AIAA JOURNAL
Vol. 53, No. 4, April 2015

Low-Dimensional Approach for Reconstruction of Airfoil Data via Compressive Sensing

Zhe Bai,* Thakshila Wimalajeewa,† Zachary Berger,‡ Guannan Wang,‡ Mark Glauser,§ and Pramod K. Varshney¶
Syracuse University, Syracuse, New York 13244-1240

DOI: 10.2514/1.J053287

Compressive sensing is used to compress and reconstruct a turbulent-flow particle image velocimetry database over a NACA 4412 airfoil. The spatial velocity data at a given time are sufficiently sparse in the discrete cosine transform basis, and the feasibility of compressive sensing for velocity data reconstruction is demonstrated. Application of the proper orthogonal decomposition/principal component analysis on the dataset works better than the compressive-sensing-based reconstruction approach with discrete cosine transform as the basis in terms of the reconstruction error, although the performance gap between the two schemes is not significant. Using the proper orthogonal decomposition/principal component analysis as the sparsifying basis, compressive-sensing-based velocity reconstruction is implemented, which outperformed discrete cosine transform. Compressive sensing preprocessing (filtering) with discrete cosine transform as the basis is applied to a reduced number of particle image velocimetry snapshots (to mimic conditions with limited time support) before application of proper orthogonal decomposition/principal component analysis. Using only 20 particle image velocimetry snapshots with a 10% compressive sensing compression, it is found that the proper orthogonal decomposition/principal component analysis modes 1 and 2 of the streamwise velocity component are very close to those extracted from full time support data (1000 particle image velocimetry snapshots in this case). Results demonstrate the feasibility and utility of a compressive-sensing-based approach for reconstruction of compressed or limited time support particle image velocimetry flow data.

Nomenclature

A	=	projection/measurement matrix
a_n	=	time-dependent proper orthogonal decomposition/principal component analysis coefficients
CR	=	compression ratio (M/N for discrete cosine-transform-based approach, and M/T for snapshot proper orthogonal decomposition/principal component analysis based approach)
$C(t, t')$	=	two-point temporal correlation matrix
K	=	number of nonzero elements in sparsifying basis domain coefficients
M	=	number of compressed measurements of velocity data
N	=	number of particle image velocimetry velocity vectors
N_m	=	number of proper orthogonal decomposition/principal component analysis modes
N_{tot}	=	total number of proper orthogonal decomposition/principal component analysis modes
$R_{ij}(x, x')$	=	ensemble-averaged two-point spatial velocity correlation tensor
s, r	=	streamwise, wall-normal component coefficients in the transformed domain
T	=	number of snapshots of velocity data
U, V	=	streamwise, wall-normal components of the actual velocity
y, z	=	projection of streamwise, wall-normal component velocities
λ	=	eigenvalues of the correlation matrix
Φ	=	basis in which the velocity data are sparse
$\phi_i^{(n)}(x)$	=	proper orthogonal decomposition/principal component analysis mode n eigenfunction, $i(= 1, 2)$th component of velocity

Presented as Paper 2013-0772 at the 51st AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, Grapevine (Dallas/Ft. Worth Region), TX, 7–10 January 2013; received 15 December 2013; revision received 5 June 2014; accepted for publication 14 August 2014; published online 25 November 2014.

*Graduate Student, Department of Mechanical and Aerospace Engineering. Student Member AIAA.
†Research Assistant Professor, Department of Electrical Engineering and Computer Science.
‡Post-Doctoral Researcher, Department of Mechanical and Aerospace Engineering. Student Member AIAA.
§Professor, Department of Mechanical and Aerospace Engineering. Fellow AIAA.
¶Distinguished Professor, Department of Electrical Engineering and Computer Science.

I. Introduction

IN THE past decade, larger and larger datasets have been (and are being) generated in wind tunnels and from flight tests due to enhanced capabilities to measure various aerodynamic variables, including spatially and temporally resolved velocity and pressure, resulting in datasets on the order of terabytes and beyond. As a result, we need faster and more efficient methods for processing and understanding them so that critical data can be extracted quickly to make timely decisions.

The aim of this study is to explore the application of a new efficient approach, compressive sensing (CS) [1,2], to reconstruct particle image velocimetry (PIV) data from a NACA 4412 airfoil. Implementation of efficient control of turbulent flow over the airfoil requires high-frequency sensing and actuating capabilities to capture the instantaneous velocity field data [3]. However, increasing the sampling rate is quite difficult and expensive, which makes the real-time processing a challenge. We therefore need low-dimensional tools [4–6] that can reduce the complexity of data collection and analysis in high-Reynolds-number turbulent flows. Proper orthogonal decomposition (POD)/principal component analysis (PCA) with the modified linear stochastic measurement low-dimensional techniques have been used to capture and estimate the most energetic structures [7,8]. Pinier et al. [8] performed closed-loop feedback control of the separated flow over the airfoil using these methods. These approaches require significant time support (large number of PIV snapshots) to determine the space-time correlations so that the POD/PCA eigenfunctions can be extracted. Further, it is required to compute all the eigenfunctions (modes), although only

significant ones in terms of the turbulent kinetic energy are used in representing and reconstructing the original data. Over the past couple of years, our group [9,10] and several others [11–19][**] have been exploring the application of CS to turbulence and general fluid dynamics. These include compressive sensing with dynamic mode decomposition to improve temporal resolution or rebuild low-dimensional modes [17–19], compressive sensing to rebuild a pressure field with a limited number of sensors [13,15], and improving the spatial and temporal resolution of tomographic PIV data through compressive sensing concepts [14].[**] In the CS framework, the original data are projected onto a lower-dimensional subspace based on a simple linear projection scheme to achieve compression. A signal is said to be sparse if represented in a proper basis, all but a few coefficients in that basis representation are negligible, and the sparsity index refers to the number of such significant coefficients. If the original data are sufficiently sparse in a particular domain, the sparse signal can be reliably reconstructed with a very small number of compressed measurements (which is slightly greater than the sparsity index of the signal and much less than the original signal dimension) using several approaches including optimization techniques [2,20,21] and greedy techniques [22].

