MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Crowd Flow Completion From Partial Spatial Observations Using Kernel DMD

Mansour, Hassan; Benosman, Mouhacine; Huroyan, Vahan

TR2017-085 July 03, 2017

Abstract

In this paper, we address the problem of estimating the total flow of a crowd of pedestrians from spatially limited observations. Our approach relies on identifying a dynamical system regime that characterizes the observed flow in a limited spatial domain by solving for the modes and eigenvalues of the corresponding Koopman operator. We develop a framework where we first approximate the Koopman operator by computing the kernel dynamic mode decomposition (DMD) operator for different flow regimes using fully observed training data. We then pose flow completion as a least squares problem constrained by the one step evolution of the kernel DMD operator. We present numerical experiments with simulated pedestrian flows and demonstrate that the proposed approach succeeds in completing the flow from limited spatial observations.

International Conference on Sampling Theory and Applications (SampTA)

Crowd Flow Completion From Partial Spatial Observations Using Kernel DMD

Hassan Mansour, Mouhacine Benosman
Mitsubishi Electric Research Laboratories
Cambridge, MA 02139
{mansour, benosman}@merl.com

Vahan Huroyan*
University of Minnesota
Minneapolis, MN 55455
huroy002@umn.edu

***Abstract*—In this paper, we address the problem of estimating the total flow of a crowd of pedestrians from spatially limited observations. Our approach relies on identifying a dynamical system regime that characterizes the observed flow in a limited spatial domain by solving for the modes and eigenvalues of the corresponding Koopman operator. We develop a framework where we first approximate the Koopman operator by computing the kernel dynamic mode decomposition (DMD) operator for different flow regimes using fully observed training data. We then pose flow completion as a least squares problem constrained by the one step evolution of the kernel DMD operator. We present numerical experiments with simulated pedestrian flows and demonstrate that the proposed approach succeeds in completing the flow from limited spatial observations.**

***Index Terms*—Dynamic flow completion, Koopman analysis, dynamic mode decomposition, crowd analysis.**

I. Introduction

Dynamic flow completion from limited observation data is a common problem that arises in many fields such as crowd flow estimation, air flow sensing, and weather and ocean current prediction. The problem entails estimating temporally evolving flows on a spatial grid that is not fully observed. The observations are limited since it is infeasible or too costly to have enough sensors to cover the entire grid.

In the context of crowd flow completion, the sensors are cameras that monitor a portion $\mathbf{\Gamma}$ of a wide surveillance area such as an exhibition hall or a large public event, denoted by $\mathbf{\Omega}$, where $\mathbf{\Gamma} \subset \mathbf{\Omega}$. The flows can be the pedestrian velocities $\mathbf{v}$ and densities $\boldsymbol{\rho}$ at all grid locations in $\mathbf{\Omega}$. Given partial observations of the flows on $\mathbf{\Gamma}$, we wish to estimate the complete flows. More formally, let $\boldsymbol{x}_t$ be the vector of complete flow observables at time instance $t \in \{1 \dots T\}$ defined over all $\mathbf{\Omega}$, and let $\boldsymbol{o}_t$ be the the observed entries of $\boldsymbol{x}_t$ restricted to the set $\mathbf{\Gamma}$, i.e. $\boldsymbol{o}_t = \boldsymbol{x}_t|_{\mathbf{\Gamma}} = M\boldsymbol{x}_t$, where $M : \mathbf{\Omega} \rightarrow \mathbf{\Gamma}$ is a spatial mask. Given prior information about the flow of crowds in $\mathbf{\Omega}$ in the form of training data or a function $p(\cdot)$ that estimates the temporal dynamics $\boldsymbol{x}_{t+1} = p(\boldsymbol{x}_t) + \boldsymbol{e}_t$, we address the task of recovering the complete flows $\boldsymbol{x}_t$ from observations $\boldsymbol{o}_t$ by casting it as the following problem

$$\min_{\boldsymbol{x}_t \dots \boldsymbol{x}_T} \sum_{t=1}^{T} \frac{1}{2}\|\boldsymbol{o}_t - M\boldsymbol{x}_t\|_2^2 + \frac{\gamma}{2}\sum_{t=1}^{T-1}\|\boldsymbol{x}_{t+1} - p(\boldsymbol{x}_t)\|_2^2, \tag{1}$$

where γ is a regularization parameter.

