

Dynamic Mode Decomposition and Robust Estimation: Case Study of a 2D Turbulent Boussinesq Flow

Vijayshankar, Sanjana; Nabi, Saleh; Chakrabarty, Ankush; Grover, Piyush; Benosman, Mouhacine

TR2020-091 July 03, 2020

Abstract

This paper focuses on an application of dynamic mode decomposition (DMD) identification methods and robust estimation theory to thermo-fluid systems modelled by the Boussinesq equations. First, we use Dynamic Mode Decomposition with control (DMDc) to construct a reduced order linear model for the Boussinesq equations. Due to inherent model uncertainties in real applications, we propose robust estimators that minimize an H infinity norm from disturbance to estimation error. The disturbances we consider here stem from uncertainty in boundary conditions and unknown inputs acting on walls. Numerical simulations on a challenging turbulent flow, of the 2D Boussinesq equations, is used to demonstrate the potential of our approach.

American Control Conference (ACC) 2020

Dynamic Mode Decomposition and Robust Estimation: Case Study of a 2D Turbulent Boussinesq Flow

Sanjana Vijayshankar¹, Saleh Nabi^{2,*}, Ankush Chakrabarty², Piyush Grover^{2,3}, Mouhacine Benosman⁴

Abstract—This paper focuses on an application of dynamic mode decomposition (DMD) identification methods and robust estimation theory to thermo-fluid systems modelled by the Boussinesq equations. First, we use Dynamic Mode Decomposition with control (DMDc) to construct a reduced order linear model for the Boussinesq equations. Due to inherent model uncertainties in real applications, we propose robust estimators that minimize an \mathcal{H}_∞ norm from disturbance to estimation error. The disturbances we consider here stem from uncertainty in boundary conditions and unknown inputs acting on walls. Numerical simulations on a challenging turbulent flow, of the 2D Boussinesq equations, is used to demonstrate the potential of our approach.

I. INTRODUCTION

Buildings contribute 40% of energy consumption in the United States. Building simulation, composed of a coupled dynamics of heating, ventilation and air-conditioning (HVAC) equipment and airflow within the built environment, is essentially complex, multi-scale, multi-physics, and subject to a wide variety of disturbances and uncertainties [1]. Designing computationally tractable tools for analysis and synthesis of controllers and estimators for energy-efficient buildings is a challenging problem due to inherent complexity of airflow models. A reasonable mathematical model for the airflow dynamics is described by Boussinesq equations, which are partial differential equations (PDEs) [2], [3]. Boussinesq equations are a combination of the Navier-Stokes (NS) equation and heat transfer. The numerical simulation of PDEs is carried out by means of discretization, which results in a large-scale discrete system of the order of $10^4 - 10^6$ states. The challenge is to develop scalable computational algorithms, this requires using reduced-order model approximation for estimation and control.

Several observers have been proposed for the NS equation, see for example [4]–[6]. For the Boussinesq equation, far fewer estimation results are available due to the presence of a coupling nonlinearity between the NS equation and the thermal equation; this nonlinearity makes the estimation problem more challenging. In [7], the authors proposed a

learning-based robust observer design for the 2D Boussinesq equation under model parametric uncertainties. They proved that the closed loop system for the observer error state satisfies an estimate of \mathcal{L}_2 norm in a sense of locally input-to-state stability (LISS) with respect to parameter uncertainties. Then they proposed to learn the uncertain parameters estimate using a data-driven extremum seeking (ES) algorithm. In [8], the authors introduced a method for designing robust, proper orthogonal decomposition (POD)-based, low-order observers for a class of spectral infinite-dimensional nonlinear systems, with application to the 2D Boussinesq equation. Robustness to bounded model uncertainties was incorporated using the Lyapunov reconstruction approach from robust control theory. The gains of the observer were optimized online using a data-driven learning approach.

In this paper, we adopt the method of DMDc to model such complex and high-dimensional systems. The obtained model is then used to estimate the airflow and the temperature in whole space, based on localized sensing points. Due to inherent uncertainties in the obtained data-driven model, originating from uncertain boundary conditions and unknown disturbances on the wall, we propose to use robust observers for estimation. Robust estimators have proven useful not only in providing robustness against uncertainties but in estimating dynamical information without strong statistical assumptions [9], [10]. Finally, we test the obtained DMD-based model and the associated robust observers on a challenging turbulent flow case.

II. PROBLEM STATEMENT

We consider the problem of data-driven modeling and robust estimation of the airflow velocity and temperature in a built environment based on a two-dimensional turbulent flow. Fig. 1 illustrates a schematic of a representative case study in a built environment along with the corresponding boundary conditions. Specifically, cold air of larger density enters the room through a distributed inlet at the top vent Γ_i , subject to Dirichlet boundary condition of prescribed inlet velocity and temperature and leaves the room through the exhaust located at the top left corner of the domain Γ_o , subject to zero-Neumann boundary condition. The exterior serves as the heat load of the domain at boundaries $\Gamma_{w,1} - \Gamma_{w,4}$. In practice, the wall temperature is not known and is modeled as a disturbance input as demonstrated in the following.