It has been shown that many signals are either sparse in a certain domain or can be transformed to be sparse in some specific domain [23], which motivates the use of low-dimensional schemes. Turbulent flows that contain large-scale coherent structures fall into this category. In this study, we reconstruct a given N-dimensional physical field data (PIV data from the NACA 4412 airfoil) from a small number of data samples M obtained via random projections by exploiting sparsity. Within the fluid mechanics community and beyond, there is a growing need for large datasets with sufficient time support and spatial resolution. As the community continues to investigate more complex flowfields, PIV is an efficient means of probing the flow because it is spatially resolved and nonintrusive. For many applications, a large number of snapshots are desired for statistical convergence as well as for control purposes. Many times, however, sufficient time support is not easily achievable under normal operating conditions. For example, when studying wind turbines in the field, PIV would be a desirable means for acquiring velocity measurements. In this case, setting up and acquiring measurements in the field would be quite difficult, possibly resulting in limited number of snapshots. In addition to POD, several other compression schemes such as discrete cosine transform (DCT) and discrete wavelet transform (DWT) can be used to represent data in a lower dimension. To compress data with such schemes, the original data have to be first processed, and then one finds the most significant coefficients along with their locations in the given basis. Although these schemes are promising under certain circumstances, they require a higher computational burden at the compression stage. In contrast, in the compressive-sensing-based scheme investigated in this paper, the compression is performed with much less computational burden at the compression stage.

To reliably reconstruct the original velocity field (described in Sec. II) based on CS with a small number of compressive measurements, we need to identify the bases in which the velocity data are sparse. The approaches used to identify the bases are presented in Sec. III. Section IV describes the results of the study. We show that the spatial velocity data at a given time snapshot are considerably sparse in the DCT domain. Using DCT as the sparsifying basis, the velocity field at a given time snapshot is reconstructed using CS-based measurements via l_1 norm minimization. The performance of the CS-based velocity field reconstruction is compared to that with the traditional snapshot POD/PCA method [8] using the mean-squared error as the performance metric. We have the following comparison results of the two schemes.

1) When the compression ratio is small, the performance of the CS-based approach for velocity field reconstruction is slightly worse than that with the standard POD/PCA approach. However, as the compression ratio increases, both schemes depict identical performance.

2) To reconstruct the original velocity field at a given time snapshot based on the POD/PCA approach, it is required to compute the correlation matrix of the original fluctuating velocity and thus the need for a large time support. However, to do the same based on CS, only the projected data at one particular time snapshot and the knowledge of the basis in which the signal is sparse are used, and large time support is not required. Thus, the CS-based scheme does not need to store large datasets for any offline computation.

3) In the POD/PCA approach, the compression stage (or the computation of the significant eigenfunction modes) is computationally expensive, whereas the reconstruction phase is fairly simple. In contrast, with the CS-based approach, the compression stage is very simple and universal, whereas the reconstruction requires some computational effort.

In Sec. IV, we also explore combining the two schemes in different ways to reconstruct the velocity field appealing to the advantages of each scheme. For example, we employ POD/PCA to generate the sparsifying basis for the original data, where the POD/PCA modes are computed based on some training data and then apply CS. Because POD/PCA is a data-dependent scheme, it is capable of representing the original data more sparsely compared to DCT if the stochastic properties of the data remain unchanged over time. Thus, we observe an improved performance in velocity field reconstruction based on CS with POD/PCA as the sparsifying basis compared to that with DCT. We also lay out a method for extracting POD/PCA modes with limited time support using CS preprocessing on the snapshots with DCT as the sparsifying basis. With this approach, the small-scale structures in the velocity field are filtered out, and the large-scale ones are preserved. In this way, the first few POD/PCA modes can be extracted from a very small subset of PIV snapshots (the limited time support condition).

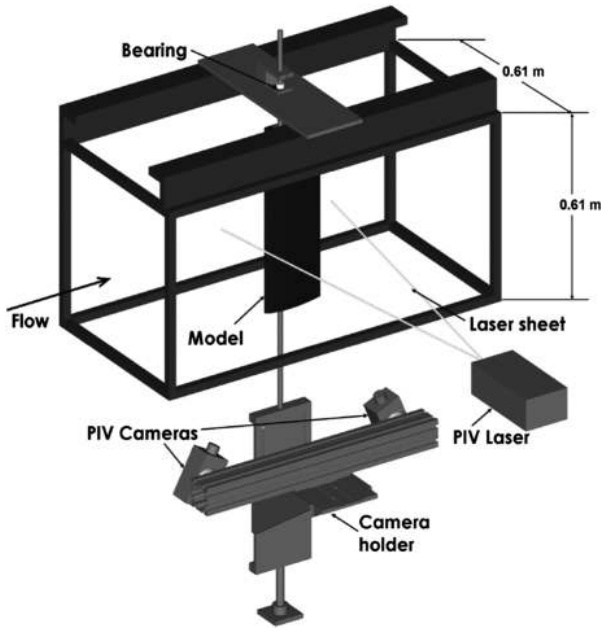


Fig. 1 Overall view of experimental setup located in the test section of the Syracuse University wind tunnel.

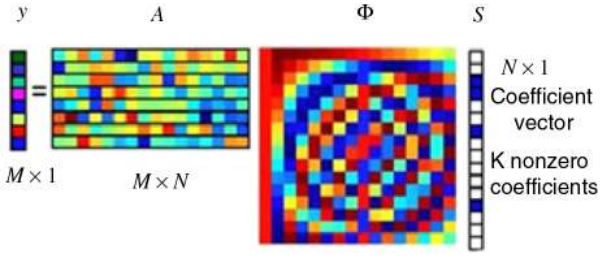


Fig. 2 Compressive sensing measurement process with a random Gaussian measurement matrix and DCT matrix (Baraniuk [20]).

**Petra, S., and Schnörr, C., "Tomopiv Meets Compressed Sensing," 2009, http://archiv.ub.uni-heidelberg.de/volltextserver/9760/1/puma09_submission.pdf [retrieved April 2014].

Table 1 Description of CS-based approach and direct snapshot POD/PCA approach

	CS_DCT/DWT basis	CS_POD/PCA basis	CS_Modified POD/PCA basis	Direct snapshot POD/PCA method
$\Phi(u(x,t_0)=\Phi s)$	DCT/DWT matrix $\Phi = f(N)$	$C = U(x,t)^T U(x,t) C\phi_n = \lambda\phi_n$ $\Phi^{(n)} = \frac{\sum_{n=1}^{T} U(x,t)\cdot\phi_n}{\left\|\sum_{n=1}^{T} U(x,t)\cdot\phi_n\right\|_2}, n = 1, \ldots, T$	$C_1 = y_1^T y_1 (y_1 = AU(x,t))$ $C_1\phi_n = \lambda\phi_n$	$C = U(x,t)^T U(x,t)$
$A\ (y = Au(x,t_0))$	$A = g(M,N)\ M < N$	Gaussian matrix $A = g(M,N)\ M < T$		$C\phi_n = \lambda\phi_n$ $a_n(t_0) = \int_D u(x,t_0)\phi^{(n)}(x)\,dx$
$\hat{s}\ (y = A\Phi s)$	$\hat{s}(N,1)$	l_1 norm minimization $\hat{s} = \arg\min \|s\|_1$ such that $y = A\Phi s$ $\hat{s}(T,1)$		$u(x,t_0) = \sum_{n=1}^{N_m} a_n(t_0)\phi^{(n)}(x)\,dx$
$\hat{u}\ (\hat{u}(x,t_0) = \Phi\hat{s})$		$\hat{u}(x,t_0) = \Phi\hat{s}$		