The above formulation relates to the inference problem from a partially observed Markov process [1], [2]. In particular, problem (1) resembles a maximum *a posteriori* (MAP) formulation of Kalman filtering under nonlinear dynamics and Gaussian noise assumptions, see e.g. [3]. However, when the function $p(\cdot)$ is not known, data driven methods are needed to estimate the temporal dynamics of the flow observables $\boldsymbol{x}_t$ from training data.

While there has been extensive research in the field of crowd analysis, little has been done in terms of crowd flow completion. The general focus has involved finding dynamical models that govern the flow of a crowd [4]–[8]. Completing a partially observed flow requires solving an inverse problem that identifies the initial and boundary conditions of the partial differential equation (PDE) that governs the flow, as well as the flow parameters [9]. Alternatively, data-driven approaches in the form of dynamic mode decomposition (DMD) and its nonlinear extensions [10]–[13] have been proposed recently, which allow for approximating a system dynamics by estimating its infinite dimensional Koopman operator [14]. Finally, [15] proposed a related numerical procedure to synthesize an observer form for discrete time autonomous nonlinear systems based on Koopman operator theoretic framework.

In this paper, we propose a flow completion framework that utilizes the prediction capabilities of the Koopman operator. We describe in Section II the data-driven kernel DMD method of [13] that approximates

*This work was performed in part while Vahan Huroyan was an intern at MERL.

the Koopman operator which captures nonlinear system dynamics. In Section III we develop our flow completion framework and propose a method for estimating basis coefficients of the flow. For a particular scene $\mathbf{\Omega}$, we first simulate complete flow data using a macroscopic pedestrian flow model with different initializations. Next, we use the simulated data to learn a kernel DMD operator that captures the dynamics experienced in the scene. For every new flow observed in the spatially limited region $\mathbf{\Gamma}$, we cast the flow completion problem as a least squares inverse problem reguraized by the prediction equations of the kernel DMD operator. More precisely, the kernel DMD operator allows us to learn a basis representation of the flows in $\mathbf{\Omega}$. Therefore, instead of estimating the flow observables $\boldsymbol{x}_t$ directly, our problem becomes that of estimating the coefficients of the basis expansion of $\boldsymbol{x}_t$ from incomplete spatial observations $\boldsymbol{o}_t$. Finally, we present numerical results that validate our approach in Section IV and draw conclusions in Section V.

II. Koopman and dynamic mode decomposition

A. The Koopman Operator

Let $(\mathcal{M}, n, \boldsymbol{F})$ be a discrete time dynamical system, where $\mathcal{M} \subseteq \mathbb{R}^N$ is the state space, $n \in \mathbb{N}$ is the time parameter and $\boldsymbol{F} : \mathcal{M} \to \mathcal{M}$ is the system evolution operator. The Koopman operator $\mathcal{K}$ is defined on the space of functions $\mathcal{F}$, where $\mathcal{F} = \{\phi | \phi : \mathcal{M} \to \mathbb{C}\}$, as follows:

$$\mathcal{K}\phi = \phi \circ \boldsymbol{F}. \tag{2}$$

The Koopman operator is linear and infinite dimensional, and is characterized by the triple; eigenvalues, eigenvectors, and eigenfunctions [10], [14]. For vector valued observables $g : \mathcal{M} \to \mathbb{R}^{N_o}$, the Koopman operator also admits Koopman modes. The Koopman operator specifies a new discrete time dynamical system on the function space $(\mathcal{F}, n, \mathcal{K})$. Let $\varphi_k(x), 1 \leq k \leq K$ be the first K eigenfunctions of $\mathcal{K}$. Without loss of generality, let the system variable be $\boldsymbol{x} \in \mathcal{M}$ and assume that the function $g(\boldsymbol{x}) = \boldsymbol{x}$. Then, [12] shows that $g(\boldsymbol{x}) = \boldsymbol{x} = \sum_{k=1}^{K} \boldsymbol{\xi}_k \varphi_k(\boldsymbol{x})$ and the future state $\boldsymbol{F}(\boldsymbol{x})$ can be estimated as

$$\boldsymbol{F}(\boldsymbol{x}) = (\mathcal{K}g)(\boldsymbol{x}) = \sum_{k=1}^{K} \boldsymbol{\xi}_k (\mathcal{K}\varphi_k)(\boldsymbol{x}) = \sum_{k=1}^{K} \lambda_k \boldsymbol{\xi}_k \varphi_k(\boldsymbol{x}), \tag{3}$$

where $\boldsymbol{\xi}_k$ and λ_k are the Koopman modes and Koopman eigenvalues, respectively.

