Integration of non-time-resolved PIV and time-resolved velocity point sensors for dynamic estimation of time-resolved velocity fields

Jonathan H. Tu,† Clarence W. Rowley†
Princeton University, Princeton, NJ 08544

John Griffin,* Louis Cattafesta‡
Florida Center for Advanced Aero-Propulsion (FCAAP)
University of Florida, Gainesville, FL 32611

Adam Hart,* Lawrence S. Ukeiley†
Florida Center for Advanced Aero-Propulsion (FCAAP)
University of Florida, Shalimar, FL 32579

We demonstrate a three-step method for estimating time-resolved velocity fields using time-resolved point measurements and non-time-resolved particle image velocimetry (PIV) data. First, we use linear stochastic estimation to obtain an initial set of time-resolved estimates of the flow field. These initial estimates are then used to identify a linear model of the flow physics. The model is incorporated into a Kalman smoother, which is used to make a second, improved set of estimates. We verify this method with an experimental study of the wake behind a 7.1% thick elliptical-leading-edge flat plate at a chord Reynolds number of 50,000. Time-resolved PIV data are acquired synchronously with a point measurement of velocity. The above method is then applied to a non-time-resolved subset of the PIV data, along with the time-resolved probe signal. This produces a time-resolved, low-dimensional approximation of the velocity field. The original, time-resolved PIV snapshots are then used to compare the accuracy of the initial, stochastic estimates to the latter, dynamic ones, as measured by the kinetic energy contained in the estimation errors. We find that, for this particular flow, the Kalman smoother is able to utilize the dynamic model and the non-time-resolved PIV data to produce estimates that are not only more accurate, but also more robust to noise. Consequently, modal decompositions based on these estimates more accurately identify coherent structures in the flow.

I. Introduction

Knowledge of the full velocity field can aid greatly in elucidating the fundamental dynamics of a fluid flow. With this information, pertinent spatial structures (modes) can be identified using methods such as proper orthogonal decomposition (POD), balanced proper orthogonal decomposition (BPOD), or dynamic mode decomposition (DMD).1–4 Full-field information may also be required for implementing active flow control or for simply visualizing a flow.5 Unfortunately, time-resolved velocity fields are difficult to obtain.

Particle image velocimetry (PIV) is the standard technique for capturing velocity fields, but time-resolved PIV (TRPIV) systems are costly and thus uncommon. In addition, such systems are generally restricted...
to low-speed flows due to the larger time interval needed between snapshots when using a high-speed laser. The required sampling rates can also limit the spatial extent of the data that can be captured.\textsuperscript{6} As such, typical PIV systems are not time-resolved, and as a result are often incapable of resolving the characteristic frequencies of a flow.

On the other hand, many instruments exist for capturing time-resolved “point” measurements, including hot-wire probes and unsteady pressure sensors. Arrays of such instruments can be used to collect information from various parts of a flow field, but it is not always possible to resolve all the spatial scales of interest. The dense arrays necessary to capture small spatial scales in the flow would likely be too intrusive, and any measurement will be limited by spatial averaging on the scale of the sensor’s size. Furthermore, the resulting measurements can be sensitive to the location of the instruments, which are generally predetermined.

In this work, we demonstrate a three-step method that integrates time-resolved point measurements of velocity, non-time-resolved PIV snapshots, and a flow physics model to estimate the time evolution of a velocity field. As we only wish to resolve the dominant coherent structures, we use proper orthogonal decomposition (POD) to obtain a low-order description of the flow. First, linear stochastic estimation (LSE) is used to compute an initial set of time-resolved estimates of the velocity field. We then form a model of the flow physics by combining an analytic characterization of the flow with a stochastic one identified from the initial estimates. The resulting model is used as the basis for a Kalman smoother, with which a second set of estimates is computed.

Whereas the initial LSE estimates are determined by the point measurements alone, the Kalman smoother also incorporates the non-time-resolved PIV snapshots. These two sets of measurements are used to correct an internal, model-based prediction of the estimate. The dynamics of the model prevent the Kalman smoother estimates from evolving on time scales that are fast with respect to the flow physics, filtering out measurement noise. Thus we can leverage a knowledge of the flow physics and a non-time-resolved description of the velocity field to obtain a more accurate and robust set of estimates.

This approach is fundamentally different from LSE, which does not rely on, nor provide, a model of the flow physics. LSE is only intended to capture those features of the flow that are correlated with the measurement signal. Our approach also differs from the recent work by Legrand et al., in which a phase-averaged description of a velocity field is obtained directly from a large ensemble of PIV data.\textsuperscript{7, 8} Theirs is a post-processing technique that does not make use of any measurement signal, and as such is less amenable to flow control applications.

As a proof of concept, we apply this method in a bluff-body wake experiment. A finite-thickness flat plate with an elliptical leading edge and a bluff trailing edge is placed in a uniform flow, producing oscillatory wake dynamics. We collect TRPIV snapshots, from which we extract the velocity at a single point in the wake, simulating a probe signal. POD modes are computed from the TRPIV data and a set of basis modes is chosen for describing the flow field. The TRPIV snapshots are then down-sampled (in time) and these non-time-resolved data are fed to the dynamic estimator along with the time-resolved probe signal. This generates a time-resolved trajectory for each POD mode coefficient.

The estimation error is quantified using the original TRPIV data. For each TRPIV snapshot, we compute the difference between the estimated POD representation of the velocity field and its projection onto the POD basis. The kinetic energy contained in this difference is then calculated. We collect the values and find the mean value of the error-energy and its distribution. This procedure is then repeated with various levels of artificial noise injected into the probe signal, in order to test each method’s sensitivity to noise. Finally, the estimated flow fields are used to compute DMD modes, testing whether or not the estimates are accurate enough to identify the oscillatory structures in the flow.

The rest of this paper is structured as follows: Sec. II provides a brief introduction to the theory of stochastic and dynamic estimation. These estimation techniques are implemented using the numerical methods detailed in Sec. III and demonstrated using the experiment described in Sec. IV. The results of this experiment are discussed in Sec. V, and conclusions drawn therefrom are presented in Sec. VI.

\section{Background}

\subsection{Stochastic estimation}

Suppose that given some event, we would like to estimate the value of another event. For instance, we may wish to use the velocity measurement at one point in a flow, $u(x)$, to estimate the velocity at another point, $u(x')$. This can be accomplished using stochastic estimation techniques. One such technique is linear stochastic estimation (LSE), which is used in this paper to provide initial estimates of the velocity field. The LSE estimates are then used as the basis for a Kalman smoother, which provides a more accurate and robust set of estimates.

The LSE estimates are determined by the point measurements alone, and do not rely on, nor provide, a model of the flow physics. In contrast, the Kalman smoother incorporates a model of the flow physics, and uses this model to correct the initial estimates. The dynamics of the model prevent the Kalman smoother estimates from evolving on time scales that are fast with respect to the flow physics, filtering out measurement noise. Thus we can leverage a knowledge of the flow physics and a non-time-resolved description of the velocity field to obtain a more accurate and robust set of estimates.

As a proof of concept, we apply this method in a bluff-body wake experiment. A finite-thickness flat plate with an elliptical leading edge and a bluff trailing edge is placed in a uniform flow, producing oscillatory wake dynamics. We collect TRPIV snapshots, from which we extract the velocity at a single point in the wake, simulating a probe signal. POD modes are computed from the TRPIV data and a set of basis modes is chosen for describing the flow field. The TRPIV snapshots are then down-sampled (in time) and these non-time-resolved data are fed to the dynamic estimator along with the time-resolved probe signal. This generates a time-resolved trajectory for each POD mode coefficient.

The estimation error is quantified using the original TRPIV data. For each TRPIV snapshot, we compute the difference between the estimated POD representation of the velocity field and its projection onto the POD basis. The kinetic energy contained in this difference is then calculated. We collect the values and find the mean value of the error-energy and its distribution. This procedure is then repeated with various levels of artificial noise injected into the probe signal, in order to test each method’s sensitivity to noise. Finally, the estimated flow fields are used to compute DMD modes, testing whether or not the estimates are accurate enough to identify the oscillatory structures in the flow.