II. NACA 4412 Airfoil Particle Image Velocimetry Dataset

The PIV data used in this study were obtained in the subsonic wind-tunnel facility at Syracuse University, and a schematic of the experimental setup is shown in Fig. 1. The freestream velocity was set at 10 m/s, resulting in a Reynolds number of 135,000 based on a chord length of 20.32 cm. The extensive dataset was obtained as part of a large-scale closed-loop flow control set of experiments [3,8], and the particular dataset used here is from a baseline case at an angle of attack of 16 deg where the flow is separated. The in-plane components of velocity, U and V, are both measured using a Dantec Dynamics PIV system with two cameras side-by-side to double the measurement area size in the $X-Y$ (streamwise wall-normal) plane above the airfoil. The PIV sampling frequency was 4 Hz, which is much slower than the time scales in the 10 m/s flow, and so all sample data in the study can be considered statistically independent [8].

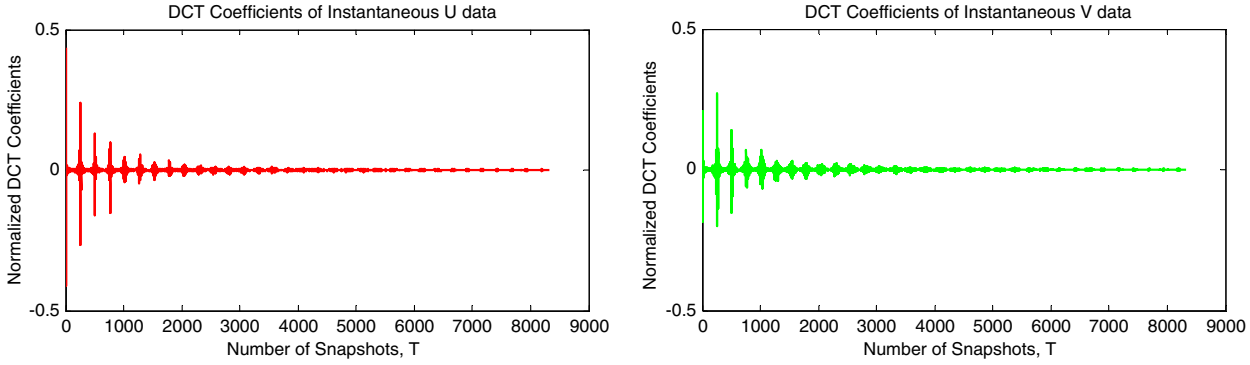
III. Reconstruction of the Velocity Field

In many fields, POD/PCA has been widely applied for the extraction of important features from large datasets. It is used for compression and for successful reconstruction from compressed data including turbulence, where it has been used to reconstruct velocity, pressure, etc. [3,8]. In this study, the reconstruction results based on compression via the CS methodology are compared with that using the POD/PCA method by conducting an error analysis. Compression and reconstruction using a combination of these two approaches is also explored.

A. Proper Orthogonal Decomposition/Principal Component Analysis

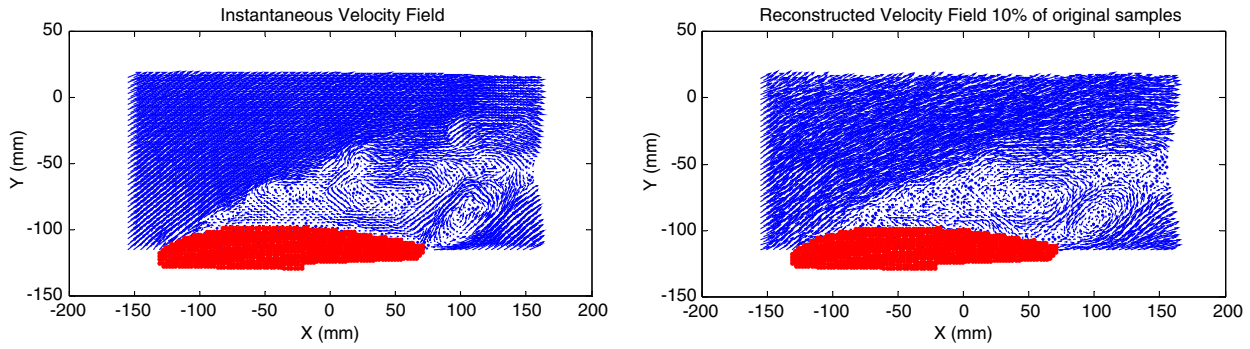
POD/PCA (we use these interchangeably) was first introduced into the turbulence community by Lumley [24] in 1967 as an unbiased method to study coherent structures in turbulent flows. It is a logical way to construct the basis functions to capture the most energetic features in a small number of modes, in which the first principal component (mode) has the largest energy. Thus, it enables an optimal low-dimensional representation of the energy contained in the flow.

The POD/PCA method maximizes the energy content of the flow in an orthogonal basis functions $\phi^{(n)}(x)$. These are obtained from the solution of the following integral eigenvalue problem (one-dimensional case shown here):



a) U component **b) V component**

Fig. 3 DCT coefficients of streamwise velocity U data and wall-normal velocity V data.



a) Instantaneous velocity field **b) Reconstructed velocity field, CR=10%**

Fig. 4 Instantaneous velocity field and CS-based reconstructed velocity field taking DCT basis with 10% of the original samples.

$$\int_D R_{ij}(x,x')\phi_j^{(n)}(x')\,dx' = \lambda^{(n)}\phi_i^{(n)}(x) \tag{1}$$

where $R_{ij}(x,x')$ is the ensemble-averaged two-point spatial velocity correlation tensor, defined as

$$R_{ij}(x,x') = \langle u_i(x,t_0)u_j(x',t_0)\rangle \tag{2}$$

where t_0 is a given PIV-snapshot time, and $\langle\cdot\rangle$ denotes the average over time. We can then extract the time-dependent (not time-resolved) expansion coefficients, by projecting the PIV velocity field onto the eigenfunctions, as follows:

$$a_n(t_0) = \int_D u_i(x,t_0)\phi_i^{(n)}(x)\,dx \tag{3}$$

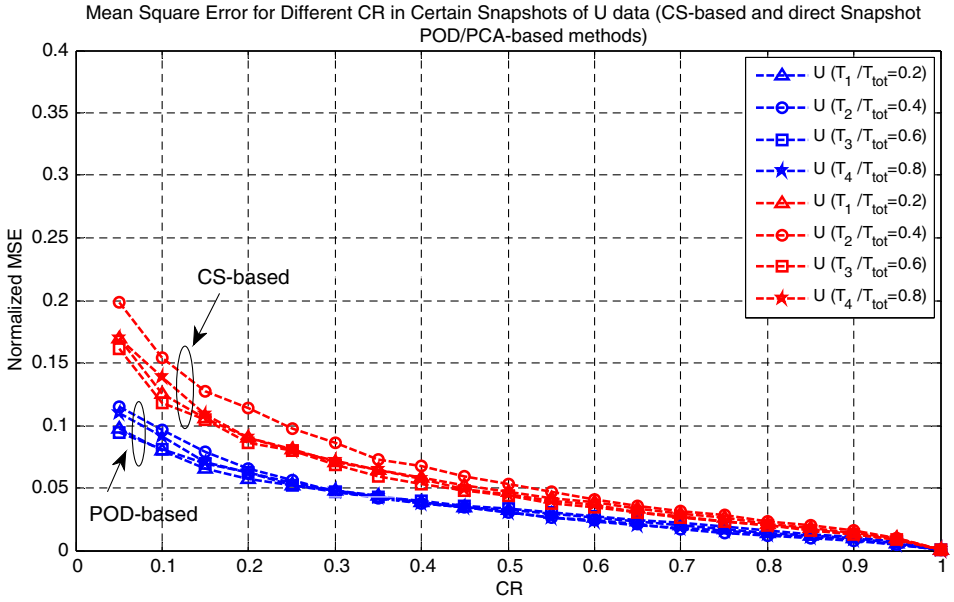
where $u_i(x,t_0)$ is the velocity field from a particular PIV time snapshot.

The eigenfunctions $\phi^{(n)}(x)$ of Eq. (1) are the optimal basis functions in terms of turbulent kinetic energy that are extracted from the PIV measurement ensemble. The Hilbert–Schmidt theory ensures that, if the random field occurs over a finite domain, an infinite number of orthonormal solutions can be used to express the original random velocity field $u_i(x,t_0)$; therefore, we can partially or completely reconstruct the original velocity field by projecting an $a_n(x,t_0)$ onto the eigenfunctions:

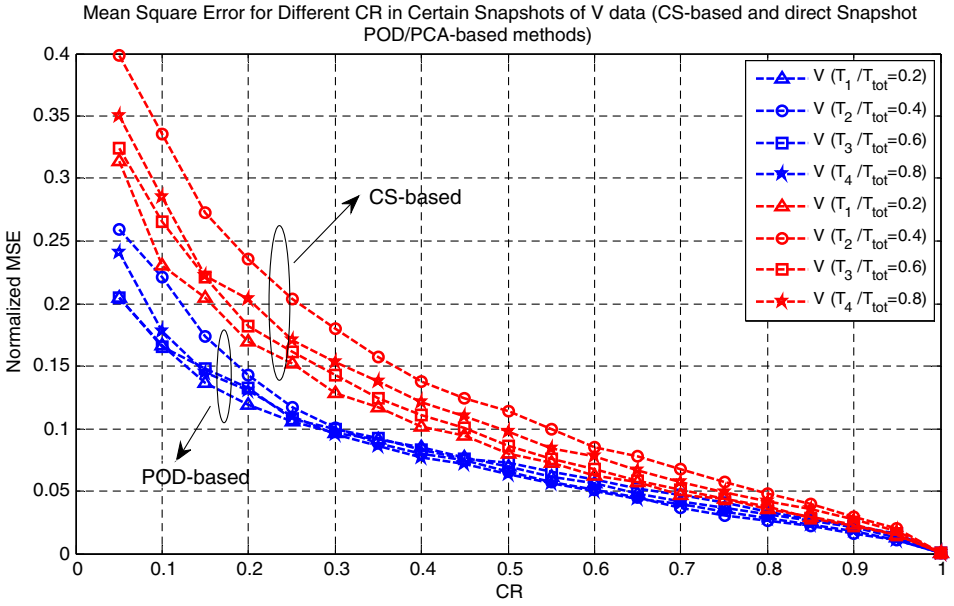
$$u_i(x,t_0) = \sum_{n=1}^{N_m} a_n(t_0)\phi_i^{(n)}(x)\,dx \tag{4}$$

where N_m is the number of modes with which we wish to reconstruct the velocity field. If N_m is chosen to be N_{tot}, the velocity field is completely and exactly reconstructed. Using POD/PCA, most of the kinetic energy can be captured in a small number of modes relative to the total number of modes, enabling an effective and accurate representation of the large coherent structures in terms of energy.

A modification to the classical POD/PCA approach was introduced by Sirovich [25] in 1987 to reduce the dimension of the eigenvalue problem, which is more widely used in the cases in which



a) U component



b) V component

Fig. 5 Normalized MSE comparison of CS-based and snapshot POD/PCA-based reconstruction.

the number of snapshots is much smaller than the number of grid points. In our case, the velocity vector map is composed of $N = 8320$ vectors in both X and Y components with $T = 1000$ statistically independent PIV velocity vector snapshots for each case; therefore, it is more appropriate and equivalent mathematically to use the snapshots POD/PCA method [8]. In this way, the eigenvalue problem is reduced to the dimension of T, instead of N as with the classical POD/PCA method [24]. The integral eigenvalue problem from Eq. (1) is equivalent to

$$\int_T C(t,t')a_n(t')\,\mathrm{d}t' = \lambda^{(n)}a_n(t) \tag{5}$$

where $C(t,t')$ is defined as

$$C(t,t') = \frac{1}{T}\int_D u_i(x,t)u_i(x,t')\,\mathrm{d}x \tag{6}$$

For the reason of consistency with the classical POD/PCA, the temporal eigenfunctions and eigenvalues are arbitrarily chosen to satisfy the following relation:

$$\langle a_m \cdot a_n \rangle = \lambda^{(m)}\delta_{mn} \tag{7}$$

The spatial eigenfunctions $\phi^{(n)}(x)$ must then be defined as

$$\phi_i^{(n)}(x) = \frac{1}{T\cdot\lambda^{(n)}}\int_T a_n(t)u_i(x,t) \tag{8}$$

so that they become orthogonal and satisfy

$$\int_D \phi_i^{(m)}(x)\phi_i^{(n)}(x)\,\mathrm{d}x = \delta_{mn} \tag{9}$$

The low-dimensional velocity fields are then reconstructed by projecting the temporal eigenfunctions $a_n(t)$ onto the spatial eigenfunctions $\phi_i^{(n)}(x)$, using Eq. (4).

