of output autocorrelations is shown in Fig. 5 and the comparison of the relative error of state autocorrelations is shown in Fig. 6.

We can see that the input autocorrelation relative error by using ROM is almost the same as the relative error using full order system, and as a consequence, the state autocorrelation relative error using the ROM is a little worse compared to using the full order system. In this simulation, the output autocorrelation relative error using ROM is even smaller than using the full order system, in common, the errors are on the same scale, while the computations are reduced.

Next, we apply the Kalman filter to the augmented system, and the comparison of the state estimation is shown in Fig. 7. We randomly choose two states and show the comparison of the actual state with the estimated states. The state estimation error and $3\sigma$ bounds are shown.

The state estimation using ROM is shown in Fig. 8. It can be seen that the kalman filter using the ROM performs reasonably well, hence, for a large scale system, the computational complexity of ASKF can be reduced by using the

---

**Fig. 4.** Comparison of input autocorrelation relative error

**Fig. 5.** Comparison of output autocorrelation relative error

**Fig. 6.** Comparison of state autocorrelation relative error

**Fig. 7.** Comparison of state estimation

**Fig. 8.** State estimation using ROM

---

-BPOD-
B. Orr-Sommerfeld Equation

Consider the three-dimensional flow between two infinite plates (at $y = \pm 1$) driven by a gradient in the streamwise $x$ direction. The mean velocity profile is given by $U(y) = 1 - y^2$. At each wavenumber pair $(\alpha, \beta)_{mn}$, the wall-normal velocity $v(x, y, z, t)$ and wall-normal vorticity $\eta(x, y, z, t)$ are:

$$v(x, y, z, t) = \tilde{v}_{mn}(y, t)e^{i(\alpha x + \beta z)}, \quad (43)$$

$$\eta(x, y, z, t) = \tilde{\eta}_{mn}(y, t)e^{i(\alpha x + \beta z)}. \quad (44)$$

Denote

$$\tilde{q}_{mn}(y, t) = \left( \tilde{v}_{mn}(y, t) \tilde{\eta}_{mn}(y, t) \right), \quad (45)$$

where $\tilde{\cdot}$ denotes the Fourier transformed variable, and $(\cdot)_{mn}$ denotes the wavenumber pair $(\alpha, \beta)_{mn}$.

The evolution of the flow in Fourier domain can be written as:

$$\frac{d}{dt} M \tilde{q}_{mn} + L \tilde{q}_{mn} = T f(y, t), \quad (46)$$

where

$$M = \begin{pmatrix} -\Delta & 0 \\ 0 & I \end{pmatrix}, \quad (47)$$

$$L = \begin{pmatrix} -i\alpha U \Delta + i\alpha U'' + \Delta^2/Re & 0 \\ i\beta U' & i\alpha U - \Delta/Re \end{pmatrix}. \quad (48)$$

Operator $T$ transforms the forcing $f = (f_1, f_2, f_3)^T$ on the evolution equation for the velocity vector $(u, v, w)^T$ into an equivalent forcing on the $(v, \eta)^T$ system [18],

$$T = \begin{pmatrix} i\alpha D & k^2 \\ i\beta & 0 \end{pmatrix} \quad (49)$$

where

$$k^2 = \alpha^2 + \beta^2, \quad (50)$$

$$\Delta = D^2 - k^2, \quad (51)$$

and $D, D^2$ represent the first and second order differentiation operators in the wall-normal direction. The forcing $f(y, t)$ accounts for the nonlinear terms and the external disturbances via an unknown stochastic model.

The boundary conditions on $v$ and $\eta$ correspond to no-slip solid walls

$$v(\pm 1) = Dv(\pm 1) = \eta(\pm 1) = 0. \quad (52)$$

System (46) can be discretized using Chebyshev polynomials, and in the simulation, we assume there are two unknown inputs and two measurements.

In the simulation, the design parameters $M = 1000$, $N_t = N_o = 100$ are chosen by trial and error as explained before. The unknown input $f$ is assumed to be a colored noise generated by a third order linear complex system. The realization of the unknown inputs is a second order system. The measurement noise is white noise with covariance $0.1I_{2\times 2}$.