B. Kernel Dynamic Mode Decomposition

Williams et al. [13] proposed the Kernel DMD (KDMD) algorithm as a low complexity method for approximating the Koopman operator. Let $f : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ be a kernel function, and define the following data matrices

$$\hat{\boldsymbol{G}}_{ij} = f(\boldsymbol{x}_i, \boldsymbol{x}_j), \;\; \hat{\boldsymbol{A}}_{ij} = f(\boldsymbol{y}_i, \boldsymbol{x}_j), \tag{4}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{y}_j$ are column vectors of the data sets $\mathcal{X}$ and $\mathcal{Y}$. A rank-r truncated singular value decomposition of the symmetric matrix $\hat{\boldsymbol{G}}$ results in the singular vector matrix $\boldsymbol{Q}$ and the singular value matrix $\boldsymbol{\Sigma}$. The KDMD operator $\hat{\boldsymbol{K}}$ is then computed using

$$\hat{\boldsymbol{K}} = (\boldsymbol{\Sigma}^\dagger \boldsymbol{Q}^T) \hat{\boldsymbol{A}} (\boldsymbol{Q}\boldsymbol{\Sigma}^\dagger). \tag{5}$$

An eigenvalue decomposition of $\hat{\boldsymbol{K}}$ results in the eigenvector matrix $\hat{\boldsymbol{V}}$ and eigenvalue matrix $\boldsymbol{\Lambda}$. It was shown in [13] that $\boldsymbol{\Lambda}$ approximates the Koopman eigenvalues. Moreover, the Koopman eigenfunctions are approximated by the matrix $\boldsymbol{\Psi} = \hat{\boldsymbol{V}}^T \boldsymbol{\Sigma}^T \boldsymbol{Q}^T$. Since every data point $\boldsymbol{x}_i = \sum_k \lambda_k \boldsymbol{\xi}_k \varphi_k$, the Koopman modes are approximated by the matrix $\boldsymbol{\Xi} = \boldsymbol{X}\boldsymbol{\Psi}^\dagger = \boldsymbol{X}\boldsymbol{Q}\boldsymbol{\Sigma}^\dagger \hat{\boldsymbol{V}}^\dagger$, where $\boldsymbol{X} = [\boldsymbol{x}_1 \ldots \boldsymbol{x}_T]$.

Next consider the basis matrix $\boldsymbol{\Phi} = \hat{\boldsymbol{V}}^T \boldsymbol{\Sigma}^\dagger \boldsymbol{Q}^T$. Then, for every new data point $\boldsymbol{x}^*$, [13] shows that the corresponding prediction $\boldsymbol{y}^* \approx \boldsymbol{F}(\boldsymbol{x}^*)$ can be approximated using KDMD by first estimating the eigenfunction

$$\boldsymbol{\varphi}(\boldsymbol{x}^*) = \boldsymbol{\Phi}[f(\boldsymbol{x}^*, \boldsymbol{x}_1), f(\boldsymbol{x}^*, \boldsymbol{x}_2), \ldots, f(\boldsymbol{x}^*, \boldsymbol{x}_T)]^T, \tag{6}$$

and using the Koopman prediction relation

$$\begin{cases} \boldsymbol{x}^* \approx \boldsymbol{\Xi}\boldsymbol{\varphi}(\boldsymbol{x}^*), \\ \boldsymbol{y}^* \approx \boldsymbol{\Xi}\boldsymbol{\Lambda}\boldsymbol{\varphi}(\boldsymbol{x}^*). \end{cases} \tag{7}$$

Consequently, the temporal dynamics of a time series can be predicted starting from an initial observation $\boldsymbol{x}_0^*$ by first computing $\boldsymbol{\varphi}(\boldsymbol{x}_0^*)$, then driving the system using

$$\boldsymbol{x}_t^* = \boldsymbol{\Xi}\boldsymbol{\Lambda}^t \boldsymbol{\varphi}(\boldsymbol{x}_0^*). \tag{8}$$

III. Crowd Flow Completion

We propose a data-driven framework for flow completion where training data defined over the complete scene $\mathbf{\Omega}$ are available to learn a kernel DMD operator that captures the dynamics of flows in the scene. Then, for new test flows defined over a subset $\mathbf{\Gamma} \subset \mathbf{\Omega}$, we solve a least squares minimization problem constrained by the one step prediction of kernel DMD operator.