B. Velocity Field Reconstruction Based on Compressive Sensing

According to the Nyquist sampling theory, one must sample at least two times faster than the signal bandwidth to avoid losing information when capturing a signal [20]. In many modern applications, the signal bandwidths have increased tremendously, requiring very high sampling rates, whereas the acquisition capabilities have not scaled sufficiently fast [26]. Handling such large amounts of data is difficult in many applications because it requires large processing power, space for storage, and time. It has been observed that lower-dimensional representations of large datasets are possible for many signals. A signal is said to be sparse if it can be represented as a combination of fewer basis vectors in an appropriate basis compared to the original signal dimension. The traditional way of representing such compressible or sparse signals in a lower dimension is to first "sample" and then "compress" or identify the most significant coefficients with respect to the particular domain. Although many signals of interest are known to be sparse after representing in a proper basis, a considerable computational effort is required at the compression stage.

In contrast, CS provides a new way to capture and represent compressible signals at a rate significantly below the Nyquist rate. More specifically, in the CS framework, the original sparse signal is projected onto a lower-dimensional subspace via a random projection scheme. It has been shown in [1,2,20] that the sparse signal can be reliably reconstructed based on a small number of such random projections via optimization techniques (such as l_1 norm minimization). Generally, CS relies on two principles: sparsity, which pertains to the signals of interest, and incoherence, which pertains to the sensing modality [21]. Sparsity expresses the idea that a discrete time (or space) signal depends on a number of degrees of freedom, which are comparatively much smaller than its (finite) length. More precisely, CS exploits the fact that many natural signals may be sparse or compressible in the sense that they have concise representations when expressed in the proper basis (Φ). Incoherence extends the duality between time and frequency and expresses the idea that objects having a sparse representation in Φ must be spread out in the domain in which they are acquired, just as a Dirac or a spike in the time domain is spread out in the frequency domain. Put differently, incoherence says that, unlike the signal of interest, the sampling/sensing waveforms have an extremely dense representation in Φ [21].

From the PIV system, we obtained the velocity vector map composed of $N = 8320(128\times 65)$ vectors with $T = 1000$ snapshots for each component. This PIV velocity database is not time-resolved, and hence we concentrate on the spatial problem. For the spatial velocity field, we focus on one snapshot, processing U and V velocity components individually, and consider each one of the two as a one-dimensional column vector $u_i \in R_n(i = 1,2)$.

For the original velocity data, we are collecting them as a time series, and CS is applied to the spatial data collected at one time snapshot. A single velocity snapshot $u_i(x,t_0)(i = 1,2)$ is processed at one time, and it is assumed that the one-dimensional column vector

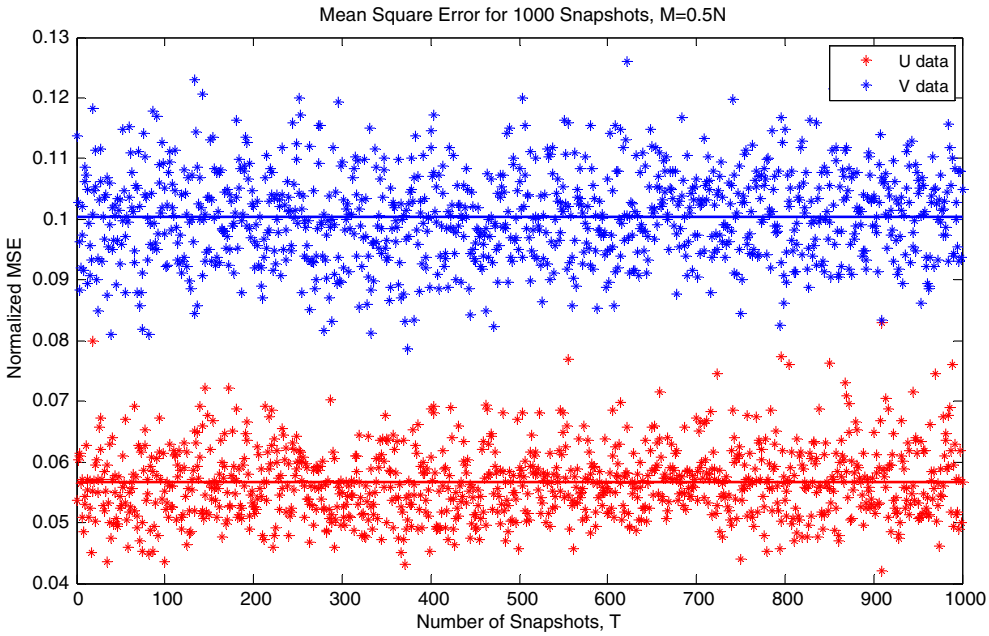


Fig. 6 Normalized MSE for 1000 snapshots with CS-based reconstruction.

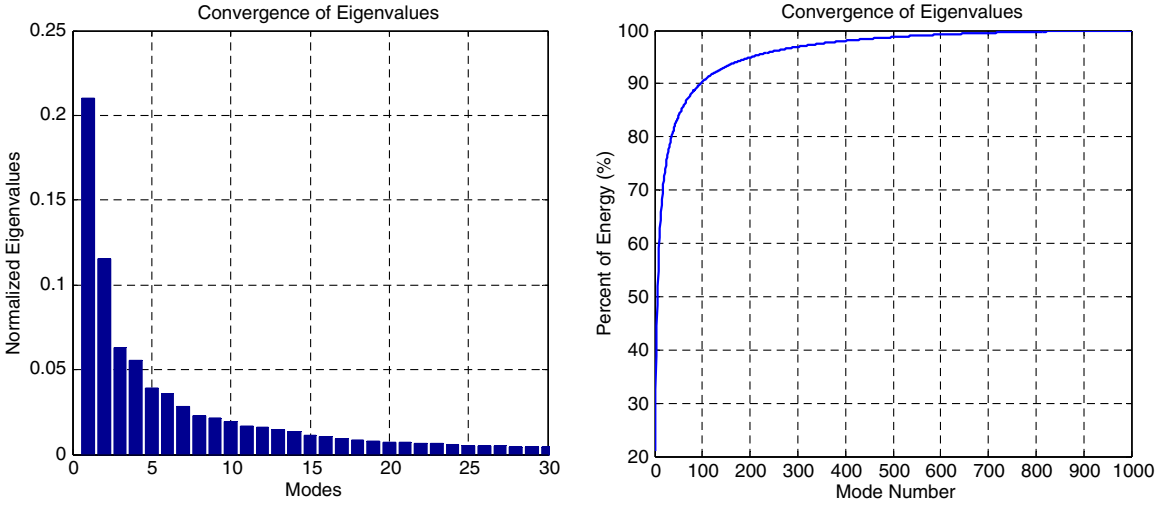
is such that there exists a transformation under which it is sparse. In this study, we assume only orthogonal bases. Specifically, there is an invertible transformation matrix Φ of size $N \times N$ such that

$$u_i = \Phi s_i \tag{10}$$

s_i is an $N \times 1$ column vector, and the basis Φ is taken as a sparsifying matrix such as a DCT matrix, or a DWT matrix. We can regard u_i and s_i as equivalent representations of the signal, with u_i in the space or time domain and s_i in the Φ domain. We say that s_i is K-sparse when it has at most K nonzero elements, with $K < N$, and u_i can be represented as the linear combination of only K vectors in the basis. The observation model is given by