The rest of this paper is structured as follows: Sec. II provides a brief introduction to the theory of stochastic and dynamic estimation. These estimation techniques are implemented using the numerical methods detailed in Sec. III and demonstrated using the experiment described in Sec. IV. The results of this experiment are discussed in Sec. V, and conclusions drawn therefrom are presented in Sec. VI.
\( u(x') \). The conditional average
\[
\hat{u}_i(x') = \langle u_i(x') | u_i(x) \rangle \tag{1}
\]
provides the expected value of \( u(x') \) given the measurement \( u(x) \), which is the least-mean-square estimate of \( u(x') \).

We can estimate the conditional average by measuring \( u(x') \) repeatedly and averaging over those values that occur whenever \( u(x) \) is near a nominal value \( u^*(x) \).\(^6\) Adrian (1977) proposed an alternative, stochastic estimation technique in which the conditional average is approximated by the power series
\[
\langle u_i(x') | u_i(x) \rangle = A_{ij}(x') u_j(x) + B_{ijk}(x') u_j(x) u_k(x) + \ldots, \tag{2}
\]
where summation over repeated indices is implied.\(^11\) In the case of linear stochastic estimation (LSE), only the linear coefficients \( A_{ij} \) are retained. These can be computed from the two-point, second-order correlation tensor \( R_{ij}(x') \).

Similar procedures exist for higher order estimations, making use of higher-order two-point correlations. While Tung & Adrian found in 1980 that higher-order estimations did not provide much additional accuracy,\(^12\) later studies show that this is not always the case. For instance, quadratic estimation can be more effective when the estimation of a given quantity (e.g., velocity) is based on the measurement of another (e.g., pressure).\(^5\) In particular, for a cavity flow, Murray & Ukeiley (2003) found that while a linear estimate captured the large scale features of the flow field, the turbulent kinetic energy in the quadratic estimate was much more accurate.\(^5\) The quadratic estimate was also able to better predict features of the flow on smaller scales.

Other studies incorporate time offsets into the stochastic estimates, resulting in better performance under the right conditions.\(^10\) Ewing & Citriniti (1999) developed a multi-time LSE technique in the frequency domain that was a significant improvement over single-time LSE.\(^14\) This multi-time formulation also incorporated global analysis tools, namely proper orthogonal decomposition (POD), that yielded low-dimensional representations of the turbulent jet flows being studied.\(^6, 14\) The multi-time approach was translated into the time domain as well, and used for predicting pressure from past measurements of pressure.\(^15\) Durgesh & Naughton (2010) demonstrated the existence of optimal time-delays for the so-called LSE-POD (also referred to as modified LSE) technique, where they estimated the POD mode coefficients of a bluff-body wake in a non-causal, post-processing fashion.\(^16\)

It is important to note that stochastic estimation does not involve any conditional modeling of a system’s dynamics. It simply provides a statistical estimate of a random variable given the knowledge of other random variables.\(^17\) We can think of stochastic estimation as a mapping, computed using a pre-existing dataset, that yields the most statistically likely value of some unknown (conditional) variable, given some other measured (unconditional) data. For a fluid flow, such a method can produce visual representations of the flow field, but cannot suggest, without further analysis, what events should be observed, nor how they might be related to the underlying flow physics.\(^18\) Furthermore, for LSE, the estimated values will lie in a subspace whose dimension is limited to the number of conditions. This is especially important when using a small number of measurements to predict a high-dimensional variable, such as a velocity field. Depending on the application, it can be either a limitation or an advantage, unnecessarily restricting the estimates, or capturing only the features of interest.

B. Dynamic estimation

Dynamic estimators are a foundational topic in control theory. They estimate a system’s state using a model for its dynamics along with real-time measurement updates. The measurement updates are used to correct the trajectory of the model, which will drift from the true trajectory due to parameter uncertainty, unmodeled dynamics, and external disturbances (process noise). This is in contrast to static estimation techniques, such as stochastic estimation, which use a fixed relationship to estimate the system state from a set of measurements. Such an approach does not take advantage of the fact that the equations governing a system’s evolution are often known.

In this work, we focus on the Kalman filter and smoother, both standard subjects in the study of estimation. (For a more in-depth discussion, see any standard text on estimation, for instance the book by Simon.)\(^19\) Suppose we are interested in the evolution of a system described by a linear model
\[
\begin{align*}
\xi_k &= F \xi_{k-1} + d_{k-1} \\
\bar{y}_k &= H_k \xi_k + \bar{w}_k,
\end{align*}
\tag{3}
\]

American Institute of Aeronautics and Astronautics Paper 2012-0033
where $\xi$ is the system state, $\eta$ is a measurement of the state, $d$ represents process noise, and $n$ represents sensor noise. At any given time $k$, we assume that we can observe the measurement $\eta_k$. From this knowledge, we would like to estimate the value of $\xi_k$, which is otherwise unknown.

The dimension of $\eta$ may be smaller than that of $\xi$, meaning that even without sensor noise, the matrix $H$ relating the two may not be invertible. However, if the system is observable, we can use a knowledge of the system dynamics $F$ and the time history of $\eta$ to produce an estimate $\hat{\xi}_k$ that converges, in the case of no noise, to the true value $\xi_k$. In the presence of noise, the Kalman filter will minimize the expected value of the error

$$
(\xi_k - \hat{\xi}_k)^T (\xi_k - \hat{\xi}_k).
$$

The Kalman filter is a causal filter, meaning that only measurements made up to and including time $k$ are available in forming the estimate $\hat{\xi}_k$. In some situations, we may also have access to measurements occurring after time $k$, for instance in a post-processing application. We can use that information to improve our estimate of $\xi_k$. This yields a non-causal filter, generally referred to as a smoother. We will use a variant of the Kalman smoother developed by Rauch, Tung, & Striebel, known as the RTS smoother.\textsuperscript{19} The RTS smoother is a fixed-interval smoother, meaning that all measurements taken over a fixed time interval are used to estimate the state evolution within that interval. Algorithmically, it consists of a forwards pass with a Kalman filter followed by a backwards smoothing pass. The specifics of the Kalman filter and smoother algorithms are described in Sec. III.C.

### III. Numerical methods

#### A. Modal analysis

1. Proper orthogonal decomposition (POD)

Proper orthogonal decomposition (POD), also known as principal component analysis (PCA) or Karhunen-Loeve analysis, is a data analysis method that identifies the dominant structures in a dataset according to their energy content.\textsuperscript{1,20,21} More precisely, suppose we wish to project the dataset \{\xi_j\}\textsubscript{j=0}^m onto an $r$-dimensional subspace. Let $P_r$ be the corresponding projection operator. Then the first $r$ POD modes form the orthogonal basis that minimizes the sum-squared error

$$
\sum_{j=0}^m \|\xi_j - P_r \xi_j\|_2^2.
$$

As such, POD modes are naturally ordered, with a smaller index corresponding to a higher energy content.

When analyzing an incompressible fluid flow, we generally take the data elements to be mean-subtracted velocity fields at given instants in time. These elements $\xi_j = \mathbf{u}'_j = \mathbf{u}'(t_j)$ are commonly referred to as “snapshots.” In this work, the snapshots are collected experimentally using particle image velocimetry (PIV). The POD modes are computed efficiently using the method of snapshots.\textsuperscript{1} Each velocity field is reshaped into a one-dimensional vector, and the vectors are stacked in a data matrix

$$
X = \begin{bmatrix}
\mathbf{u}'_0 \\
\mathbf{u}'_1 \\
\vdots \\
\mathbf{u}'_m
\end{bmatrix}.
$$

The singular value decomposition (SVD) of the correlation matrix $X^TMX$ is then computed,

$$
X^TMX = W\Sigma W^T,
$$

where $M$ is a matrix of inner product weights. For numerical data, this matrix generally contains grid weights, for instance the scaled identity matrix $Idx\,dy$. The inclusion of $M$ allows us to interpret the vector norm as the integrated kinetic energy:

$$
\|\mathbf{u}'_j\|_2^2 = (\mathbf{u}'_j)^T M \mathbf{u}'_j = \iint (u'(t_j)^2 + v'(t_j)^2) \, dx \, dy.
$$
(We note that in this work we measure only two components of velocity, assuming that the third is negligible due to the symmetry of the flow in the spanwise direction.)