A. Training data generation

The dynamics of a crowd can be modeled at the micro and macro scales as the motion of particles in a fluid flow. Alternatively, when real video surveillance data that cover all of $\mathbf{\Omega}$ are available, flow information can be extracted by computing video motion vectors.

We consider in this paper the macroscopic scale, namely variations on the Hughes model [16]–[18] which characterizes the flow in terms of the flow density and velocity. We use the finite volume method proposed in [19], [20] to numerically solve the corresponding PDE. We run the model with L different initial conditions for $T+1$ time steps and collect the data as follows. For the l-th initial condition, $1 \leq l \leq L$, we first form the data matrix $\boldsymbol{Z}^l \in \mathbb{R}^{D\times(T+1)}$ by vertically stacking the horizontal and vertical velocities for each spatial grid point in $\boldsymbol{\Omega}$, where the scene grid is of size $N_x \times N_y$ and $D = 2N_xN_y$. Let $\boldsymbol{Z}^l = [\boldsymbol{z}_1^l, \ldots, \boldsymbol{z}_{T+1}^l]$, we rearrange the columns of $\boldsymbol{Z}^l$ to construct current and future training data matrices $\boldsymbol{X}_{tr}^l$ and $\boldsymbol{Y}_{tr}^l$, such that $\boldsymbol{X}_{tr}^l = [\boldsymbol{z}_1^l, \ldots \boldsymbol{z}_T^l]$, $\boldsymbol{Y}_{tr}^l = [\boldsymbol{z}_2^l, \ldots \boldsymbol{z}_{T+1}^l]$. The data matrices associated with each of the L initial conditions are then stacked to form the total training data matrices $\boldsymbol{X}_{tr} = [\boldsymbol{X}_{tr}^1, \ldots, \boldsymbol{X}_{tr}^N]$ and $\boldsymbol{Y}_{tr} = [\boldsymbol{Y}_{tr}^1, \ldots, \boldsymbol{Y}_{tr}^N]$. The kernel DMD operator is then learned by computing the matrices $\boldsymbol{\Xi}$, $\boldsymbol{\Phi}$, and $\boldsymbol{\Lambda}$ as described in Section II.

B. Flow completion

Suppose that we observe the velocities of a new flow $\boldsymbol{X}^* = \{\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_{T'}^*\} \in \mathbb{R}^D$ on the region $\boldsymbol{\Gamma} \subset \boldsymbol{\Omega}$ that shares the dynamics of the training set. Denote by $\boldsymbol{O} = \{\boldsymbol{o}_1, \ldots \boldsymbol{o}_{T'}\}$, $\boldsymbol{O} \subset \mathbb{R}^d$, $d < D$ the sequence of observed velocities over T snapshots in time. Also let $M : \mathbb{R}^D \to \mathbb{R}^d$ be a masking operator that maps from $\boldsymbol{\Omega}$ to $\boldsymbol{\Gamma}$. We represent the masking operator by the $d \times D$ matrix with the same name M. Then $\forall\ 1 \leq t \leq T'$, $\boldsymbol{o}_t = M\boldsymbol{x}_t^*$. Our goal is to estimate $\boldsymbol{x}_t^*$ from the partial observations $\boldsymbol{o}_t$ and using the approximated Koopman modes, eigenvalues, and eigenfunctions as shown in (7).

Since the vectors $\boldsymbol{x}_t^*$ are not available to evaluate the kernel functions $f(\boldsymbol{x}_t^*, \boldsymbol{x}_j),\ \forall j \in \{1 \ldots (T+1)L\}$, we model $\boldsymbol{x}_t^*$ as $\begin{cases} \boldsymbol{x}_t^* = \boldsymbol{\Xi}\boldsymbol{c}_t, \\ \boldsymbol{x}_{t+1}^* = \boldsymbol{\Xi}\boldsymbol{\Lambda}\boldsymbol{c}_t, \end{cases}$ where $\boldsymbol{c}_t \in \mathbf{R}^K$ approximate the kernel synthesis coefficients of $\boldsymbol{x}_t^*$. Hence, the flow completion problem becomes that of estimating the coefficient matrix $\boldsymbol{C} = [\boldsymbol{c}_1 \ldots \boldsymbol{c}_{T'}] \in \mathbb{R}^{K\times T'}$ from the observations $\boldsymbol{O}$.