The governing equations are turbulent transient Boussinesq equations, which are incompressible Navier-Stokes equations

* Corresponding author. Email: nabi@merl.com, Phone: +1 (617) 621-7589.

¹Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA. This work was completed during an internship at MERL.

²Control and Dynamical Systems Group, Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA.

³Department of Mechanical and Materials Engineering, University of Nebraska-Lincoln, NE 68588, USA.

⁴Data Analytics Group, Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA.

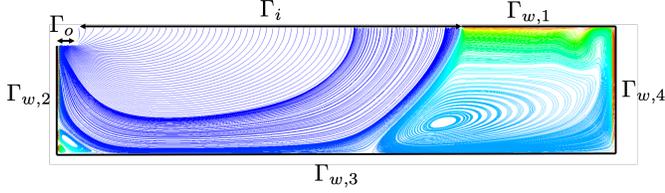


Fig. 1. Exemplar solution of the 2-D Boussinesq equation in a representative built environment. Streamlines of velocity field superposed by temperature pseudo-color at long-time simulation (from red to blue, the temperature varies from hot to cold).

coupled with the heat transfer equations, and are given below, using Einstein notation

$$0 = \frac{\partial u_j}{\partial x_j}, \quad (1a)$$

$$0 = \frac{\partial u_i}{\partial \tau} + \frac{\partial u_i u_j}{\partial x_j} + \frac{\partial p_i}{\partial x_i} - Ri \delta_{i2} \theta - \frac{\partial}{\partial x_j} \left(\frac{1}{Re} \frac{\partial u_i}{\partial x_j} \right), \quad (1b)$$

$$0 = \frac{\partial \theta}{\partial \tau} + \frac{\partial u_j \theta}{\partial x_j} - \frac{\partial}{\partial x_j} \left(\frac{1}{Pe} \frac{\partial \theta}{\partial x_j} \right), \quad (1c)$$

with \mathbf{u}, p, θ denoting scaled ensemble-averaged velocity, pressure, and temperature; δ_{ij} is the Kronecker delta. Non-dimensional numbers in (1) are given by

$$Re = \frac{u_{in} L_{in}}{\nu_{eff}}, \quad Pe = \frac{u_{in} L_{in}}{\kappa_{eff}}, \quad Ri = \frac{u_{in} \theta_{in} L_{in}}{V_{in}^2},$$

where subscript ‘in’ denotes the values at the inlet. The effective viscosity and diffusivity are ν_{eff} and κ_{eff} , respectively, and are computed by unsteady Reynolds-averaged Navier Stokes (URANS) approach with $k-\epsilon$ closure model. A finite-volume solution to Eq. 1 with appropriate boundary conditions in turbulent flow regime has been extensively validated in previous studies, e.g. [11]–[13] for both steady and transient scenarios.

Robust estimation strategies typically require knowledge of a state-space model for the underlying dynamics of the flow. A state-space of representation of (1) needs to express the PDEs in the form of an operator. Unfortunately, the system (1) is highly nonlinear. Furthermore, the lack of explicit dynamical characterizations of pressure necessitates that the pressure term has to be removed for operator extraction. While eliminating pressure has been investigated in laminar and linearized flows [14], this task is extremely difficult when the flow is turbulent due to numerical challenges. Dynamic mode decomposition methods provides a data-driven framework for obtaining state-space representations of highly-nonlinear large-scale dynamical systems [15], [16] such as turbulent flows as in (1); this is discussed next.

III. DATA-DRIVEN MODELING VIA DMD

A. Motivation

Dynamic mode decomposition (DMD) can be thought of as a combination of proper orthogonal decomposition (POD) with Fourier transforms in time [15]. We consider an ordered time-series of snapshots $\{\mathbf{x}_k, \mathbf{u}_k\}_{k=0}^m$ obtained

from a thermo-fluid system, where k is a time index. The k -th snapshot is comprised of column-vectors of measured outputs $\mathbf{x}_k \in \mathbb{R}^{n_x}$ and inputs $\mathbf{u}_k \in \mathbb{R}^{n_u}$, and $(m+1) \in \mathbb{N}$ denotes the total number of snapshots. With these data points, one can construct a discrete-time model of the form

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \quad (2)$$

such that the approximation error

$$\|\mathbf{x}_{k+1} - \mathbf{A}\mathbf{x}_k - \mathbf{B}\mathbf{u}_k\|_F \quad (3)$$

is minimized over all snapshots for $k = 1, \dots, m$. In (2), the system matrix \mathbf{A} is of size $n_x \times n_x$ and the input matrix \mathbf{B} is of size $n_x \times n_u$.

The model (2) may not always be amenable to controller or estimator design, for example, when n_x is large. This is typically the case in PDE systems due to the resulting large dimension incurred by discretization in space. Therefore, it becomes necessary to design models that provide good fits to the snapshot data and resides in a lower-dimensional state-space $r_x \leq n_x$. To this end, we will construct a reduced-order model of the form

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{A}}\hat{\mathbf{x}}_k + \hat{\mathbf{B}}\hat{\mathbf{u}}_k, \quad (4)$$

where $\hat{\mathbf{x}} \in \mathbb{R}^{r_x}$, $\hat{\mathbf{u}} \in \mathbb{R}^{n_u}$, are reduced-order states and inputs, respectively. The reduced-order system matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ have dimensions $r_x \times r_x$ and $r_x \times n_u$, respectively.