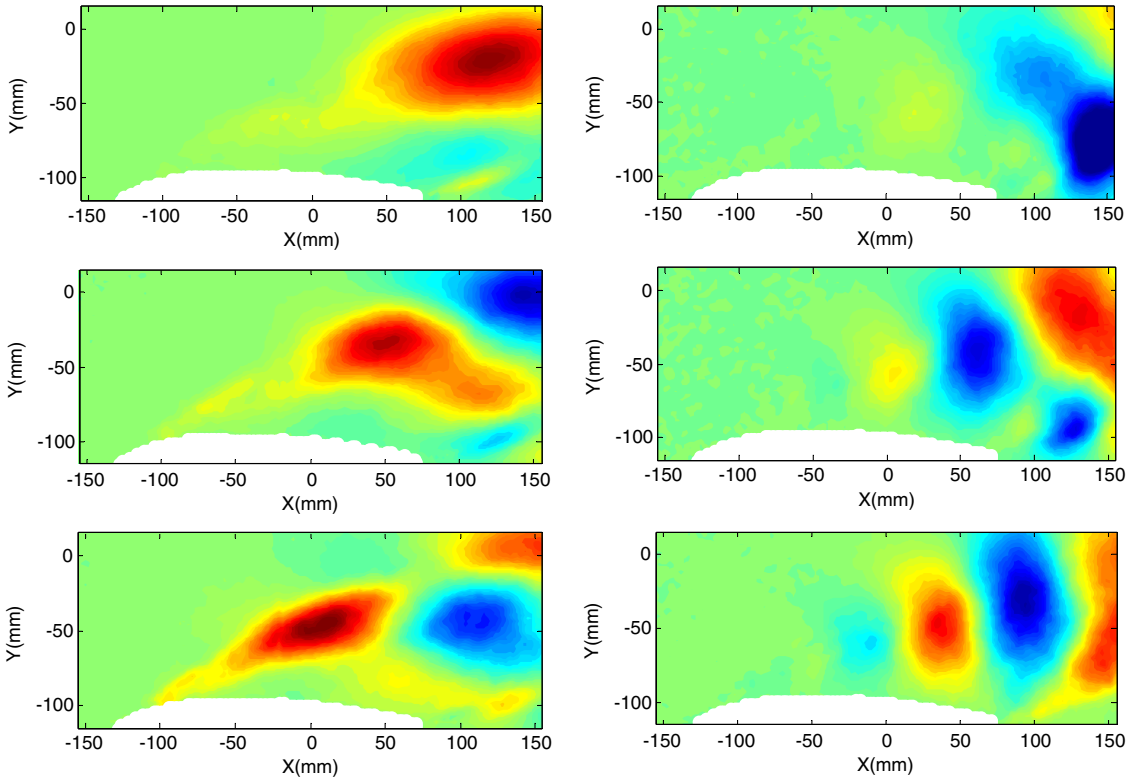
$$y_i = Au_i \tag{11}$$

where y_i is a column vector of size $M(< N)$, and A is an $M \times N$ projection matrix. To reliably reconstruct length N column vector u_i from length M observation vector y_i as in Eq. (11), the projection matrix A has to satisfy a certain stability condition that is commonly known as restricted isometry property (RIP) [27]. It has been shown that, when elements of A are taken as realizations of zero mean random variables (e.g., Gaussian, Bernoulli), the RIP condition is satisfied with high probability when $M > cK \log(N/K)$, where c is a small constant [27,28]. Although there are other possible matrices (such as random Bernoulli and random rows of DCT matrix) that satisfy RIP with high probability, we use Gaussian matrices in this paper. When using random projections, CS is universal in the sense



a) Normalized eigenvalues **b) Percentage of the energy distribution**

Fig. 7 Convergence of the eigenvalues, using the snapshot POD/PCA.



a) U component **b) V component**

Fig. 8 First three POD/PCA spatial eigenfunctions.

that the same mechanism for acquiring measurements can be used irrespective of the sparsity level or the basis in which the signal is sparse [29]. In our work, A is taken as a random matrix, whose elements are independent and identically distributed random variables from a Gaussian probability density function with mean zero and variance $1/N$ [30]. Then, by substituting Φ from Eq. (10), Eq. (12) can be expressed as

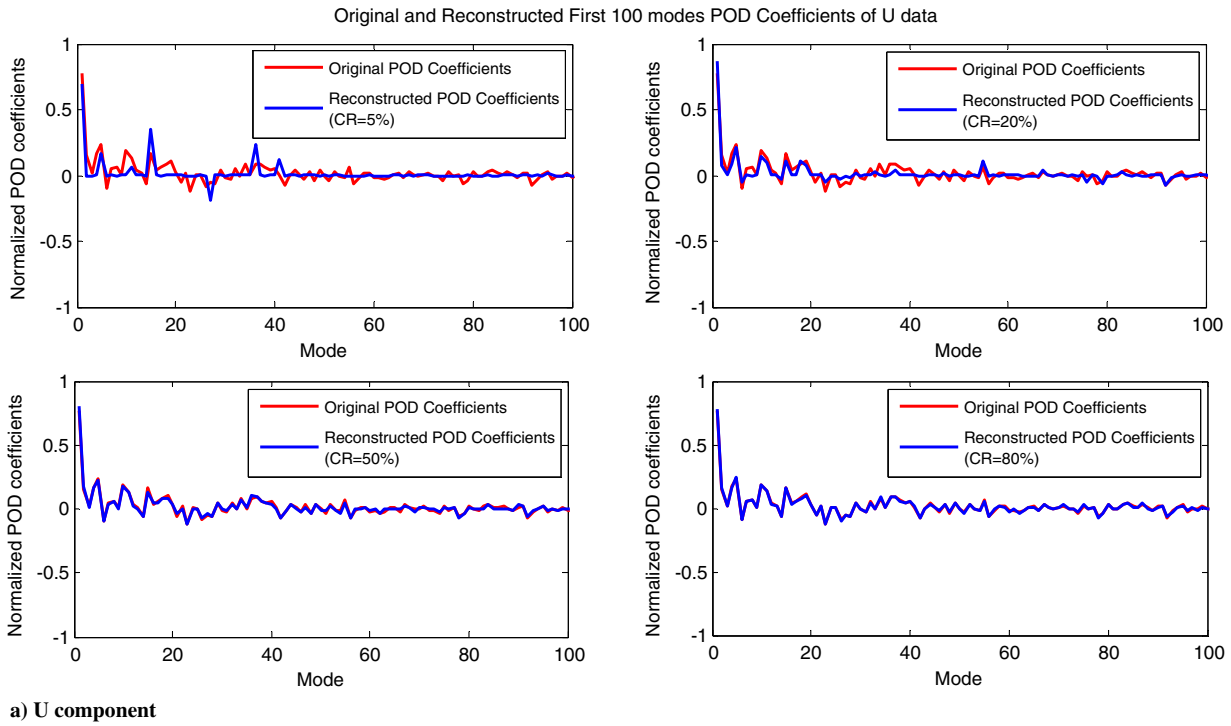
$$y_i = Au_i = A\Phi s_i = \Theta s_i \tag{12}$$

where $\Theta = A\Phi$ is an $M \times N$ matrix as shown in Fig. 2, and A is fixed when M is selected, which does not depend on the signal u_i.