First, we show the comparison of the input autocorrelations we recover with the actual input autocorrelations in complex plane. Since there are two inputs, thus, the cross-correlation function between input 1 and input 2 are also included in the input autocorrelations.

![Fig. 9. Comparison of input autocorrelations](image)

![Fig. 10. Comparison of output autocorrelations](image)

Before we apply the ASKF for the state estimation, we compare the statistics of the states and outputs of the system perturbed by the unknown inputs we construct and the actual system. Fig. [10] shows the comparison between the estimated output autocorrelations and the actual autocorrelations. The comparison of the state autocorrelations is shown in Fig. [11] for some randomly chosen states.
The comparison of the relative error of state autocorrelations is shown in Fig. 13.

Fig. 14. Comparison of state autocorrelation relative error

We can see that the statistics reconstructed by using the ROM is not as accurate as using the full order system, however, the relative error is on the same scale, and hence, the computational cost is reduced without losing too much accuracy.

The comparison of the state estimation using the ASKF is shown in Fig. 15. We randomly choose two states and show the comparison of the actual state with the estimated states. The state estimation error and $3\sigma$ bounds are shown. Since the error is complex valued, only the absolute value of the error is shown.

Fig. 15. State estimation using full order system

The state estimation using ROM is shown in Fig. 16. It can be seen that the kalman filter using the ROM perform well, and hence, for a large scale system, the computational complexity of ASKF can be reduced by using the BPOD.

VI. Conclusion

In this paper, we have proposed a balanced unknown input realization method for the state estimation of system.
with unknown stochastic inputs. The unknown inputs are assumed to be a wide sense stationary process with a rational power spectrum, and no other prior information about the unknown inputs needs to be known. We recover the unknown inputs statistics from the output data using a least-squares procedure, and then construct a balanced minimal realization of the unknown inputs using an AR model and the ERA technique. The recovered innovations model is used for state estimation, and the standard Kalman filter is applied on the augmented system. The next step in this process would require us to consider more complex realistic problems in fluid flow applications, and cases where the unknown numbers of inputs/outputs are large, and also cases where the locations of the inputs are unknown.

APPENDIX I

PROOF OF PROPOSITION 4

Proof: The output autocorrelation function using the first $M$ Markov parameters is:

$$\hat{R}_{yy}(m) = \sum_{i=1}^{M} \sum_{j=1}^{M} h_i R_{uu}(m + i - j) h_j^*.$$  (53)

Comparing with (5), the output autocorrelation errors resulting from using $M$ Markov parameters is:

$$\Delta_1(m) = \sum_{i=M+1}^{\infty} \sum_{j=1}^{M} h_i R_{uu}(m + i - j) h_j^* + \sum_{i=M+1}^{\infty} \sum_{j=M+1}^{M} h_i R_{uu}(m + i - j) h_j^* + \sum_{i=1}^{M} \sum_{j=M+1}^{\infty} h_i R_{uu}(m + i - j) h_j^*.$$  (54)

From assumption A5, by choosing $M$ large enough, we have $|h_i| \leq \delta$, $i > M$, where $\delta$ is small enough, thus,

$$||\Delta_1(m)|| \leq \sum_{i=M+1}^{\infty} \sum_{j=1}^{M} \delta \times ||R_{uu}(m + i - j)|| |h_j^*| + \sum_{i=M+1}^{\infty} \sum_{j=M+1}^{M} \delta \times ||R_{uu}(m + i - j)|| \times \delta + \sum_{i=1}^{M} \sum_{j=M+1}^{\infty} |h_i| ||R_{uu}(m + i - j)|| \times \delta \leq k_3 \delta,$$  (55)

where $k_3$ is some constant.

Denote $C_{yu}$ as the “true” coefficient matrix and $C_{yu}^M$ as the coefficient matrix using $M$ Markov parameters, we need to solve the least squares problem:

$$\text{vec}(\hat{R}_{yy}) = C_{yu}^M \text{vec}(R_{uu}^M).$$  (56)

where $R_{uu}^M$ is the input autocorrelation we recover from using $M$ Markov parameters, and $\text{vec}(\hat{R}_{yy})$ is defined in (15).