B. DMD With Control (DMDc)

In order to construct the model (4), we need to define a few quantities. Let

$$\begin{aligned} \mathbf{X}_+ &= [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_m], \\ \mathbf{X} &= [\mathbf{x}_0 \quad \mathbf{x}_1 \quad \cdots \quad \mathbf{x}_{m-1}], \\ \mathbf{U} &= [\mathbf{u}_0 \quad \mathbf{u}_1 \quad \cdots \quad \mathbf{u}_{m-1}] \end{aligned}$$

denote data matrices constructed using the available snapshots. By construction of these matrices, and by the recurrence relation (2), we obtain

$$\mathbf{X}_+ = \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{U}, \quad (5)$$

which can be rewritten as

$$\mathbf{X}_+ = [\mathbf{A} \quad \mathbf{B}] \begin{bmatrix} \mathbf{X} \\ \mathbf{U} \end{bmatrix} =: \bar{\mathbf{A}}\bar{\mathbf{X}}. \quad (6)$$

Note that the augmented matrix $\bar{\mathbf{X}}$ is of size $(n_x + n_u) \times m$. We deduce from (5) that minimizing the approximation error (3) is tantamount to minimizing $\|\mathbf{X}_+ - \bar{\mathbf{A}}\bar{\mathbf{X}}\|_F$.

To this end, we employ DMDc, as reported in [17], with unknown state and input matrices. We begin by taking a truncated SVD of $\bar{\mathbf{X}}$ up to a truncation value of $r' > r_x$, that is $\bar{\mathbf{X}} \approx \bar{\mathbf{U}}\bar{\mathbf{\Sigma}}\bar{\mathbf{V}}^\top$, where $\bar{\mathbf{\Sigma}}$ has r' non-zero diagonal entries. This yields

$$\bar{\mathbf{A}} \approx \mathbf{X}_+ \bar{\mathbf{V}}\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{U}}^\top.$$

This can be partitioned into the state and input matrices \mathbf{A} and \mathbf{B} described in (2) as follows:

$$[\mathbf{A} \quad \mathbf{B}] = [\mathbf{X}_+ \bar{\mathbf{V}}\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{U}}_1^\top \quad \mathbf{X}_+ \bar{\mathbf{V}}\bar{\mathbf{\Sigma}}^{-1}\bar{\mathbf{U}}_2^\top], \quad (7)$$

where $\bar{\mathbf{U}}_1 \in \mathbb{R}^{n_x \times r'}$ and $\bar{\mathbf{U}}_2 \in \mathbb{R}^{n_u \times r'}$.

As explained earlier, the ambient state space n_x may be prohibitively large for estimator and controller design. Therefore, we need an additional projection step to bring this state-dimension down to r_x . Unlike conventional DMD, we cannot use $\bar{\mathbf{U}}$ to find the projection subspace basis because $\bar{\mathbf{U}}$ contains both state and input data. Instead, we find the basis from the pure state data matrix \mathbf{X}_+ . This involves computing another truncated SVD, that is,

$$\mathbf{X}_+ = \hat{\mathbf{U}}\hat{\Sigma}\hat{\mathbf{V}}^\top,$$

where $\hat{\Sigma}$ has r_x non-zero diagonal elements due to truncation.

The projected state is then given by

$$\hat{\mathbf{x}} := \hat{\mathbf{U}}^\top \mathbf{x}. \quad (8a)$$

Consequently, the reduced-order system matrices are computed using

$$\hat{\mathbf{A}} = \hat{\mathbf{U}}^\top \bar{\mathbf{A}} \hat{\mathbf{U}} = \hat{\mathbf{U}}^\top \mathbf{X}_+ \bar{\mathbf{V}} \bar{\Sigma}^{-1} \bar{\mathbf{U}}_1^\top \hat{\mathbf{U}}, \quad (8b)$$

$$\hat{\mathbf{B}} = \hat{\mathbf{U}}^\top \bar{\mathbf{B}} = \hat{\mathbf{U}}^\top \mathbf{X}_+ \bar{\mathbf{V}} \bar{\Sigma}^{-1} \bar{\mathbf{U}}_2^\top. \quad (8c)$$

Therefore, the equations (8), along with $\hat{\mathbf{u}} := \mathbf{u}$ yields the desired reduced-order model (4). The complete pseudocode is provided in Algorithm 1 for convenience.