The POD modes \( \phi_j \) are then given by the columns of the matrix

\[
\Phi = XW\Sigma^{-1/2}.
\]

They form an orthogonal set, satisfying the identity

\[
\phi_j^T M \phi_i = \delta_{ij},
\]

where \( \delta \) is the Kronecker delta function. As such, the projection of a snapshot \( u_k' \) onto the first \( r \) POD modes is given by

\[
P_r u_k' = \Phi_r \Phi^T M u_k',
\]

where \( \Phi_r \) contains only the first \( r \) columns of \( \Phi \). The projected state is then defined by a vector of POD coefficients \( \mathbf{a} \):

\[
P_r u_k' = \sum_{j=1}^{r} a_j \phi_j = \Phi_r \mathbf{a}.
\]

For a spatially discretized velocity field, the dimension of a POD mode \( \phi_k \) is twice the number of grid points (again assuming we only keep track of two components of velocity). In contrast, \( \mathbf{a} \) has only dimension \( r \).

We observe that due to the orthogonality of the POD modes (see Eq. (7)), the energy in any POD approximation of a velocity field is simply given by

\[
\|P_r u_k'\|_2^2 = \mathbf{a}^T \Phi_r^T M \Phi_r \mathbf{a} = \mathbf{a}^T \mathbf{a} = \|\mathbf{a}\|_2^2.
\]

We emphasize that the first \( r \) POD modes form the orthogonal basis that best captures the kinetic energy in a set of velocity fields. Because we would like our estimators to reproduce velocity fields with as much accuracy as possible, POD modes are a natural choice of basis vectors. However, we note that in flow control applications, other bases may be more suitable, as high energy modes may not always capture the input-output behavior of a system well.\(^2\) In these cases, control-oriented methods such as balanced POD or the eigensystem realization algorithm (ERA) may be advantageous.\(^2,23\)

2. Dynamic mode decomposition (DMD)

Dynamic mode decomposition (DMD) is a snapshot-based method that identifies oscillatory structures in a flow based on their frequency content, as opposed to POD, which identifies modes based on their energy content.\(^3\) When a temporal (as opposed to spatial) analysis is performed, these structures can be interpreted in terms of Koopman operator theory.\(^3\) For a wake flow, which exhibits clear oscillatory behavior, it is natural to apply DMD when trying to identify coherent structures.

To compute the DMD modes from a set of velocity fields \( \{u_j\}_{j=0}^m \), where again \( u_j = u(t_j) \), we form the data matrices

\[
K = \begin{bmatrix} u_0 & \cdots & u_{m-1} \\ \hline u_1 & \cdots & u_m \end{bmatrix}, \quad K' = \begin{bmatrix} u_1 & \cdots & u_m \\ \hline 1 & \cdots & 1 \end{bmatrix}.
\]

(Note that for DMD, in general the mean is not subtracted from the dataset.)

We then compute the SVD \( K = U\Sigma W^T \) using the method of snapshots, taking advantage of the equivalence of left singular vectors and POD modes:

\[
K^T M K = W\Sigma^2 W^T, \quad U = KW\Sigma^{-1},
\]

where \( M \) is again a matrix of grid weights (see the discussion of POD above). These matrices are used to solve the eigenvalue problem

\[
(U^T MK'W\Sigma^{-1})V = V\Lambda,
\]

where the columns of \( V \) and diagonal elements of \( \Lambda \) are the eigenvectors and eigenvalues, respectively, of \( U^T MK'W\Sigma^{-1} \).
The DMD modes $\psi_j$ are then given by the columns of the matrix

$$\Psi = UV,$$

scaled such that

$$\sum_{j=0}^{m-1} \psi_j = \mathbf{u}_0.$$  

With this scaling, the DMD modes are related to the original snapshots by the equations

$$u_k = \sum_{j=0}^{m-1} \lambda_j^k \psi_j, \quad k = 0, \ldots, m - 1$$

$$u_m = \sum_{j=0}^{m-1} \lambda_j^m \psi_j + \xi$$

where the values $\lambda_j$ are the eigenvalues lying on the diagonal of $\Lambda$. Each of these eigenvalues is associated with a particular DMD mode $\psi_j$, giving each mode an associated growth rate $\|\lambda_j\|$ and oscillation frequency $\arg(\lambda_j)$.

### B. Stochastic estimation

Stochastic estimation is a means of approximating a conditional average using a knowledge of unconditional statistics. The conditional mean of event $S$ (the conditional event) given event $P$ (the unconditional event) occurs and is equal to $p$ is written as

$$\langle S|P = p \rangle = \int s \frac{f_{SP}(s, p)}{f_P(p)} ds,$$

where $\langle \cdot \rangle$ is the expected value, $f_{SP}$ denotes the joint probability density function of events $S$ and $P$, and $f_P$ denotes the probability density function of only event $P$.\footnote{Adrian proposed a stochastic estimate of the conditional average by means of a Taylor series expansion.}

$$\hat{s}_i = \langle S_i|P_j = p_j \rangle = A_{ij}p_j + B_{ijk}p_jp_k + \ldots,$$

where $\hat{s}_i$ is the estimate of the conditional average and indicial notation has been included to indicate multiple variables within an event. (This definition of stochastic estimation is more general than Eq. (2) because $s$ and $p$ have not been assigned to fluid-specific quantities (e.g., velocity) and can even represent different variables, with respect to each other.) An estimate of a particular order is set by truncating the series after the corresponding term. The linear and nonlinear coefficients are determined by minimizing the mean square error of the estimate

$$\left\langle (\hat{s}_i - s_i)^2 \right\rangle,$$

which requires solving a set of linear equations.\footnote{1. Linear stochastic estimation (LSE)}

1. **Linear stochastic estimation (LSE)**

In linear stochastic estimation (LSE), only the linear term in Eq. (18) is retained. Focusing specifically on an application to fluid dynamics, we let the velocity field $\mathbf{u}(t)$ be the conditional event ($s = \mathbf{u}$), with the unconditional event $p$ representing a vector of probe measurements. Then the LSE estimate of the velocity field given the value of the probe measurements is

$$\hat{u}_i(t) = A_{ij}p_j(t).$$

(Note that here, the indices $i$ and $j$ describe the spatial dependence of the velocity field and the number of probe measurements, respectively, as opposed to an instant in time.) To compute the coefficients $A_{ij}$, we must minimize the mean-square error of the estimates, which requires solving the equation

$$A^T = [PP]^{-1}[SP],$$

6 of 22
where

\[ A^T = \begin{bmatrix} A_{1,1} & \cdots & A_{1,N_p} \\ \vdots & \ddots & \vdots \\ A_{N_p,1} & \cdots & A_{N_p,N_p} \end{bmatrix}, \quad [PP] = \begin{bmatrix} p_1p_1 & \cdots & p_1p_{N_p} \\ \vdots & \ddots & \vdots \\ p_{N_p}p_1 & \cdots & p_{N_p}p_{N_p} \end{bmatrix}, \quad [SP] = \begin{bmatrix} u_1p_1 \\ \vdots \\ u_{N_p}p_{N_p} \end{bmatrix}, \quad (21) \]

and \( N_p \) is the number of probe measurements. The stochastic estimation procedure can also be applied in the case of multiple conditions, as would be the case for a rake of experimental sensors.\(^{18}\)

This is the classic, or single-time, form of LSE, but several variations exist. For instance, we can introduce a time delay \( \tau \) into the estimate to account for a known lead or lag between the conditional and unconditional variables:\(^{15,16}\)

\[ \hat{u}_i(t) = A_{ij}p_j(t - \tau_j). \quad (22) \]

Accounting for the time delay increases the strength of the correlations between \( \hat{u} \) and \( p \).

We can also account for multiple time delays, summing the correlations over several values of \( \tau.\)\(^{15,16}\) This technique is advantageous if the exact time delay is unknown, not constant, or not resolved well-enough in time. This has been developed for purely negative time delays, requiring only past data,\(^{15}\) as well as for two-sided delays that use probe data taken before and after the estimation time.\(^{16}\) The latter method is applied in this work and is hereafter referred to as multi-time-delay LSE. The use of both past and future data results in stronger correlations, but at the cost of requiring non-causal (future) data. As such, it cannot be used in real-time estimation or flow control applications. (Using a one-sided delay would result in a causal filter, but this is not pursued here as active flow control is outside the scope of this work.) For a derivation of the multi-time-delay LSE algorithm, the reader is referred to the work by Durgesh & Naughton.\(^{16}\)

We note that Eqs. (19) and (22) provide static maps from the measurement \( p \) to the estimate \( \hat{u} \). When computing the coefficients \( A_{ij} \), we make sure to average over large datasets, mitigating the effects of sensor noise. However, in using those coefficients to compute an estimate, the static map will respond directly to the probe measurements (without averaging), making the estimates sensitive to noise. The use of a multi-time-delay formulation reduces these effects somewhat, but cannot completely overcome this inherent limitation of static estimators.