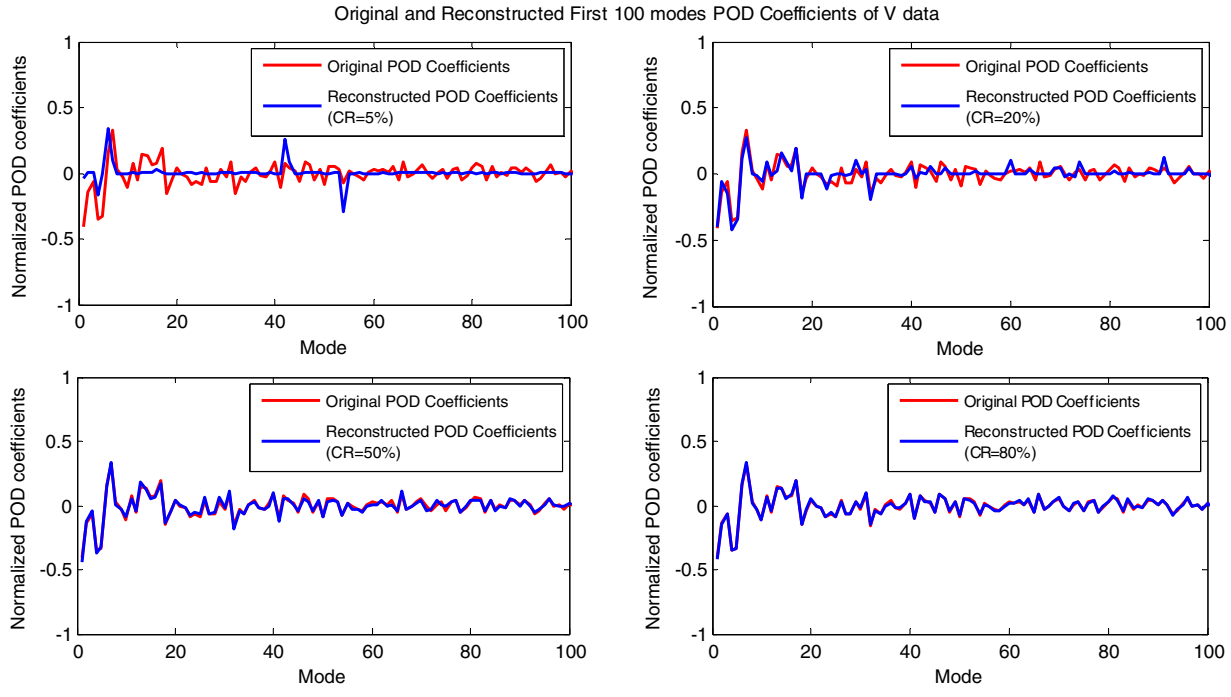
The measurement matrix A must allow the reconstruction of the length-N signal u_i from the length M measurement y_i. Because $M < N$, this problem appears ill-conditioned. However, if s is sparse, it has been shown that the sparse signal s can be recovered exactly with high probability based on length M vector y, if A satisfies RIP. Reconstruction can be performed based on optimization techniques or greedy algorithms [20–22,31]. A widely used optimization-based approach is to consider the following l_1 norm minimization algorithm [31]

$$\hat{s} = \arg \ \min \|s\|_1 y = A\Phi s \tag{13}$$

where $\|s\|_1$ denotes the l_1-norm of the column vector s. This is a convex optimization problem that reduces to a linear program [20]. We used the



a) U component



b) V component

Fig. 9 Original and reconstructed first 100 coefficients based on CS using snapshot POD/PCA basis for compression ratios $M/N = 5, 20, 50$, and 80%: a) U component, and b) V component.

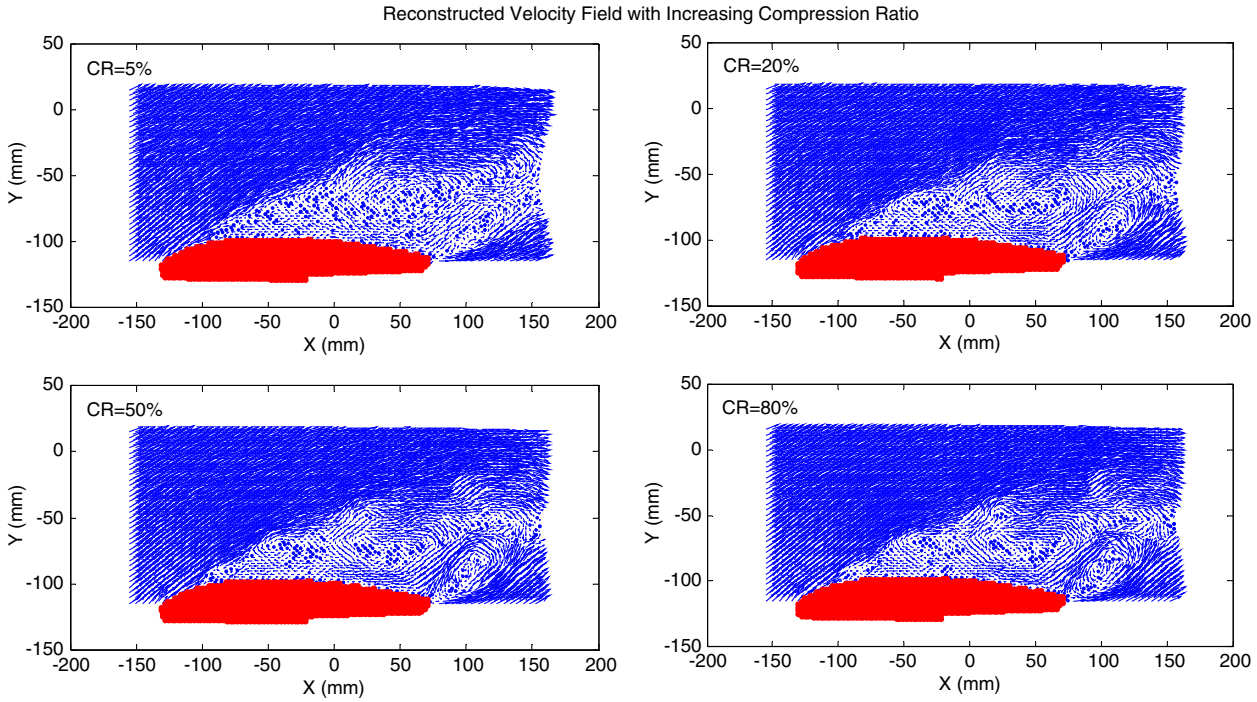


Fig. 10 Reconstructed velocity field based on CS with snapshot POD/PCA basis for compression ratios $M/T = 5, 20, 50$, and 80%.

interior point method to solve this problem, and Matlab codes are available.[††] Then, s is retrieved as the reconstructed $\hat{s}$, and the original signal is found through Eq. (10). We apply this technique for both U and V data at a given time snapshot.