Therefore, we propose the following optimization problem to recover the matrix $\boldsymbol{C}$:

$$\min_{\boldsymbol{C}} \|\boldsymbol{O} - M\boldsymbol{\Xi}\boldsymbol{C}\|_F^2 + \tfrac{\gamma}{2}\|\boldsymbol{\Lambda}\boldsymbol{C}E_1 - \boldsymbol{C}E_2\|_F^2, \tag{9}$$

where E_1 and E_2 are restriction operators that select columns 1 to $T'-1$ and 2 to T' of $\boldsymbol{C}$, respectively, and γ is a regularization parameter. The regularization term $\frac{\gamma}{2}\|\boldsymbol{\Lambda}\boldsymbol{C}E_1 - \boldsymbol{C}E_2\|_F^2$ ensures that the coefficients $\boldsymbol{c}_t$ behave as the kernel synthesis coefficients by satisfying the one step prediction relation $\boldsymbol{c}_{t+1} \approx \boldsymbol{\Lambda}\boldsymbol{c}_t$.

Problem (9) can now be recast as a simple linear least squares problem in terms of the vector $\bar{\boldsymbol{c}} = \begin{bmatrix} \boldsymbol{c}_1 \\ \vdots \\ \boldsymbol{c}_{T'} \end{bmatrix}$. Denote by the matrix $\boldsymbol{B} = M\boldsymbol{\Xi}$, and let $\boldsymbol{I}$ be the identity matrix of dimension K. The matrix of coefficient vectors $\widetilde{\boldsymbol{C}} := [\tilde{\boldsymbol{c}}_1 \ldots \tilde{\boldsymbol{c}}_{T'}]$ is then computed by solving

$$\widetilde{\boldsymbol{C}} = \arg\min_{\bar{\boldsymbol{c}}} \left\| \begin{bmatrix} \bar{\boldsymbol{o}} \\ \boldsymbol{0} \end{bmatrix} - \begin{bmatrix} \overline{\boldsymbol{B}} \\ \sqrt{\gamma}\overline{\boldsymbol{D}} \end{bmatrix} \bar{\boldsymbol{c}} \right\|_2^2, \tag{10}$$

where $\bar{\boldsymbol{o}}$ is the vector formed by vertically stacking the observations $\boldsymbol{o}_1 \ldots \boldsymbol{o}_{T'}$, $\boldsymbol{0}$ is the vector of zeros, and the matrices $\overline{\boldsymbol{B}}$ and $\overline{\boldsymbol{D}}$ are defined as follows

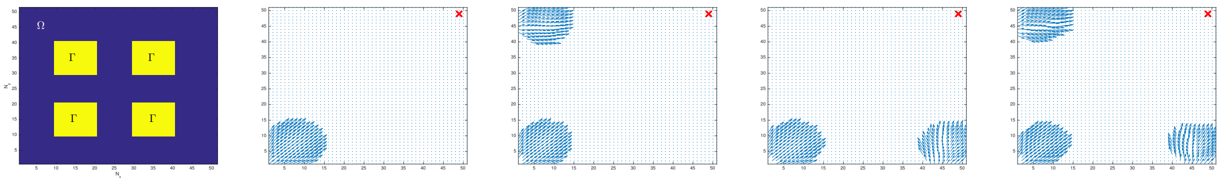
$$\begin{aligned} \overline{\boldsymbol{B}} &= \begin{bmatrix} \boldsymbol{B} & & & \\ & \boldsymbol{B} & & \\ & & \ddots & \\ & & & \boldsymbol{B} \end{bmatrix} \\ \overline{\boldsymbol{D}} &= \begin{bmatrix} \boldsymbol{\Lambda} & -\boldsymbol{I} & & & \\ & \boldsymbol{\Lambda} & -\boldsymbol{I} & & \\ & & \ddots & & \\ & & & \boldsymbol{\Lambda} & -\boldsymbol{I} \end{bmatrix} \end{aligned} \tag{11}$$

Finally, the completed flow $\tilde{\boldsymbol{x}}_t$ can be recovered by multiplying the computed $\widetilde{\boldsymbol{C}}$ with the Koopman mode matrix $\boldsymbol{\Xi}$, i.e. $\tilde{\boldsymbol{x}}_t = \boldsymbol{\Xi}\tilde{\boldsymbol{c}}_t$ for all $t \in \{1 \ldots T'\}$.