2. Low-dimensional estimation (LSE-POD)

A typical PIV velocity field can consist of thousands of data points. In contrast, for many flows the dominant behavior can be captured by a handful of POD modes. LSE can be applied to estimate the value of the corresponding POD coefficients, producing a low-dimensional estimate of the velocity field. This method is referred to as LSE-POD, and has been applied by Bonnet et al. (1994) and Taylor & Glauser (2004) for linear estimates,\(^{26,27}\) Naguib et al. (2001) and Murray & Ukeiley (2007) for quadratic stochastic estimation,\(^{13,28}\) and Durgesh & Naughton (2010) for multi-time-delay LSE-POD.\(^{16}\) In this manuscript, we take the same low-dimensional approach to stochastic estimation, focusing specifically on multi-time-delay LSE-POD.

C. Dynamic estimation

1. Model identification

Our goal is to use a dynamic estimator to estimate the state a bluff-body wake experiment. We assume that a time-resolved velocity probe signal is available, as well as particle image velocimetry (PIV) velocity fields captured at a slower, non-time-resolved frequency. To implement a dynamic estimator, we need a model for the time-evolution of the system. A high-fidelity numerical discretization of the Navier-Stokes equation is far too computationally intensive for this purpose, and would in any case be difficult to match to the experiment. As such, we develop an empirically-derived, low-order model. We focus on POD-based models, as the first \( r \) POD modes optimally capture the kinetic energy contained in a set of snapshots, for any model order \( r \).

From a large, statistically independent ensemble of PIV snapshots, we can compute a single set of well-converged POD modes. For the model identification procedure, we assume only non-time-resolved data are available. (See Sec. IV.B for a detailed description of the particular dataset used for this computation.) We fix a desired model order \( r \) based on the energy content of the modes, which can be determined from the
POD eigenvalues. For instance, we may choose \( r = 7 \) if it is observed that the first seven modes contain 85% of the energy in the PIV snapshots. These \( r \) modes form a basis for our low-order model.

Due to noise and low spatial resolution, methods such as Galerkin projection can be difficult to apply when using experimentally-acquired velocity fields. As such we take a stochastic approach in identifying a dynamic model. First, we collect a set of non-time-resolved PIV snapshots synchronously with a time-resolved probe signal. The PIV data are projected onto the POD basis to get a non-time-resolved set of POD coefficients \( a \). These coefficients are used along with the probe signal as “training data” to implement multi-time delay LSE-POD, as described above in Sec. III.B. LSE-POD is then applied to the time-resolved probe data to generate time-resolved estimates of the POD coefficients \( \hat{a} \).

We then apply LSE to these vectors, recalling that LSE estimates the expected value of a conditional variable as a linear function of an unconditional variable. If we take \( \hat{a}_k \) to be the conditional variable and \( \hat{a}_{k-1} \) to be the unconditional variable, then we can use LSE to identify a linear, discrete-time dynamical map:

\[
\hat{a}_k = \langle a_k | a_{k-1} \rangle \approx F_{\text{LSE}} a_{k-1}.
\]  

(23)

So long as the LSE-POD estimates of the POD coefficients are accurate enough, then the resulting model will capture enough of the true dynamics to be used as the basis for a Kalman filter.

Finally, we note that it can be shown that the solution to the above LSE problem is the same as the least-squares, least-norm solution to the problem

\[
B = F_{\text{LSE}} A,
\]

where the columns of \( B \) are the vectors \( \{ \hat{a}_j \}_{j=1}^m \) and the columns of \( A \) are the vectors \( \{ \hat{a}_j \}_{j=0}^{m-1} \), collected over all runs. (The proof is simple and omitted here.) As such, \( F_{\text{LSE}} \) is nothing but the Moore-Penrose pseudo-inverse. However, the equivalence to LSE provides an additional interpretation to the dynamics it defines, as it provides a linear estimate of the most statistically likely value of \( \hat{a}_k \) given a value of \( \hat{a}_{k-1} \), according to the ensemble defined by \( A \) and \( B \). Based on this interpretation, this modeling procedure can naturally be extended using quadratic stochastic estimation (QSE), or even higher order methods, for which there are no analogs to the pseudo-inverse.

The bluff-body wake studied in this work is dominated by a Kármán vortex street. A computational study of a very similar flow shows that this behavior is captured well by the first two POD modes alone, which by virtue of their similarity to the dominant DMD modes, have purely oscillatory dynamics.\(^{29}\) To take advantage of this knowledge when developing a model, we decouple the dynamics into two parts: an analytic, oscillatory component describing the Kármán vortex shedding, and a stochastic component describing the dynamics of all the other POD modes. This yields a system with dynamics

\[
\hat{a}_k = \begin{bmatrix} F_{\text{osc}} & 0 \\ 0 & F_{\text{LSE}} \end{bmatrix} \hat{a}_{k-1},
\]  

(24)

where

\[
F_{\text{osc}} = \begin{bmatrix} \lambda^\text{re} & -\lambda^\text{im} \\ \lambda^\text{im} & \lambda^\text{re} \end{bmatrix}.
\]  

(25)

We choose \( \lambda = \lambda^\text{re} + \lambda^\text{im} i \) such that \( \arg(\lambda) \) is equal to the shedding frequency (identified using an autospectrum of the probe signal), and \( \|\lambda\| \) is close to one, indicating nearly perfectly oscillatory dynamics. (In practice we choose \( \|\lambda\| = 0.999 \) to ensure stable dynamics.) The stochastic dynamics \( F_{\text{LSE}} \) are identified using the method discussed above.

2. Kalman filter

We use the procedure detailed in the preceding section to model the bluff-body wake as a dynamical system

\[
\underline{a}_k = F \underline{a}_{k-1} + \underline{d}_k
\]

\[
\underline{\eta}_k = H_k \underline{a}_k + \underline{n}_k,
\]  

(26)

where \( \underline{a} \) is a vector of POD coefficients, \( \underline{\eta} \) is some measured quantity, \( \underline{d} \) is process noise, and \( \underline{n} \) is sensor noise. The measurement matrix \( H_k \) can be varied according to the timestep. At times when a non-time-resolved
PIV snapshot are available, we choose \( H_k = I \), allowing the system access to the POD coefficients of that snapshot. Otherwise, we let \( H_k \) correspond to the velocity probe signal.

We assume that \( d \) and \( n \) are white, zero-mean, and uncorrelated, with covariances \( Q \) and \( R \). This yields the equations

\[
E[d_i d_j^T] = Q_i \delta_{i-j} \\
E[n_i n_j^T] = R_i \delta_{i-j} \\
E[d_i n_j^T] = 0,
\]

where \( \delta \) is the Kronecker delta function. \( Q \) and \( R \) are user-defined matrices, which we can consider to be design parameters. Their relative magnitudes weigh the relative accuracy of the model versus the sensor, and can be used to account for the effects of noise on the system. If we have a very noisy sensor, we want to rely more heavily on the model, and make \( R \) large to penalize the sensor noise. On the other hand, if we have an inaccurate model, then we would do better by simply following the sensor, and we increase \( Q \) to penalize process noise. For this particular experiment, we let the covariances \( Q \) and \( R \) vary in time according to which measurement is available. A higher penalty is given to the noisy probe signal, whereas the PIV data (when available) are assumed to be very accurate.