Algorithm 1 DMD with control, $\text{DMDc}(\bar{\mathbf{X}}_+, \bar{\mathbf{X}}, r_x)$

Require: Data, $\mathbf{X}_+, \bar{\mathbf{X}}$

Require: Target state dimension, r_x

Require: Truncation value, $r' > r_x$

Require: Truncated SVD subroutine, $[\mathbf{U}, \Sigma, \mathbf{V}] = \text{svd}(\cdot, \cdot)$

$[\bar{\mathbf{U}}, \sim, \sim] = \text{svd}(\bar{\mathbf{X}}, r')$

Partition $\bar{\mathbf{U}} = [\bar{\mathbf{U}}_1 \quad \bar{\mathbf{U}}_2]$

Compute $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ as in (7)

$[\hat{\mathbf{U}}, \sim, \sim] = \text{svd}(\mathbf{X}_+, r_x)$

Compute $\hat{\mathbf{A}}$ using (8b)

Compute $\hat{\mathbf{B}}$ using (8c)

Remark 1. The target state dimension r_x is a design parameter. One heuristic that is commonly used to inform the selection of r_x is by generating the spectrum of \mathbf{X}_+ and selecting r_x based on the cumulative concentration of DMD modes, for example, by only considering singular values that contribute 99% of the variance. ■

C. Randomized DMD for big data

When the number of snapshots m is large, the SVD operation becomes prohibitive. In such cases, random projections may be used to systematically reduce the ambient space of the data such that a subsequent SVD on the projected data, up to a target (usually much smaller than m) rank, is near-optimal compared to the true SVD. Thus, carefully selecting the random projection matrices can significantly reduce the computational expenditure of the SVD procedure. Randomized methods for efficient DMD without control has been studied in [18], [19].

IV. ROBUST ESTIMATION

In this section, we discuss ways in which we can use the model (4) to synthesize estimators that are robust to model uncertainties or disturbances.

We begin by converting the discrete-time system (4) into continuous-time and modeling the uncertainties acting on the system dynamics. The continuous-time model is given by

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = \hat{\mathbf{A}}_c \hat{\mathbf{x}}(t) + \hat{\mathbf{B}}_c \hat{\mathbf{u}}(t) + \mathbf{w}(t), \quad (9a)$$

$$\mathbf{y}(t) = \mathbf{C} \hat{\mathbf{U}} \hat{\mathbf{x}}(t) + \mathbf{v}(t) \quad (9b)$$

where $\mathbf{y} \in \mathbb{R}^{n_y}$ is the measured output, $\hat{\mathbf{A}}_c$ is an $r_x \times r_x$ matrix and $\hat{\mathbf{B}}_c$ is an $r_x \times n_u$ matrix. The process noise acting on the system is represented by $\mathbf{w} \in \mathbb{R}^{r_x \times 1}$ and is a manifestation of various sources of noise, discussed in the next subsection. The $\mathbf{v} \in \mathbb{R}^{n_y \times 1}$ is the measurement noise of the sensor. We assume that process and measurement noise are $\mathcal{L}_2[0, \infty)$. The matrix $\mathbf{C} \in \mathbb{R}^{n_y \times r_x}$ is the output matrix. For our specific application, \mathbf{C} is given, and $n_y = 1$. We define $\hat{\mathbf{C}} = \mathbf{C} \hat{\mathbf{U}}$.

A. Need for robust estimation

We elucidate upon some sources of uncertainty arising in built environments. As discussed in Section II, the exterior temperature at boundaries $\bigcup_{k=1}^4 \Gamma_{w,k}$ are essential to solution of the system (1). Such temperatures dictate heat transfer to the system, which is also a function of thermal and physical properties of the wall, such as the thickness of the wall and the solar gain. In absence of any physical model, one could assume a known (*e.g.* Gaussian) distribution for the boundary condition at the walls. However, in practice, the distribution of such disturbances are not usually distributed according to known stochastic characterizations and can be completely unknown and unanticipated at design time.

B. Steady-state Kalman filter

As a first step in this analysis, we consider the continuous-time Kalman filter, also known as the Kalman-Bucy filter, see for example [20], [21]. This filter requires numerical integration at every time step which is impractical. For this specific application, the error covariance reaches a steady-state value quickly. Thus, we use the steady-state counterpart of the filter, given by,

$$\frac{d\mathbf{z}(t)}{dt} = \hat{\mathbf{A}}_c \mathbf{z}(t) + \hat{\mathbf{B}}_c \hat{\mathbf{u}}(t) + \mathbf{L}(\mathbf{y}(t) - \hat{\mathbf{C}}\mathbf{z}(t)) \quad (10)$$

where $\mathbf{L} = \mathbf{P} \hat{\mathbf{C}}^\top \mathbf{R}_v^{-1}$ and \mathbf{z} is an estimate of the state $\hat{\mathbf{x}}$ in (9). Here, the matrix $\mathbf{P} = \mathbf{P}^\top \succ 0$ is the solution to

$$\hat{\mathbf{A}}_c \mathbf{P} + \mathbf{P} \hat{\mathbf{A}}_c^\top - \mathbf{P} \hat{\mathbf{C}}^\top \mathbf{R}_v^{-1} \hat{\mathbf{C}} \mathbf{P} + \mathbf{R}_w = 0,$$

where \mathbf{R}_w and \mathbf{R}_v are the covariance matrices of w and v , respectively.