IV. Numerical Results and Discussion

We evaluate the performance of our proposed flow completion framework on simulated pedestrian flow data. Let the scene $\boldsymbol{\Omega}$ be a square grid of size 51×51 pixels. The observed section of the flow is represented by the four squares labeled $\boldsymbol{\Gamma}$ in Fig. 1(a). Using the macroscopic pedestrian model [19], [20], we generate three sets of training flow data, each initialized by the three different crowd density configurations shown in Fig. 1(b)–(d). Each flow is composed of 200 time snapshots. We also generate a fourth test flow initialized by the flow distribution in Fig. 1(e). All the training and test flows have the same target of moving to the destination point marked by the red $\mathbf{X}$ in the top right corner of the scene. The test data is also contaminated with additive white Gaussian noise at 20% noise level relative to the noise free signal.

Let M be the mask that selects the velocities in the regions $\boldsymbol{\Gamma}$. Using the training flows, we compute the Kernel DMD matrices $\boldsymbol{\Xi}, \boldsymbol{\Lambda}$, and $\boldsymbol{\Phi}$ using the Gaussian kernel $f(x,y) = e^{-\|x-y\|^2/\|x\|\|y\|}$, and apply our proposed method with $\gamma = 10$ to complete the flows in all of $\boldsymbol{\Omega}$. For performance comparison, we also present



(a) (b) (c) (d) (e)

Fig. 1: (a) Simulation environment showing the total scene $\boldsymbol{\Omega}$ and the observation regions $\boldsymbol{\Gamma}$. (b)–(d) Initial conditions for the training sequences showing the velocities of different groups of pedestrians. (e) Initial conditions of the test sequence. The X in the top right corner marks the target destination of the crowd.

the flow completion result of linear approximation and sparse approximation using the training data as the signal dictionary. For that purpose, we organize the training data into a matrix $\boldsymbol{Z}$ and label the noisy test data by the matrix $\boldsymbol{Z}^*$. With linear approximation and given $\boldsymbol{O} := M\boldsymbol{Z}^*$, we wish to find the coefficient matrix $\boldsymbol{C}$ that minimizes $\frac{1}{2}\|\boldsymbol{O}-M\boldsymbol{Z}\boldsymbol{C}\|_F^2$, i.e. we want to compute $\boldsymbol{C} = (M\boldsymbol{Z})^{-1}\boldsymbol{O}$. The complete flow is then estimated by $\widehat{\boldsymbol{Z}} = \boldsymbol{Z}\boldsymbol{C}$. To compute the matrix inverse $(M\boldsymbol{Z})^{-1}$, we use a truncated SVD with rank 30 since it produced the best recovery performance for this method. For sparse approximation, we find the coefficient matrix $\boldsymbol{C}$ that minimizes the ℓ_1 regularized least squares problem: $\min_{\boldsymbol{C}} \frac{1}{2}\|\boldsymbol{O} - M\boldsymbol{Z}\boldsymbol{C}\|_F^2 + \lambda\|\boldsymbol{C}\|_1$ with $\lambda = 10^{-3}$. We also plot the prediction error from applying the classical KDMD prediction approach of [13] as described in (8). Note that in the KDMD prediction, the initial condition of the complete scene $\boldsymbol{\Omega}$ is needed to evaluate $\varphi(\mathbf{x}_0^*)$. Otherwise, the method fails completely. We plot the performance of the classical KDMD scheme to highlight the best case scenario that can be achieved by this approach. The flow completion performance is illustrated in terms of the relative error $\|\overline{\boldsymbol{Z}}-\widehat{\boldsymbol{Z}}\|_F/\|\overline{\boldsymbol{Z}}\|_F$, where $\widehat{\boldsymbol{Z}}$ is the predicted velocity matrix, and $\overline{\boldsymbol{Z}}$ is the noise-free true velocity matrix.