We initialize the Kalman filter with the values

\[
\hat{a}_{f,0} = E[a_0] \\
P_{f,0} = E[(a_0 - \hat{a}_{0})(a_0 - \hat{a}_{0})^T],
\]

where \( P \) is the covariance of the estimation error. The filter is then updated using the following equations, for \( k = 1, 2, \ldots \):

\[
P_{f,k}^{-} = F P_{f,k-1}^{-} F^T + Q_{k-1} \quad (27)
\]

\[
K_{f,k} = \frac{P_{f,k}^{-} H_k^T \left( H_k P_{f,k}^{-} H_k^T + R_k \right)^{-1}}{32}
\]

\[
\hat{a}_{f,k}^{+} = \hat{a}_{f,k}^{-} + K_{f,k} \left( \eta_k - H_k \hat{a}_{f,k}^{-} \right) \quad (28)
\]

\[
P_{f,k}^{-} = (I - K_{f,k} H_k) P_{f,k}^{-} \quad (29)
\]

\[
P_{s,k}^{+} = \frac{1}{I - F K_{f,k} H_k} \quad (30)
\]

\[
\hat{a}_{s,k}^{-} = \hat{a}_{s,k}^{+} = \hat{a}_{s,k}^{+} + K_{s,k} \left( \hat{a}_{s,k}^{+} - \hat{a}_{s,k}^{-} \right) \quad (31)
\]

3. Kalman smoother

The Kalman filter is a causal estimation technique, using only past and current data in forming a state estimate. In a pure post-processing application, we can make use of data at future timesteps to further improve the estimate. These non-causal filters are referred to as “smoothers.” We focus here on fixed-interval smoothing, in which data is available over a fixed interval (here, the length of the experiment). Specifically, we use a variant of the Kalman smoother called the Rauch-Tung-Striebel (RTS) smoother. RTS smoothing consists of a forward pass over the data using a standard Kalman filter (as described above), followed by a backwards pass with the RTS smoother.

We assume that the data are available from timesteps 0 to \( N_t \). After performing a forwards pass with a Kalman filter, the smoother is initialized with the values

\[
\hat{a}_{s,N_t} = \hat{a}_{f,N_t} \\
P_{s,N_t} = P_{f,N_t}^{+},
\]

We then iterate over \( k = N_t - 1, \ldots, 1, 0; 19 \)

\[
\mathcal{T}_{f,k+1} = \left( P_{f,k+1}^{-} \right)^{-1} \quad (32)
\]

\[
K_{s,k} = P_{f,k}^{+} F^T \mathcal{T}_{f,k+1} \quad (33)
\]

\[
P_{s,k} = P_{f,k}^{+} - K_{s,k} \left( P_{f,k+1}^{-} - P_{s,k+1} \right) K_{s,k}^T \quad (34)
\]

\[
\hat{a}_{s,k}^{+} = \hat{a}_{f,k}^{+} + K_{s,k} \left( \hat{a}_{s,k+1}^{+} - \hat{a}_{f,k+1}^{+} \right). \quad (35)
\]
IV. Experimental setup

A. Facility and instrumentation

We use time-resolved particle image velocimetry (TRPIV) to measure the velocity in the near wake behind a flat plate model with an elliptical leading edge and blunt trailing edge. The experiments are conducted in an Aerolab wind tunnel at the University of Florida Research and Engineering Education Facility. This open-return, low-speed wind tunnel has a test section that measures \(15.3 \text{ cm} \times 15.3 \text{ cm} \times 30.5 \text{ cm}\) in height, width, and length, respectively. The test section is preceded by an aluminum honeycomb, an anti-turbulence mesh screen and a 9:1 contraction section. A centrifugal fan driven by a variable frequency motor forward of the contraction section and screens controls the airspeed. The test section velocity is set by referencing the static and stagnation pressures from a Pitot-static tube placed at the beginning of the test section. The pressure differential is read by a Heise ST-2H pressure indicator with a 0–2 in-H\(_2\)O differential pressure transducer. For the experimental results that follow, the leading edge of the model is placed a few millimeters downstream of the test section entrance, as seen in Fig. 1.

For this study, we consider the von Kármán vortex street that develops behind the blunt trailing edge of the model. The two-dimensional flat plate model has a 4:1 (major axis-to-minor axis) elliptical leading edge, transitioning to a flat portion at the minor axis of the ellipse, and terminating abruptly with a flat trailing edge. Unlike other two-dimensional bluff bodies with similar wake dynamics (e.g., cylinder), the lack of surface curvature at the trailing edge simplifies the measurement of the near wake region. The PIV laser sheet illuminates the entire region behind the trailing edge surface without complex positioning or mirrors. The thickness-to-chord ratio is 7.1\%, with a chord and span of 17.9 cm and a span of 15.2 cm. For this analysis, the free-stream velocity \(U_\infty\) is set to 4.2 m/s, which corresponds to a Reynolds number of 50,000 based on chord length.

A Lee Laser 800-PIV/40G Nd:YAG system capable of up to 40 W at 10 kHz is paired with appropriate optics to generate a laser sheet for PIV measurements. As shown in Fig. 1, the light sheet enters the test section through a clear floor. The vertically oriented light sheet is aligned with the midspan of the model and angled such that the rays of light run parallel to the trailing edge without grazing the surface. This alignment prevents unwanted, high-intensity surface reflections and is necessary for well illuminated flow near the trailing edge, where particle densities can be low.

We image the seeded flow with an IDT MotionPro X3 camera with a 60 mm Nikon lens. The camera has a maximum resolution of 1280\(\times\)1024 and a sampling rate of 500 Hz for integration of all pixels. A sampling frequency of 800 Hz is achieved by reducing the number of pixels captured for each image, such that the effective image resolution is 600\(\times\)1024. The laser and cameras are synchronized by a Dantec Dynamics PIV system running Dantec Flow Manager software. The seeding for the free-stream is produced by an ATI
TDA-4B aerosol generator placed upstream of the tunnel inlet. The seed material is olive oil and the typical particle size is 1 µm.

LaVision DaVis 7.2 software is used to process the PIV data, using the following procedure. First, local minimum intensity background frames are subtracted from the raw image pairs. This step increases the contrast between the bright particles and the illuminated background by reducing the influence of static background intensities and noise bands. Then, surface regions and areas with poor particle density are masked (ignored) before computing multigrid cross-correlations. The processing consists of three passes with $64 \times 64$ pixel$^2$ interrogation windows and 75% overlap, followed by two refining passes with $32 \times 32$ pixel$^2$ interrogation windows and 50% overlap. In between passes, outliers are reduced by applying a recursive spatial outlier detection test. The final vectors are tested for outliers via the universal outlier spatial filter and the multivariate outlier detection test, an ensemble-based technique. Holes or gaps left by vector post-processing, which is less than 6% of the total vectors for all ensembles, are filled via Gappy POD. The final spatial resolution of the PIV measurements is approximately 8% of the trailing edge thickness.

B. Data acquisition

We acquire ten records of TRPIV images at a rate of 800 Hz. Each record is comprised of nearly 1400 sequential image pairs. To obtain a coarsely sampled (in time) set of velocity fields, we simply take a down-sampled subset of the original TRPIV data, such that the reduced sampling rate does not satisfy the Nyquist sampling theorem. This is intended to mimic the capture rate of a standard PIV system, which for many flows is not able to resolve all the characteristic time scales. Typical sampling rates for such commercially available systems are on the order of 15 to 50 Hz. For the estimation results that follow, one out of every 25 sequential velocity fields is used for estimation, which corresponds to a sampling rate of 32 Hz.

We also acquire a time-resolved probe signal by extracting the vertical velocity $v$ from a single point in the TRPIV velocity fields. Because this probe originates from within the velocity field, the samples are acquired synchronously with the coarsely sampled velocity fields, and span the time intervals between them (see Fig. 2). A probe within the flow field is a less-than-ideal choice for point-based estimation because non-intrusive point measurements within the flow are not viable for simultaneous PIV measurements nor for feedback flow control. Other experiments similar to this one have applied stochastic estimation successfully with non-intrusive surface pressure sensors. Unfortunately, limitations associated with the model thickness and low dynamic pressure fluctuations make such an approach unfavorable for this particular experiment. However, it seems likely that the methods developed here can be extended to work with surface-based sensors.