C. \mathcal{H}_∞ observer

A critical assumption required in the derivation of the Kalman filter is that both the process and measurement noise are drawn from distributions with known sufficient statistics

(e.g. Gaussian with known mean and covariance matrix). This assumption does not hold in many practical scenarios, especially with humans-in-the-loop. In the next subsection, we address this issue by designing robust observers that minimize the energy gain of the system to $\mathbf{w}, \mathbf{v} \in \mathcal{L}_2$.

Concretely, the \mathcal{H}_∞ observer differs from the Kalman filter in two different ways (i) unknown disturbances of finite energy replace the white-noise processes that drive the system and corrupt the observations; and, (ii) the aim of the observer is to ensure that the maximal energy gain from the disturbances to the estimation error is less than a pre-specified performance level γ .

For notational simplicity, we define \mathbf{d} as,

$$\mathbf{d} = [\mathbf{w}^\top \quad \mathbf{v}^\top]^\top. \quad (11)$$

Let $\mathcal{G}_e : \mathbf{d} \mapsto \mathbf{z} - \hat{\mathbf{x}}$ denotes a transfer function from the disturbance input to the state estimation error. We seek an asymptotic estimator gain \mathbf{L} such that \mathcal{G}_e is stable and satisfies $\|\mathcal{G}_e\|_\infty < \gamma$. The following theorem [22] gives us a tractable formulation to design an \mathcal{H}_∞ observer with pre-specified performance level.

Theorem 1. *Suppose the pair $(\hat{\mathbf{A}}, \hat{\mathbf{C}})$ is observable. There exists an observer of the form (10) with gain \mathbf{L} such that the system $\mathcal{G}_e : \mathbf{d} \mapsto (\mathbf{z} - \hat{\mathbf{x}})$ is stable and satisfies $\|\mathcal{G}_e\|_\infty < \gamma$ if and only if there exist matrices $\mathbf{P} \succ 0$ and $\mathbf{Q} \succ 0$ such that*

$$\hat{\mathbf{A}}_c \mathbf{P} + \mathbf{P} \hat{\mathbf{A}}_c^\top - \mathbf{P}(\hat{\mathbf{C}}^\top \hat{\mathbf{C}} - \gamma^{-2} \mathbf{I})\mathbf{P} + \mathbf{Q} = 0 \quad (12)$$

and $\hat{\mathbf{A}}_c - \mathbf{P}(\hat{\mathbf{C}}^\top \hat{\mathbf{C}} - \gamma^{-2} \mathbf{I})$ is asymptotically stable.

The \mathcal{H}_∞ observer dynamics have the same form as the Kalman filter, described by (10). The \mathcal{H}_∞ observer gain, however, is different, and given by $\mathbf{L} = \mathbf{P} \hat{\mathbf{C}}^\top$. Solving (12) implies that the energy gain is bounded in an \mathcal{L}_2 sense, that is

$$\|\mathcal{G}_e\|_\infty = \sup_{\mathbf{d} \neq 0} \frac{\|\mathbf{z}(\cdot) - \hat{\mathbf{x}}(\cdot)\|_{\mathcal{L}_2}}{\|\mathbf{d}(\cdot)\|_{\mathcal{L}_2}} < \gamma.$$

Remark 2. If there is no solution to (12) for any $\gamma > 0$, then no \mathcal{H}_∞ estimator exists. To reduce the conservativeness of the solution, one can solve (12) while performing a line search for γ . ■

V. RESULTS AND DISCUSSION

In this section, we discuss results obtained for modeling and estimation of 2D turbulent Boussinesq equations. First, we investigate exact and randomized DMDC for a sufficiently rich training dataset. Next, this trained linear model is validated against various testing datasets associated with different initial conditions and parameters. This model is then used to construct estimators for Boussinesq equations. Emphasis is given to estimation in the presence of disturbances with unknown statistical properties.

In the ensuing discussion, we select the sensor to measure only temperature; therefore, the velocity is estimated purely from the temperature measurements. We use a single sensor only, the location of which is based on optimal experiment design, which is performed offline. In practice, the location

of temperature sensor is mostly prescribed due to geometric and practical constraints, and is typically not a design choice: conditions that we replicate here.

A. Results for DMDC

We identify the system matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ using DMDC described in Section III. The snapshots are generated using computational fluid dynamics (CFD) simulation software via the OpenFOAM platform. The simulation time for a typical case, consisting of 39600 finite volumes, is 26742 sec on a compute node with 10 CPUs each with the maximum clock speed of 3.00 GHz. We divide the data into training and testing sets; to make the problem harder, we choose one training set and multiple test sets in the locality of the training data. The training set is generated with initial conditions 251 K and 0 m/s to replicate a common real-world condition for quiescent rooms. The sampling time for collecting the snapshots is chosen to be 0.5 s, in accordance with the Nyquist-Shannon sampling theorem. Furthermore, we verify via numerical simulations that such a sampling time is much smaller than the initial transients for both temperature and velocity fields to ensure that the collected snapshots capture the relevant (e.g. fast velocity) dynamics. We set the time horizon of simulations to be long enough time to capture all the dominant phenomena. The CFD simulations illustrate that a time horizon of 1500 s is well beyond the settling times of the relevant dynamics. The snapshots from CFD simulations are then transferred to MATLAB for further analysis using an intermediate C++ software.