Since $||\text{vec}(\hat{R}_{yy}(m)) - \text{vec}(\hat{R}_{yy}^M(m))||_2 = ||\Delta_2(m)|| \leq k_3 \delta$, we have $\text{vec}(\hat{R}_{yy}(m)) = \text{vec}(\hat{R}_{yy}^M(m)) + \Delta_2(m)$, where $||\Delta_2(m)||_2 \leq k_3 \delta$, or equivalently

$$\text{vec}(\hat{R}_{yy}) = \text{vec}(\hat{R}_{yy}^M) + \Delta_2,$$  (57)

Consider (57). $\text{vec}(\hat{R}_{yy})$ and $\text{vec}(\hat{R}_{yy}^M)$ can be written as:

$$\text{vec}(\hat{R}_{yy}) = C_{yu} \text{vec}(R_{uu}),$$
$$\text{vec}(\hat{R}_{yy}^M(m)) = C_{yu}^M \text{vec}(R_{uu}),$$  (58)

Substitute into (57), we have:

$$C_{yu} \text{vec}(R_{uu}) - C_{yu}^M \text{vec}(R_{uu}) = \Delta_2.$$  (59)

Since $(C_{yu}^M)^{-1}$ exists, we have:

$$\text{vec}(R_{uu}) - \text{vec}(R_{uu}^M) = (C_{yu}^M)^{-1} \Delta_2,$$  (60)

which means:

$$||\text{vec}(R_{uu}) - \text{vec}(R_{uu}^M)||_2 \leq k_M \delta,$$  (61)

where $k_M$ is some constant. Thus, we have $||\Delta_M(m)|| \leq k_M \delta$, where $k_M$ is some constant.

APPENDIX II

PROOF OF PROPOSITION 4

Proof: (10) can be separated into two parts:

$$\text{vec}(\hat{R}_{yy}(m)) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \tilde{h}_j \otimes h_i \text{vec}(R_{uu}(m + i - j))$$
$$\quad + \sum_{i=1}^{N_i} \sum_{j=1}^{\infty} \tilde{h}_j \otimes h_i \text{vec}(R_{uu}(m + i - j)),$$  (62)

Thus, it can be written as:

$$\text{vec}(\hat{R}_{yy}(m)) = \text{vec}(\hat{R}_{yy}^N(m)) + \Delta_4(m),$$  (63)
where
\[ \| \Delta_4(m) \|_2 = \left\| \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \hat{h}_j \otimes h_i \text{vec}(R_{uu}(m + i - j)) \right\|_2 \]
\[ \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \| \hat{h}_j \otimes h_i \|_2 \leq k_4 \delta, \quad (64) \]

where \( k_4 \) is some constant. \( \| A \|_2 \) denotes the induced 2-norm of matrix \( A \). Following the same procedure as in Proposition [3] it can be proved that \( \| \Delta_N(m) \| \leq k_N \delta \), where \( k_N \) is some constant.

**APPENDIX III**

**Proof of Proposition [5]**

*Proof:* Denote output autocorrelation in (14) as \( \hat{R}_{yy}(m) \), comparing (14) with (10), the output autocorrelation error resulting from assumption A5 and A6 is:
\[ \text{vec}(\hat{R}_{yy}) - \text{vec}(\hat{R}_{yy}^c) = \Delta_2 + \sum_{i=1}^{M} \sum_{j=1}^{M} \hat{h}_j \otimes h_i \text{vec}(R_{uu}(m + i - j)) \leq \Delta_2 + \Delta_4. \quad (65) \]

Thus
\[ \| \text{vec}(\hat{R}_{yy}) - \text{vec}(\hat{R}_{yy}^c) \|_2 \leq \| \Delta_2 \|_2 + \| \Delta_4 \|_2 \leq k_5 \delta, \quad (66) \]

where \( k_5 \) is some constant. Following the same procedure as in Proposition [3] we can prove:
\[ \| \Delta(m) = R_{uu}(m) - \hat{R}_{uu}(m) \| \leq k \delta. \quad (67) \]

**APPENDIX IV**

**Brief Description of ERA and BPOD**

The Eigensystem Realization Algorithm is summarized as follows.