The estimation algorithm developed in this manuscript relies on the stochastic relationship between the point measurement and the time-varying POD coefficients. As such, the time-resolved probe measurement must correlate to structures within the flow field in order for the estimation to work properly. After all, the time-varying POD coefficients must be estimated using this measurement, which implies a dependence or relationship between the two. Consequently, the outcome of the estimation can be sensitive to the placement of the sensors. Motivated by the work of Cohen et al., we place our sensor at the node of a POD mode. In particular, we choose the point of maximum $v$ velocity in the third POD mode (see Fig. 5), as a heuristic analysis determined that the dynamics of this mode were the most difficult to estimate.

![Figure 2](image-url)  
**Figure 2.** Example cartoon of data acquisition method. Probe data is collected synchronously with TRPIV snapshots. The TRPIV are down-sampled to get a non-time-resolved dataset (red). This subset of the TRPIV data is used to develop static and dynamic estimators. Cartoon does not depict actual sampling rates.
C. Estimation procedure

Four of the ten TRPIV records are used as a “training set” for model identification. The remaining data are reserved for estimation and error analysis. We first analyze these data using proper orthogonal decomposition (POD), seeking a low-dimensional basis with which to approximate the flow field. For this computation, the time-resolved aspect of the records is not utilized. The key assumption here is that the POD modes computed from the time-resolved velocity fields are the same as those generated from randomly sampled velocity fields. This is valid in the limit of a large, statistically independent snapshot ensemble. The intention of this study, after all, is to approximate the core structures of time-resolved velocity fields from coarsely sampled (in time) velocity fields and time-resolved point measurements. Therefore, the estimation methods applied here do not rely on time-resolved, full-field data nor any derived quantities. The time-resolved attributes of the full-field data are used only for error analysis of the low-order estimates.

Once the POD modes are computed, the TRPIV snapshots are down-sampled to obtain a coarsely sampled subset. These coarsely sampled vector fields are projected onto the first seven POD modes, yielding a coarsely sampled time history of POD coefficients. We take these POD coefficients along with the synchronous, time-resolved probe signals and compute coefficients for linear stochastic estimation (LSE). Specifically, we use the multi-time-delay variant of LSE-POD.\textsuperscript{15,16} Again, for this part of the analysis, only coarsely sampled data from the training set are used. At a sampling rate of 32 Hz, this amounts to about 220 snapshots.

We then use the LSE-POD coefficients to estimate a time-resolved history of the POD coefficients. With these initial, stochastic estimates, we identify a dynamic model as described in Sec. III.C. Using this model, we implement a Kalman smoother and compute a second estimate of the time history of the POD coefficients. The results of the two estimation approaches are then evaluated through a comparison with the original, time-resolved velocity fields and POD coefficients.

Once more, we emphasize that though a time-resolved set of velocity fields (and thus POD coefficients) is available, we do not use any such data for any part of the model identification procedure. All computations are carried out as if the PIV data were acquired by a standard (non-time-resolved) system that yields independent samples. The only time resolved data used for this portion of the analysis is a velocity probe signal.

V. Results and discussion

The results of the experiment described in Sec. IV are discussed below. This discussion is broken into two main parts. First, we analyze the dynamics of the wake flow, using proper orthogonal decomposition (POD), dynamic mode decomposition (DMD), and standard spectral analysis methods. An effort is made to identify key characteristics of the wake, including the dominant frequencies and any coherent structures. In doing so, we allow ourselves access to the time-resolved PIV velocity fields, taken at 800 Hz.

Then the PIV data are down-sampled, leaving snapshots taken at only 32 Hz. These velocity fields are combined with a time-resolved point measurement of velocity (again at 800 Hz) for use in estimating the time-resolved flow field. We compare the results of linear static estimation (LSE) on the POD coefficients (LSE-POD) to those of dynamic estimation using a Kalman smoother.

A. Wake characteristics

1. Global/modal analysis

Global/modal analysis can be useful in identifying coherent structures and characteristic flow frequencies. These methods require full-field data with high spatial and/or temporal resolution, which are not always available. For instance, POD is a method that decomposes a flow into orthogonal spatial modes based on their energy content. It is well-suited for PIV snapshots because it requires only a statistically-independent set of data. On the other hand, techniques such as discrete Fourier transforms or DMD, which identify global structures of fixed temporal frequency, require time-resolved data. When these data are not available, they can be approximated using an estimation procedure like the one developed in this work. The aforementioned tools can then be applied to the approximated, time-resolved data.

We first perform these analyses using the time-resolved particle image velocimetry (TRPIV) data. (The results are later compared to an analysis using the estimated fields.) At a chord Reynolds number of 50,000 (based on $U_{\infty}$), the wake behind the flat plate displays a clear Kármán vortex street, as seen in Fig. 3. POD
analysis shows that 79.61% of the energy in the flow is captured by the first two modes (Fig. 4). The structure of these dominant modes, illustrated in Fig. 5, (b) and (c), demonstrates that they capture the dominant vortex shedding behavior. Each subsequent mode contributes only a fraction more energy, with the first seven modes containing 84.99% in total. (For the remainder of this analysis, we consider this seven-mode POD basis.) The lower energy modes also contain coherent structures, with those shown in Fig. 5(e)–(h) resembling the modes computed by Tu et al. for a simulation of a similar flow.29 However, without further analysis, their physical significance is unclear.

DMD analysis of the time-resolved velocity fields reveals that the flow is in fact dominated by a single frequency. The spectrum shown in Fig. 6 has a clear peak at a Strouhal number St = 0.27 (based on plate thickness h and freestream velocity $U_{\infty}$), with secondary peaks at the near-superharmonic frequencies of 0.52 and 0.79. The corresponding DMD modes (Fig. 7) show structures that resemble the POD modes discussed above. Because DMD analysis provides a frequency for every spatial structure, we can clearly identify the harmonic nature of the modes, with the modes in Fig. 7(a) corresponding to the dominant frequency, those in Fig. 7(b) corresponding to its first superharmonic, and those shown in Fig. 7(c) corresponding to its second superharmonic.

Furthermore, because DMD identifies structures based on their frequency content, rather than their energy content (as POD does), these modes come in clean pairs. Both DMD and POD identify similar structures for the dominant shedding modes (Fig. 5(b) and (c), Fig. 7(a)), but the superharmonic pairs identified by DMD do not match as well with their closest POD counterparts. For instance, the POD mode shown in Fig. 5(e) closely resembles the DMD modes shown in Fig. 7(b), whereas the mode shown in Fig. 5(g) does not. Similarly, Fig. 5(h) depicts a mode resembling those in Fig. 7(c), while Fig. 5(f) does not.

Interestingly, the third POD mode is not observed as a dominant DMD mode. This suggests that the structures it contains are not purely oscillatory, or in other words, that it has mixed frequency content. As such, its dynamics are unknown a priori. This is in contrast to the other modes, whose dynamics should be dominated by oscillations at harmonics of the wake frequency, based on their similarity to the DMD modes. This motivates our placement of a velocity probe at the point of maximum $v$-velocity in the third POD mode (as suggested by Cohen et al.), in an effort to better capture its dynamics.34 This location can be seen (plotted over a vorticity field) in Fig. 5(d).

2. Point measurements

Fig. 8 shows a time trace of the probe signal collected in the flat plate wake. We recall that there is no physical velocity probe in the wake. Rather, we simulate a probe of $v$-velocity by extracting its value from the TRPIV snapshots (see Fig. 5(d) for the probe location). Because PIV image correlations are both a spatial average across the cross-correlation windows and a temporal average over the time interval between image laser shots, PIV probe measurements typically do not resolve the fine scale structures of turbulence. To simulate a more realistic probe, Gaussian white noise is artificially injected into this signal, at various levels. We define the “noise level” $\gamma$ as the ratio of the squared root-mean-square (RMS) value of the noise to the squared RMS value of the probe signal:

$$\gamma = \frac{\langle n^2 \rangle_{\text{RMS}}}{\langle v^2 \rangle_{\text{RMS}}}.$$  \hspace{1cm} (36)
Figure 4. Energy content of the first $r$ POD modes, normalized with respect to the mean perturbation energy.