Fig. 2 shows the training error for exact and randomized DMDC. We note that the dimension of the full-order system, $n_x = 118800$. Both exact and randomized DMDC show relative errors of order 10^{-3} which confirms that the model is a good fit on the training set. As expected, we see a significant decrease in the computational time required to compute $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ using randomized DMDC, compared with DMDC, for the same target rank of $r_x = 100$ states: the decrease is $\approx 100\times$. It remains to show that the model does not over-fit, that is, we have good generalization capabilities.

Since the system of interest is nonlinear, we need to ensure that our model performs well when the initial conditions are local, but not composed heavily of training set data. Concretely, we validate the model against CFD simulated trajectories generated from different initial conditions. Figure 3 shows the relative error for three unique initial conditions chosen either randomly within 10% of the initial conditions for the training set or initial conditions that are solutions to the 2-D Boussinesq equations. The latter is achieved by first solving (1) with given input and then at the end of time horizon, for which steady state is reached, we set the input to zero and let the velocity and temperature evolve freely in the domain for a given time t_{off} . Physically, such initial conditions mimic the environment in the room when the air conditioner is off for a certain period of time and then it is turned on. As shown in Fig. 3, for these three initial cases, the relative error is consistently small (less than 0.25), demonstrating the potential of our proposed

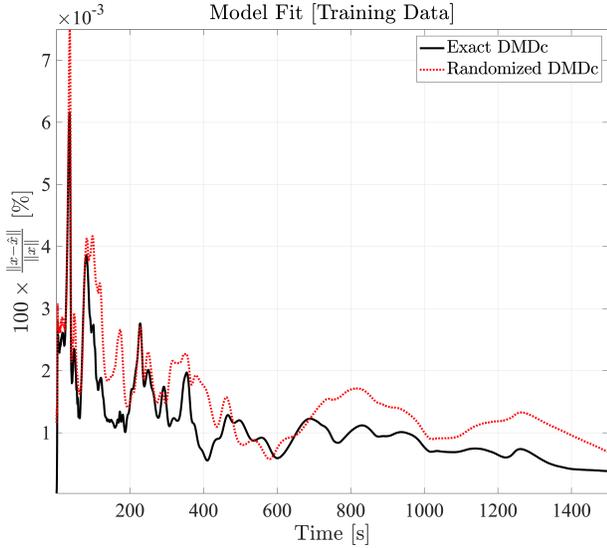


Fig. 2. Relative training error % over 1500 s for models constructed using exact (black continuous line) and randomized DMD (red dashed line).

system identification approach for locally reconstructing the dynamics for the challenging case of a turbulent flow.

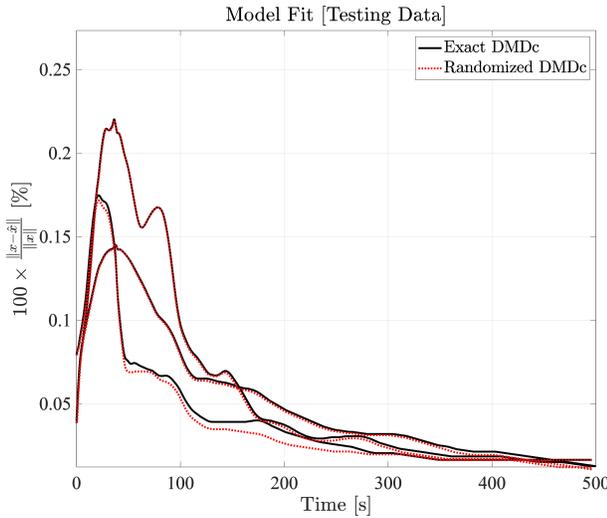


Fig. 3. Relative testing error % for the exact and randomized DMDc models against CFD simulations generated from multiple unique initial conditions.

The Reynolds number, Re , is a non-dimensional scalar that quantifies the impact of input, that is, the air conditioner velocity, on the built environment. To assess the robustness of the identified model to changes in input conditions, we compare the exact DMDc predictions with that of CFD corresponding to Re values that are different from the trained data. Fig. 4 shows that the testing error is reasonably small when the Reynolds number is either increased by 10% or 15% more than the nominal value. When the Reynolds number is deviated significantly, e.g. $\geq 50\%$ compared to Re used in the trained data, the reconstruction error becomes large and unbounded, that is, it increases with time. This is justified because the DMDc is a local model approximation

and more data needs to be used for DMDc constructed in order to extrapolate to Re far from the training set.