Fig. 2 shows the change in the relative error over time for each of the above mentioned methods. For a visual assessment, we also show the reconstructed flows at time steps 20, 60, and 140 in Fig. 3. It can be seen that our proposed method results in the smallest relative error especially in the regions where the transient dynamics are active, i.e. time steps 20 through 120. In some cases, the linear and sparse approximations even reverse the flow directions. It is also striking that our proposed method outperforms the KDMD approach which used the complete observation of the initial flow distribution. Finally, we illustrate in Fig. 4 the effect of changing the DMD rank on the reconstruction performance. For comparison with linear approximation, we set the rank of the truncated SVD used in the reconstruction equal

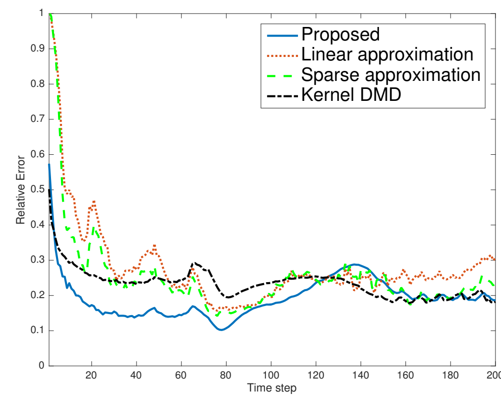
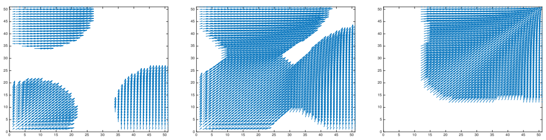


Fig. 2: Relative error after flow completion where the DMD rank is 100 and the observation noise level is 20%.

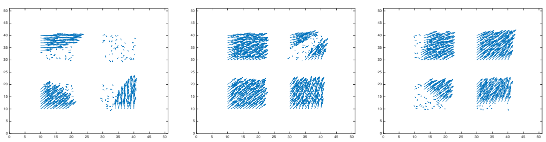
to the DMD rank. The figure shows that our proposed scheme is less sensitive to the choice of the the DMD rank compared to the KDMD approach once it exceeds a certain level, in this case 80. On the other hand, setting the rank too high for the linear approximation results in numerical instability and inclusion of the noise subspace.

V. Conclusion

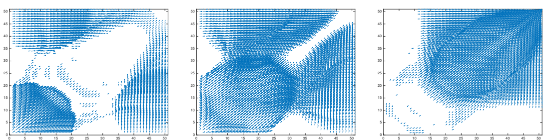
The flow completion framework proposed in this paper is capable of predicting the transient and steady state behavior of a dynamical system from partially observed data. The proposed scheme relies on using complete training data to learn the Koopman operator that captures the dynamics of the system. Flow completion is then performed by solving an inverse problem constrained by the learned Koopman dynamics. While the numerical experiments were limited to the macroscopic crowd flow model, the proposed flow completion framework is model independent and can be applied to a multitude of flow models. Future work will experiment with more complicated crowd flows and real surveillance video data. We will also explore the use of observation control theory using dynamical observers based on the Koopman operator, similar to the investigations in [15].



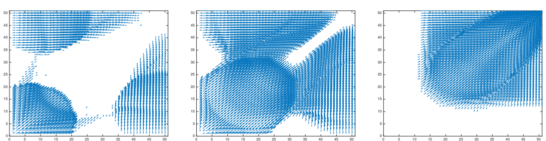
(a)



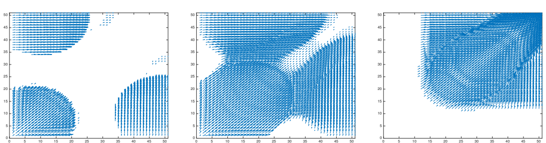
(b)



(c)



(d)



(e)

Fig. 3: Time snapshots 20, 60 and 140 (left to right) of (a) the ground truth flow, (b) the observed flow, and the completed flows using (c) linear approximation, (d) sparse approximation, and (e) the proposed method.