Figure 5. Spanwise vorticity of POD modes computed from TRPIV fields. The modes are arranged in order of decreasing energy content, labelled with mode indices matching those used in Fig. 4. Most resemble modes computed by Tu et al. in a computational study of a similar flow. (a) Mean flow. (b), (c) Dominant shedding modes. (d) Unfamiliar structure, with $v$-velocity probe marked by $(\circ)$. (e), (g) Anti-symmetric modes. (f), (h) Spatial harmonics of dominant shedding modes.

where the prime notation indicates the mean-subtracted value. This noise level is the reciprocal of the traditional signal-to-noise ratio. Note that the noise level only reflects the contribution of noise from the additive noise and does not reflect the inherent noise in the velocity probe signal. We consider six noise levels ranging from 0.01 to 0.36, in addition to the the original signal ($\gamma = 0$). Fig. 8 shows a comparison of the original signal to artificially noisy signals. We see that though the addition of noise produces random fluctuations, the dominant oscillatory behavior is always preserved.

A spectral analysis of the probe data, seen in Fig. 9, confirms that even when noise is added, the shedding frequency is preserved. This is to be expected, as the addition of white noise only adds to the broadband spectrum. The dominant peaks in the autospectra lie at $St = 0.27$, in agreement with the dominant DMD frequency. This confirms our previous characterization of the dominant DMD (and POD) modes as structures capturing the dominant vortex shedding in the wake.

The autospectra also show clear harmonic structure, again confirming the behavior seen in the DMD spectrum. However, as the broadband noise levels increase, the third and fourth harmonics of the wake frequency become less prominent relative to the noise floor. This loss of the harmonic structure carries certain implications for estimation. Most notable is that the these harmonic fluctuations in the probe signal will not correlate as strongly with the time-varying POD coefficients. Consequently, the flow field estimates based on the noisy probe signals may not capture the corresponding harmonic structures as well. The inclusion of noise is designed to be a test of estimator robustness, as future applications of the static and dynamic estimators presented here will be applied to pressure and shear stress sensors that are inherently noisy.
Figure 6. DMD spectrum. Clear harmonic structure is observed, with a dominant peak at St = 0.27, followed by superharmonic peaks at 0.52 and 0.79.

(a) St = 0.27
(b) St = 0.52
(c) St = 0.79

Figure 7. Spanwise vorticity of DMD modes computed from TRPIV velocity. Note the similarity of these modes to the POD modes depicted in Fig. 5. (a) Dominant shedding modes. (b) Temporally superharmonic modes; spatially anti-symmetric. (c) Further superharmonic modes; spatial harmonics of dominant shedding modes.

B. Low-dimensional flow-field estimation

1. Determining an optimal delay interval for LSE-POD

We find that with $\tau^* = 0$, multi-time-delay LSE-POD estimation of the first two POD coefficients is poor, unless multiple probes are used. Here, $\tau^*$ is the non-dimensional time-delay, defined as

$$\tau^* = \frac{\tau U_{\infty}}{h},$$

where $U_{\infty}$ is the freestream velocity and $h$ is the model thickness. This follows the results of Durgesh & Naughton (2010), who conducted a very similar bluff-body wake experiment. The cause lies in the fact that there is often a $\pm 90^\circ$ phase shift between the probe signal and the time history of one of the POD coefficients, nulling the LSE cross-correlations. Without a strong correlation for both of these coefficients, we are unable to accurately estimate the correct phase between them.

Durgesh & Naughton (2010) consequently derived and demonstrated the accuracy of a multi-time-delay approach that accounts for this phase mismatch and estimates the time-varying coefficients for the first two POD modes very well. Their work motivates the use of multi-time-delay LSE-POD in this manuscript. (For brevity we refer to this method as LSE-POD hereafter, assuming that the multi-time-delay variant is used.) In this method, rather than using the probe measurement with zero delay to estimate the POD coefficients, we also include probe data with multiple delays collected within some time interval. This improves the cross-correlations between the time-varying POD coefficients and the probe measurements. The delays span the interval

$$-\tau_{\text{min}} U_{\infty}/h \leq \tau^* \leq \tau_{\text{max}} U_{\infty}/h,$$

where $\tau_{\text{min}}$ and $\tau_{\text{max}}$ are the minimum and maximum time-delays used, respectively.
where \( \tau_{\text{max}} \) and \( \tau_{\text{min}} \) are greater than zero. (In this document, only symmetric intervals are considered, such that the intervals are bounded by \( \pm \tau^*_r \).)

Durgesh & Naughton demonstrate that an optimum time delay exists for estimating the unknown POD coefficients.\(^\text{16}\) In order to determine the optimal value, an estimation error must be computed and evaluated. For the present study, the flow field is approximated using the first seven POD modes. The corresponding vector \( \vec{a} \) of POD coefficients encodes a low-dimensional representation of the velocity field, with a corresponding kinetic energy \( E = \| \vec{a} \|^2 \) (see Eq. (10)). We wish to quantify the error between the true coefficients \( \vec{a} \) and the estimated POD coefficients \( \hat{\vec{a}} \). One way to do so is to simply compute the kinetic energy contained in the difference of the two corresponding velocity fields. If we normalize by the mean kinetic energy of the snapshot set, this gives us an error metric

\[
e(t) = \frac{\| \vec{a}(t) - \hat{\vec{a}}(t) \|^2}{\langle \| \vec{a}(t) \|^2 \rangle} = \frac{\sum_{i=1}^{r} \left[ \hat{a}_i(t) - a_i(t) \right]^2}{\sum_{i=1}^{r} \langle a_i(t)^2 \rangle}. \tag{39}
\]

This can be interpreted as the fraction of the mean kinetic energy contained in the estimation error.

In finding the optimal interval of delays for LSE-POD, we use only down-sampled PIV data from the training set (see Sec. IV.C) to compute the LSE-POD coefficients. The estimation error is then assessed by taking another PIV record (outside the training set), and estimating its POD coefficients \( \hat{\vec{a}} \). These other PIV velocity fields are also projected onto the POD modes to get the true coefficients \( \vec{a} \), which we then compare to the estimated coefficients. The mean energy in the error \( \overline{\varepsilon} \) is calculated from these coefficients for values of \( \tau^* \) ranging from 0 to 12. The results are plotted in Fig. 10.
Figure 10. Mean energy in the LSE-POD estimation error for various time delay intervals. An optimal value of $\tau_{\text{max}}^*$ is observed.

The minimum value of $\bar{\sigma}$ occurs for a delay interval with $\tau_{\text{max}}^* = 2.48$. However, we note that due to the shallow minimum seen in Fig. 10, similar results would be expected for delays ranging between 1.7 and 3.0 (roughly). Note that the case of zero delay empirically demonstrates that LSE-POD without any time delay performs very poorly in this experiment (as argued theoretically above).

2. Kalman smoother design

The model for the Kalman smoother is identified using the method described in Sec. III.C. The data used for this process come from LSE-POD estimates of the POD coefficients of the training set data. That is, the training set data (four records) are used to compute LSE-POD coefficients, where an optimal time delay is determined using a fifth, independent record. (These coefficients give us a static estimator.) Recall that this training set data has been down-sampled and is not time-resolved. A time-resolved estimate of the POD coefficients for the training data is then computed using the LSE-POD coefficients. It is these LSE-POD estimates that are used to identify the model. No time-resolved velocity fields are used in this process.

As described in Sec III.C, the model consists of two, decoupled parts. The dynamics of the two dominant POD modes are assumed to be oscillatory, with an oscillation frequency determined from the autospectra shown in Fig. 9. The dynamics of the remaining five modes are identified using the LSE-POD estimates, again as described in Sec III.C. Once the model has been obtained, the Kalman smoother is initialized with the values

$\hat{\theta}_{f,0} = \theta_0$,

$P_{f,0} = 5I$,

where $I$ is the identity matrix. We assume that the initial set of POD coefficients $\theta_0$ is known, as this is a post-processing application where the PIV data are available at certain (non-time-resolved) instances. The noise covariances are taken to be

$Q_k = \begin{bmatrix} 1 & 1 & 0.004 & 0 \\ 1 & 0.5 & \ddots & \end{bmatrix}$

$R_k = \begin{cases} 2 \times 10^4 & \text{when only probe data are available} \\ 1 \times 10^{-10} I & \text{when PIV data are available} \end{cases}$

We heavily penalize the time-resolved velocity signal to mitigate the effects of noise (large $R$), while the PIV data are assumed to be very accurate relative to the model (small $R$). In addition, the diagonal matrix $Q_k$ is designed to account for the observation that the lower energy POD modes are more sensitive to noise in the probe signal, with the third mode more sensitive than the rest.
3. Error analysis

We now compare the performance of two estimators: a static LSE-POD estimator with an optimal time delay interval and a dynamic Kalman smoother, both described above. We apply each estimator to five PIV records that were not used in any way in the development of the estimators. The estimates of the POD coefficients are evaluated using the error metric $e$ defined in Eq. (39). These results are shown in Fig. 11. By definition, $e$ is non-negative, giving it a positively skewed distribution. As such, the spread of these values is not correctly described by a standard deviation, which applies best to symmetric distributions. To account for this, the error bars in Fig. 11 are adjusted for the skewness in the distribution of $e$.