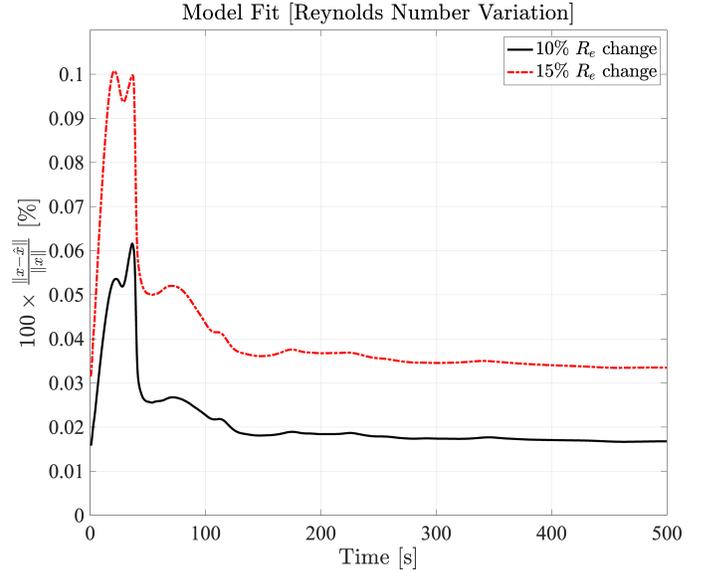


Fig. 4. Relative testing error % for various Reynolds numbers.

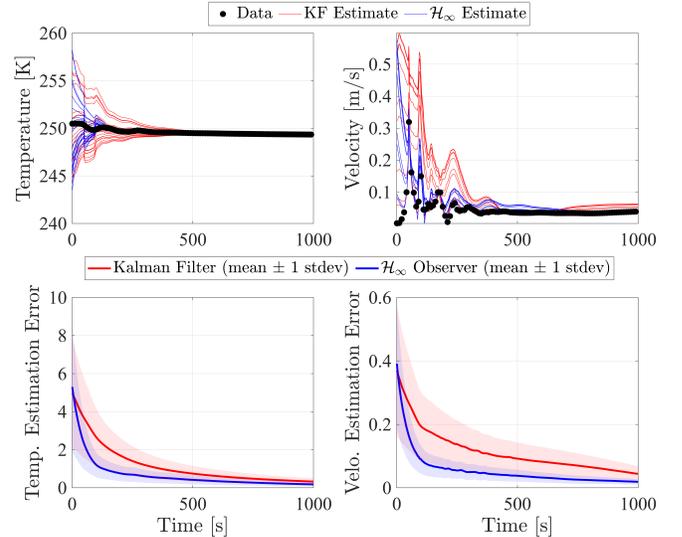


Fig. 5. Comparison between Kalman and \mathcal{H}_∞ filters for reconstruction of temperature and velocity in presence of disturbances on the wall from multiple initial conditions for the estimators. [A] Temperature estimates at a fixed location in the environment. [B] Velocity magnitude estimates at a fixed location in the environment. [C] Error norm of temperature over all locations in the environment. [D] Error norm of velocity magnitude over all locations in the environment.

B. Results for Robust Estimation

We now present our results on robust estimation using the identified system in the presence of uncertainties. As before, for generating output data the measured output y , we use CFD simulations from the case with initial condition obtained after a long t_{off} . Figure 5 shows the estimated temperature and velocity at one point in the built environment when the estimator is initialized with 20 random initial conditions such

that the estimated temperature and velocity initial conditions are extracted from the interval [242 K, 260 K] and [0.03 m/s, 0.57 m/s], respectively; such a range is architecturally relevant for ventilation in buildings exemplified by Fig. 1. The uncertainties w and v are generated with uniform random noise seeds with finite support; thus, w and v are bounded and have no useful stochastic prior information associated with them.

We compare the performance of two robust estimators in Fig. 5. The subplots [A] and [B] show the effect of robust estimation on temperature and velocity at a location in the environment, as mentioned before. It is clear that the \mathcal{H}_∞ observer converges faster to a small neighborhood of the data than the Kalman filter. In fact, for the velocity estimation problem, the Kalman filter does not enter a 5% settling zone even after 800 s. Conversely, the \mathcal{H}_∞ observer exhibits better estimation properties after 500 s. The mean and standard deviation of the estimation errors of the two robust estimators are shown in the subplots [C] and [D] in Fig. 5. This illustrates further that the rate of decay of the \mathcal{H}_∞ observer to the steady-state offset is faster than the Kalman Filter. Not only is the decay faster, but the standard deviation is tighter for the \mathcal{H}_∞ observer, indicating improved robustness. This behavior is not entirely unexpected: as explained in Section IV-A, since the underlying assumptions made on the Kalman filter do not always conform to uncertainties seen in building systems, whereas an \mathcal{L}_2 assumption on the uncertainty is more common since the uncertainty signals have finite energy and typically vanish after finite time. For example, thermal noise on the walls as an effect of incident solar energy is going to exhibit bounded energy and vanish after sunset and its characteristics depend on the wall whose material composition may be unknown and other temporal factors that are difficult to characterize.