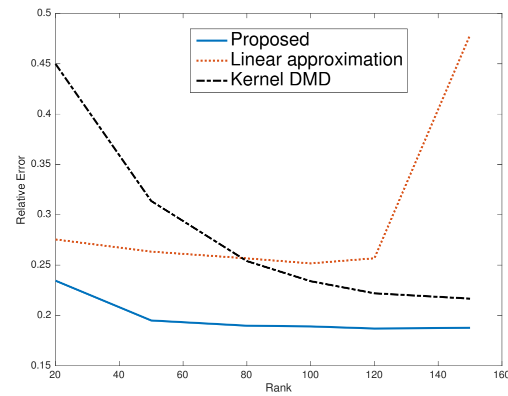


Fig. 4: Relative error after flow completion as a function of the DMD rank. For linear approximation, we set the rank of the truncated SVD equal to the DMD rank.

REFERENCES

[1] E. L. Ionides, C. Bret, and A. A. King, "Inference for nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 103, no. 49, pp. 18438–18443, 2006.

[2] E. L. Ionides, D. Nguyen, Y. Atchad, S. Stoev, and A. A. King, "Inference for dynamic and latent variable models via iterated, perturbed bayes maps," *Proceedings of the National Academy of Sciences*, vol. 112, no. 3, pp. 719–724, 2015.

[3] A. Y. Aravkin, J. V. Burke, and G. Pillonetto, *Optimization Viewpoint on Kalman Smoothing with Applications to Robust and Sparse Estimation*, pp. 237–280, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[4] K. Cao, Y. Chen, D. Stuart, and D. Yue, "Cyber-physical modeling and control of crowd of pedestrians: A review and new framework," *IEEE/CAA Journal of Automatica Sinica*, vol. 2, no. 3, pp. 334, 2015.

[5] S. Wadoo and P. Kachroo, "Feedback control of crowd evacuation in one dimension," *IEEE Transaction on Itelligent Transportation Systems*, vol. 11, no. 1, pp. 182–193, 2010.

[6] H. J. Payne, "Models of freeway traffic and control," in *In Math. Models Publ. Sys. Simul. Council Proc.*, 1971, pp. 51–61.

[7] R. M. Colombo and M. D. Rosinin, "Pedestrian flows and non-classical shocks," *Math. Mthods Appl. Sci.*, vol. 28, no. 13, pp. 1553–1567, 2005.

[8] Molna P. Helbing, D., "Self-organizing phenomena in pedestrian crowds," *ArXiv Condensed Matter e-prints*, pp. 569–577, 1997.

[9] Maria Davidich and Gerta Kster, "Predicting pedestrian flow: A methodology and a proof of concept based on real-life data," *PLOS ONE*, vol. 8, no. 12, pp. 1–11, 12 2013.

[10] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of Fluid Mechanics*, vol. 641, pp. 115–127, 12 2009.

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

[12] M. Williams, I. Kevrekidis, and C. Rowley, "A data–driven approximation of the koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.

[13] M. Williams, C. Rowley, and I. Kevrekidis, "A kernel-based approach to data-driven koopman spectral analysis," *Journal of Computational Dynamics (JCD)*, vol. 2, pp. 247–265, 2015.

[14] I. Mezić and A. Banaszuk, "Comparison of systems with complex behavior," *Physica D Nonlinear Phenomena*, vol. 197, pp. 101–133, Oct. 2004.

[15] Amit Surana and Andrzej Banaszuk, "Linear observer synthesis for nonlinear systems using koopman operator framework," *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 716 – 723, 2016.

[16] R. L. Hughes, "A continuum theory for the flow of pedestrians," *Transportation Research Part B: Methodological*, vol. 36, no. 6, pp. 507 – 535, 2002.

[17] R. L Hughes, "The flow of human crowds," *Annual review of fluid mechanics*, vol. 35, no. 1, pp. 169–182, 2003.

[18] B. D. Greenshields, "A study in highway capacity," in *Highway Research Board*, 1935, vol. 14, 458.

[19] N. Bellomo and C. Dogbé, "On the modelling crowd dynamics from scaling to hyperbolic macroscopic models," *Mathematical Models and Methods in Applied Sciences*, vol. 18, no. supp01, pp. 1317–1345, 2008.

[20] C. Dogbé, "On the numerical solutions of second order macroscopic models of pedestrian flows," *Computers & Mathematics with Applications*, vol. 56, no. 7, pp. 1884–1898, 2008.