Run impulse response simulations of the linear system (1), and collect the snapshots of the outputs \( y_k \) in the following pattern:
\[ Y_1 = CB, Y_2 = CAB, \ldots, Y_k = CA^{k-1}B, \quad (68) \]

where \( CA^kB \) are known as Markov parameters. Construct a Hankel matrix \( H(k) \)
\[ H(k - 1) = \begin{pmatrix} Y_k & Y_{k+1} & \cdots & Y_{k+\beta - 1} \\ Y_{k+1} & Y_{k+2} & \cdots & Y_{k+\beta} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{k+\alpha - 1} & Y_{k+\alpha} & \cdots & Y_{k+\alpha+\beta-2} \end{pmatrix}. \quad (69) \]

Solve the singular value decomposition (SVD) problem of \( H(0) \), i.e.,
\[ H(0) = R\Sigma S^* \quad (70) \]

Denote \( \Sigma_n \) as the first \( n \) non-zero singular value of \( \Sigma \), and \( R_n, S_n \) as the matrices formed by the first \( n \) columns of \( R \) and \( S \) respectively. Then the realization for the ERA is:
\[ \hat{A} = \Sigma_n^{-1/2}R_n^*H(1)S_n\Sigma_n^{-1/2}, \]
\[ \hat{B} = \text{first } p \text{ columns of } \Sigma_n^{1/2}S_n^* \]
\[ \hat{C} = \text{first } q \text{ rows of } R_n\Sigma_n^{1/2} \quad (71) \]

The Balanced POD procedure using the impulse response of the primal and adjoint system and is summarized below.

Consider the linear system (1), and denote \( B = [b_1, b_2, \ldots, b_p] \), \( C = [c_1, c_2, \ldots, c_q] \). We collect the impulse response of the primal system by using \( b_j, j = 1, 2, \ldots, p \), as initial conditions for the simulation of the system,
\[ x_k = Ax_{k-1}, \quad (72) \]

If we take \( \alpha \) snapshots across the trajectories at time \( t_1, t_2, \ldots, t_\alpha \), resulting an \( N \times p \alpha \) matrix
\[ X = [x_1(t_1), \ldots, x_1(t_\alpha), \ldots, x_p(t_1), \ldots, x_p(t_\alpha)], \quad (73) \]

where \( x_j(t_k) \) is the state snapshot \( x_k \) with \( b_j \) as the initial condition.

Similarly, we use the transposed rows of the output matrix \( c_i^* \), as the initial conditions for the simulations of the adjoint system \( A^* \),
\[ z_k = A^*z_{k-1}, \quad (74) \]

and take \( \beta \) snapshots across trajectories, leading to the adjoint snapshot ensemble \( Y \),
\[ Y = [z_1(t_1), \ldots, z_1(t_\beta), \ldots, z_p(t_1), \ldots, z_p(t_\beta)], \quad (75) \]

where \( z_i(t_k) \) is the state snapshot \( z_k \) with \( c_i^* \) as the initial condition.

The Hankel matrix \( H \) is constructed as:
\[ H = Y^*X. \quad (76) \]

Then we solve the SVD problem of the matrix \( H \):
\[ H = Y^*X = U\Sigma V^*. \]

Assume that \( \Sigma_1 \) consists of the first \( r \) non-zero singular values of \( \Sigma \), and \( (U_1, V_1) \) are the corresponding left and right singular vectors from \( (U, V) \), then the POD projection matrices can be defined as:
\[ T_r = XV_1\Sigma_1^{-1/2}, \quad T_l = YU_1\Sigma_1^{-1/2}, \]

and the reduced order model constructed using BPOD method is:
\[ A_r = T_l^*A T_r \]
\[ B_r = T_l^*B \]
\[ C_r = C T_r \quad (79) \]

**References**

[1] S.-H. Wang, E.J. Davison, and P. Dorato. “Observing the states of systems with unmeasurable disturbances,” IEEE Transactions on Automatic Control, vol. 20, No. 5, pp. 716–717, 1975.
[2] S. Bhattacharya, “Observer design for linear systems with unknown inputs,” IEEE Transactions on Automatic Control, vol. AC-23, No.3, pp. 483–484, 1978.