VI. CONCLUSIONS

The problem of estimating fully turbulent flows, described by Boussinesq equations (1) is studied in presence of uncertainties such as disturbance inputs on walls and unmodeled dynamics. Due to the large-scale system at hand, which are in turn result of discretized PDEs, a dynamic mode decomposition method is used for identifying a reduced-order state-space model that enables estimator design. An advantage of our proposed approach is that the model is constructed directly from data, and does not require the direct solution of Boussinesq equations which would make the problem intractable on-line due to scalability issues. Based on the DMDC model, we construct robust estimators and demonstrate the effectiveness of \mathcal{H}_∞ observers over Kalman filters when the underlying sufficient statistics of the disturbance inputs are unknown; this is expected for buoyancy-driven flows since disturbance inputs rarely conform to common probability density functions. Our proposed approach can also be used in a wide range of applications involving large-scale systems that exhibit turbulent flows such as drag reduction in aerospace systems and wind energy systems.

REFERENCES

- [1] J. Borggaard, J. A. Burns, A. Surana, and L. Zietsman, "Control, estimation and optimization of energy efficient buildings," in *Proc. of the American Control Conference (ACC)*, 2009, pp. 837–841.
- [2] Y. Zhou, M. Wang, M. Wang, and Y. Wang, "Predictive accuracy of Boussinesq approximation in opposed mixed convection with a high-temperature heat source inside a building," *Building and Environment*, vol. 144, pp. 349–356, 2018.
- [3] N. Ivanov and M. Zsimeva, "Large eddy simulation of airflow in a test ventilated room," in *Journal of Physics: Conference Series*, vol. 1038, no. 1. IOP Publishing, 2018, p. 012136.
- [4] M. Guay and N. Hariharan, "Airflow velocity estimation in building systems," in *American Control Conference*, 2009, pp. 908–913.
- [5] T. John, M. Guay, N. Hariharan, and S. Narayanan, "POD-based observer for estimation in Navier–Stokes flow," *Computers & Chemical Engineering*, vol. 34, no. 6, pp. 965–975, 2010.
- [6] N. MacKunis, S. Drakunov, M. Reyhanoglu, and L. Ukeiley, "Nonlinear estimation of fluid velocity fields," in *Proc. of the IEEE Conf. on Decision and Control*, 2011, pp. 6931–6935.
- [7] S. Koga, M. Benosman, and J. Borggaard, "Learning-based robust observer design for coupled thermal and fluid systems," in *American Control Conference (ACC)*, July 2019, pp. 941–946.
- [8] M. Benosman and J. Borggaard, "Robust nonlinear state estimation for thermal-fluid models using reduced-order models: The case of the Boussinesq equations," in *Proc. of the IEEE Conf. on Decision and Control*, December 2019.
- [9] A. Chakrabarty, M. J. Corless, G. T. Buzzard, S. H. Żak, and A. E. Rundell, "State and unknown input observers for nonlinear systems with bounded exogenous inputs," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5497–5510, Nov 2017.
- [10] A. Chakrabarty and M. Corless, "Estimating unbounded unknown inputs in nonlinear systems," *Automatica*, vol. 104, pp. 57 – 66, 2019.
- [11] S. Nabi, P. Grover, and P. C.-C. Caulfield, "Adjoint-based optimization of displacement ventilation flow," *Building and Environment*, vol. 124, pp. 342–356, 2017.
- [12] C. Laughman, P. Grover, and S. Nabi, "A study of refrigerant dispersion in occupied spaces under parametric variation," in *Proc. ASHRAE Annual Conference, Atlanta, GA*, 2015.
- [13] S. Nabi, P. Grover, and C. Caulfield, "Nonlinear optimal control strategies for buoyancy-driven flows in the built environment," *Computers & Fluids*, vol. 194, p. 104313, 2019.
- [14] W. Hu, K. Morris, and Y. Zhang, "Sensor location in a controlled thermal fluid," in *Proc. of the IEEE Conf. on Decision and Control (CDC)*. IEEE, 2016, pp. 2259–2264.
- [15] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*. SIAM, 2016.
- [16] A. Narasingam and J. S.-I. Kwon, "Development of local dynamic mode decomposition with control: Application to model predictive control of hydraulic fracturing," *Computers & Chemical Engineering*, vol. 106, pp. 501–511, 2017.
- [17] J. L. Proctor, S. L. Brunton, and J. N. Kutz, "Dynamic mode decomposition with control," *SIAM J. Applied Dynamical Systems*, vol. 15, no. 1, pp. 142–161, 2016.
- [18] N. B. Erichson, L. Mathelin, S. L. Brunton, and J. N. Kutz, "Randomized dynamic mode decomposition," *arXiv preprint:1702.02912*, 2017.
- [19] D. A. Bistrian and I. M. Navon, "Randomized dynamic mode decomposition for nonintrusive reduced order modeling," *Int. J. Numerical Methods in Engineering*, vol. 112, no. 1, pp. 3–25, 2017.
- [20] D. Simon, *Optimal State Estimation: Kalman, \mathcal{H}_∞ , and Nonlinear Approaches*. John Wiley & Sons, 2006.
- [21] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*. Chapman and Hall/CRC, 2011.
- [22] M. Green and D. J. Limebeer, *Linear Robust Control*. Courier Corporation, 2012.