We observe that for all noise levels, the mean error achieved with a Kalman smoother is smaller than that of the LSE-POD estimate. Furthermore, the rate of increase in the mean error is slower for the Kalman smoother than for the LSE-POD estimate, and the spread is smaller too. As such, we conclude that not only does the Kalman smoother produce more accurate estimates (in the mean), but it is more robust to noise. This robustness comes in two forms. The first is that for a given amount of noise in the signal, the expected value of the estimation error has a much wider distribution for LSE-POD than for a Kalman smoother. Secondly, as the noise level is increased, the LSE-POD estimation error increases more rapidly, indicating a higher sensitivity to the noise level. This is expected, as LSE-POD is a static estimation method (as discussed in Sec. III.B).

These results are further illustrated by comparing the estimated vorticity fields, for both zero and 0.36 noise contamination (as defined in Eq. (36)). Fig. 12 shows an instantaneous vorticity field and its projection onto the first seven POD modes. This projection is the optimal representation of the original vorticity field using these POD modes. We observe that the high energy structures near the trailing edge are captured well, while the far wake structures tend to be more smoothed-out.

With no noise, the LSE-POD estimate of the vorticity field (Fig. 13(a)), matches the projected snapshot quite well. The spacing and shape of the high-energy convecting structures in the Kármán vortex street are correctly identified. However, when the probe signal is contaminated with 0.36 noise, the estimated vorticity field shown in Fig. 13(b) bears little resemblance to the projection. In fact, the only structures that seem to match are features of the mean flow (Fig. 5(a)), which of course is not part of the time-varying estimate. Not only are the downstream structures captured poorly, but spurious structures are also introduced. On the other hand, the Kalman smoother estimates match the projected snapshot for both clean and noisy probe data (Fig. 13(c), (d)).

4. Estimation-based global/modal analysis

As a further investigation into the relative merits of LSE-POD and Kalman smoother estimation, we use the estimated velocity fields to perform DMD analysis. We recall that DMD analysis requires that the Nyquist sampling theorem be satisfied, where the sampling rate exceeds double the highest frequency of interest. Both the TRPIV data and the low-order estimations have a sampling frequency of 800 Hz ($St = 2.42$). The results of such an analysis from the true TRPIV data are shown in Figs. 6 and 7. The key results from the DMD analysis of the estimated flow fields (for both LSE-POD and Kalman smoother estimates) are shown.
Figure 12. Comparison of spanwise vorticity fields. (a) True PIV snapshot. (b) Projection onto a seven-mode POD basis. The first seven POD modes capture the location and general extent of the vortices in the wake, but cannot resolve small-scale features.

Figure 13. Comparison of estimated spanwise vorticity fields. Without noise, both the LSE-POD and Kalman smoother estimates match the POD projection shown in Fig. 12(b). The addition of noise to the probe signal causes the LSE-POD estimate to change dramatically, resulting in a large estimation error. In contrast, the Kalman smoother estimate remains relatively unchanged. (a) LSE-POD with $\gamma = 0$. (b) LSE-POD with $\gamma = 0.36$. (c) Kalman smoother with $\gamma = 0$. (d) Kalman smoother with $\gamma = 0.36$.

in Fig. 14. The minimum and maximum of all the considered noise levels are included in this modal analysis.

The fundamental frequency $St = 0.27$ is captured well by estimation-based DMD for both estimators and for both noise levels. The corresponding modes match as well, and illustrations are therefore neglected. (Refer to Fig. 7(a) for typical mode structures associated with the fundamental wake shedding.) For the superharmonic frequencies, however, the estimation-based DMD modes do differ in their structure, both among the various estimation cases (across methods, for varying noise levels) and in relation to the DMD modes computed directly from TRPIV data (Fig. 7).

Both the LSE-POD and Kalman smoother estimates seem to capture the first superharmonic ($St \approx 0.53$) well when no noise is added to the probe signal, but the Kalman smoother-derived modes more accurately match the anti-symmetric distribution observed in the TRPIV-based DMD modes. When the noise level is increased to 0.36, the Kalman smoother-based modes still exhibit the expected harmonic structure. For the second superharmonic, the estimation-based modes are worse for both estimation schemes. With and without noise, the DMD results from both estimators are comparable to each other. Both match the second superharmonic frequency and mode shape from the true data better without the artificial noise, but neither closely resembles the frequency and mode shape at the high noise level. This is not unexpected, as the harmonic fluctuations in the probe signals are expected to correlate less and less with the time-dependent POD coefficients as the noise floor increases. Detection of these global modes from non-time-resolved snapshots is made possible by reduced-order approximations of time-resolved, global snapshots.

Of course, DMD results from estimated, time-resolved snapshots can only be expected to extract structures based on the frequency content contained within the projections onto the reduced-order POD basis. These favorable results are not a surprise because modes associated with the fundamental shedding frequency and its harmonics are somewhat captured by the selected high-energy POD modes, although not as cleanly as the DMD modal analysis using estimated snapshots. With less mean error in the low-order estimates of
VI. Conclusions and future work

The three-step estimation procedure presented here proves to be effective in estimating the time-resolved velocity field in a bluff-body wake. Rather than estimate the flow field directly using linear stochastic estimation (LSE), we use LSE to aid in identifying a stochastic model for the lower-energy structures in the flow. This stochastic model is then combined with an analytic model of the dominant vortex shedding in the wake. The result is used to implement a Kalman smoother, whose estimates of the flow field are shown to be more accurate and robust to noise than the stochastic estimates used in the modeling process. A dynamic mode decomposition (DMD) computed using the Kalman smoother estimates identifies the same coherent structures as one computed using time-resolved PIV data, showing that the estimates correctly capture the oscillatory dynamics of the flow.

A natural next step in this work is to apply the same procedure at a higher Reynolds number. Increasing the Reynolds number will increase the complexity of the wake dynamics, with turbulent fluctuations playing a greater role. Initial investigations of the wake behind the same elliptical-leading-edge flat plate at a Reynolds number of 100,000 shows that the dynamics of the lower energy modes are indeed more difficult to capture at these conditions, as would be expected. This makes the stochastic model identification procedure even more important, as it is difficult to identify the dynamics with other methods. However, though the identified model is less accurate at a higher Reynolds number, preliminary results show that the resulting estimates are still able to capture the dominant features of the wake dynamics.

Another direction we would like to pursue is the use of surface sensors, for instance measuring pressure or shear stress, to observe the flow. These are of course more realistic devices for use in a flow control experiment than velocity probes placed in the wake. Potential difficulties lie in the fact that the value of the pressure and/or shear stress at a surface are not linearly related to the velocity field, in general. However, previous work has shown that stochastic methods can be effective in relating pressure and velocity, and may be effective if extended to shear stress. As such, a stochastic estimator could be embedded in a dynamic estimator to capture such a relationship empirically, perhaps as part of a nonlinear estimator (e.g., a sigma-point Kalman filter). For control purposes, it may also be necessary to move away from the use of
proper orthogonal decomposition (POD) to describe the flow field, as other methods have been shown to capture input-output dynamics better.\textsuperscript{2,22,23} In any case, a more accurate knowledge of the flow field offers the potential for more advanced control strategies, certainly more than would be possible with knowledge of a point measurement or the two dominant mode coefficients alone.

**Acknowledgements**

This work is supported by the Air Force Office of Scientific Research (grant FA9550-09-1-0257), under Dr. Doug Smith, whom we would like to thank for his support; the Office of Naval Research (grant N00014-11-1-0554), monitored by Dr. Ron Joslin; and the Florida Center for Advanced Aero-Propulsion. We would also like to thank Lauren Padilla for her role in many fruitful conversations.

**References**