[3] P. Kudva, N. Viswanadh, and A. Ramakrishna, “Observers for linear systems with unknown inputs,” IEEE Transactions on Automatic Control, vol. 25, No.1, pp. 113–115, 1980.

[4] M. Hou and P. Muller, “Design of observers for linear systems with unknown inputs,” IEEE Transactions on Automatic Control, vol. 37, No.6, pp. 871–875, 1992.

[5] S. Hui and S. H. Zak, “Low-order state estimators and compensators for dynamical systems with unknown inputs,” Systems & Control Letters, vol. 21, No.6, pp. 493–502, 1993.

[6] M. Darouach, M. Zasadzinski, and S. Xu, “Full-order observers for linear systems with unknown inputs,” IEEE Transactions on Automatic Control, vol. 39, No.3, pp. 606–609, 1994.

[7] S. K. Spurgeon, “Sliding mode observers: a survey,” International Journal of Systems Science, vol. 39, No.8, pp. 751–764, 2008.

[8] K. Kalsi, J. Lian, S. Hui, and S. H. Zak, “Sliding-mode observers for systems with unknown inputs: A high-gain approach,” Automatica, vol. 46, Issue 2, pp. 347–353, 2010.

[9] M. Darouach, M. Zasadzinski, A. B. Onana, and S. Nowakowski, “Kalman filtering with unknown inputs via optimal state estimation of singular systems,” International Journal of Systems Science, vol. 26(10), pp. 2015–2028, 1995.

[10] M. Hou and R. J. Patton, “Optimal filtering for systems with unknown inputs,” IEEE Transactions on Automatic Control, vol. 43, No.3, pp. 445–449, 1998.

[11] D. Koenig and S. Mammar, “Reduced order unknown input kalman filter: application for vehicle lateral control,” in Proceedings of American Control Conference, 2003, pp. 4353–4358.

[12] C.-S. Hsieh, “A unified framework for state estimation of nonlinear stochastic systems with unknown inputs,” in Proceedings of 9th IEEE Asian Control Conference, 2013.

[13] C.-S. Hsieh and F.-C. Chen, “Optimal solution of the two-stage kalman estimator,” IEEE Transactions on Automatic Control, vol. 44, pp. 194–199, 1999.

[14] S. Kanav and M. Verhaegen, “Two-stage kalman filtering via structured square-root,” Communications in information and systems, vol. 5, No.2, pp. 143–168, 2005.

[15] F. B. Hmida, K. Khemiri, J. Ragot, and M. Gossa, “Three-stage kalman filter for state and fault estimation of linear stochastic systems with unknown inputs,” Journal of the Franklin Institute, vol. 349, pp. 2369–2388, 2012.

[16] S. Gillijns and B. D. Moor, “Unbiased minimum-variance input and state estimation for linear discrete-time systems,” Automatica, vol. 43, pp. 111–116, 2007.

[17] C.-S. Hsieh, “Extension of unbiased minimum-variance input and state estimation for systems with unknown inputs,” Automatica, vol. 45, pp. 2149–2153, 2009.

[18] J. Hepffner, M. Chevalier, T. R. Bewley, and D. S. Henningson, “State estimation in wall-bounded flow systems. part 1. perturbed laminar flows,” Journal of Fluid Mechanics, vol. 534, pp. 263–294, 2005.

[19] R. K. Mehra, “On the identification of variances and adaptive kalman filtering,” IEEE Transactions on Automatic Control, vol. 15, No.2, pp. 175–184, 1970.

[20] J. Dunik and M. Simandl, “Estimation of state and measurement noise covariance matrices by multi-step prediction,” in Proceedings of the 17th IFAC World Congress, 2008, pp. 3689–3694.

[21] B. Friedlander and B. Porat, “The modified yule-walker method of arma spectral estimation,” IEEE Transactions on Aerospace and Electronic Systems, vol. AES-20, No.2, pp. 158–173, 1984.

[22] D. Yu and S. Chakravorty, “A randomized proper orthogonal decomposition technique,” arXiv: 1312.3976, Tech. Rep., 2013, submitted to Automatica.