IV. Performance Results

In this section, we provide results that demonstrate the efficiency of using CS-based schemes in recovering wind-tunnel data, obtained with PIV. The quality of the reconstruction based on CS depends on certain factors including the basis in which the data are sparse and the number of CS measurements. We use several bases to demonstrate the reconstruction capability of the CS-based approach including the DCT matrix (Sec. IV.A), DWT matrix, snapshot POD/PCA basis (Sec. IV.B), and modified snapshot POD/PCA (Sec. IV.C) as summarized in Table 1. For the velocity data considered in this paper, DWT does not transform the velocity vectors to a more sparse domain, and so only the data obtained for the DCT basis are presented in this paper. We also provide a comparison of these approaches with the direct snapshot POD/PCA method where the entire dataset is used to see the differences in terms of performance of the different data processing approaches. Finally, in Sec. IV.D, we conclude with a limited time support scenario where only a small number of PIV snapshots are compressed using the DCT basis and then used to extract estimates of the POD/PCA basis.

A. Performance of Compressive-Sensing Based Reconstructions with Discrete Cosine Transform Bases

First, DCT is used as the basis in which U and V are sparse. DCT is a commonly used low-dimensional representation for many signals because it is naturally orthogonal and only depends on the number of grid points. For a length N signal x, the mth DCT coefficient is given by

$$s(m) = \alpha(m) \sum_{k=0}^{N-1} x(k) \cos\left(\frac{(2k+1)m\pi}{2N}\right)$$

where

$$\alpha(m) = \begin{cases} \frac{1}{\sqrt{N}}, & k = 0 \\ \sqrt{\frac{2}{N}}, & \text{otherwise} \end{cases} \tag{14}$$

In matrix form, the DCT coefficient vector can be written as

$$s = \Phi^{-1} x \tag{15}$$

where the (m, n)th element of the matrix Φ is given by

$$\Phi_{m,n} = \begin{cases} \frac{1}{\sqrt{N}}, & m = 0, \quad 0 \le n \le N-1 \\ \sqrt{\frac{2}{N}} \cos\frac{\pi(2n+1)m}{2N} & \text{for } 1 \le m \le N-1, \quad 0 \le n \le N \end{cases} \tag{16}$$

DCT is quite universal and provides an easy way to perform the sparsity transformation (note that we generally would use a DCT type basis in turbulent flows if the flow is homogenous and/or periodic in a particular direction).

The DCT coefficients of U and V in one snapshot are shown in Fig. 3, from which we see that the original velocity field data are considerably sparse in the DCT domain. Taking DCT as the basis in which the velocity data are sparse, in Fig. 4b, we plot the reconstructed velocity field based on the CS mechanism with only 10% of the original samples ($M = 0.1N$) in a particular snapshot. Figure 4a shows the original instantaneous velocity field. When compared with the original instantaneous velocity field, we can see that 10% of CS measurements capture the large-scale (low-dimensional) features of the velocity field.

We further illustrate the performance of the velocity reconstruction based on CS in terms of the normalized mean-squared error (MSE), which is defined as

$$\text{MSE} = \frac{\|u_i - \hat{u}_i\|_2}{\|u_i\|_2} \tag{17}$$

where, for $U(V)$ data in one snapshot, $u_i(v_i)$ is the original velocity vector, and $\hat{u}_i(\hat{v}_i)$ is the estimated velocity vector via CS. Here, $\|\cdot\|_2$ signifies the l_2-norm of the column vector. In Fig. 5, we plot the MSE of U and V data for different snapshots when the reconstruction is performed via the CS-based scheme (as described previously) as well

[††]Data available online at http://users.ece.gatech.edu/justin/l1magic [retrieved April 2014].

as when the reconstruction is performed via the POD/PCA method applied to the entire dataset. From Fig. 5, we can see that the reconstruction performance with CS is not as good as with POD/PCA, especially when the compression ratio is small; however, the performance degradation is not significant. As M/N increases, the performance of the CS-based scheme converges to that based on the POD/PCA-based scheme. It is worth mentioning that, in contrast to the POD/PCA scheme, which requires large computation at the compression stage (finding covariance matrices, eigenfunctions, and significant modes, time support etc.), the compression stage of the CS-based scheme is very simple and universal. More specifically, the POD/PCA scheme requires processing of all the time and spatial data to find the covariance matrices and the significant eigenvalues, whereas the CS measurement scheme directly acquires lower-dimensional data for a particular snapshot via a simple random projection scheme. Thus, although the CS-based scheme does not show the same performance as the POD/PCA-based scheme at very small compression ratios, it would still appear to be attractive due to these advantages at the compression stage. However, for real-time flow control, POD/PCA only needs to be computed once, based on which stochastic estimation could be used for POD/PCA mode feedback.

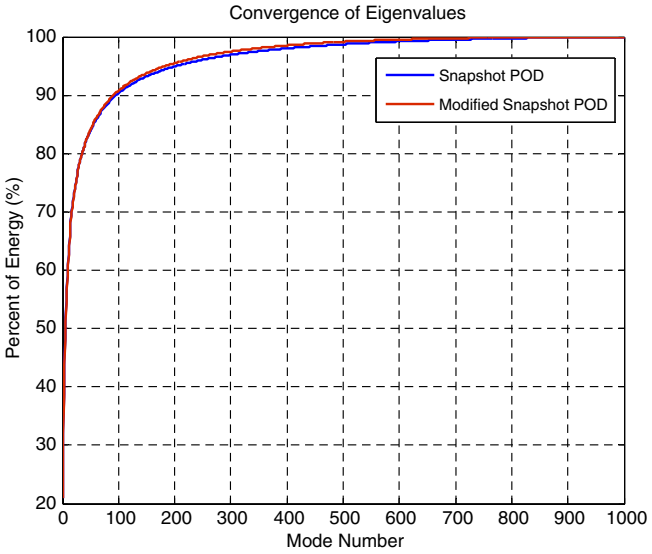


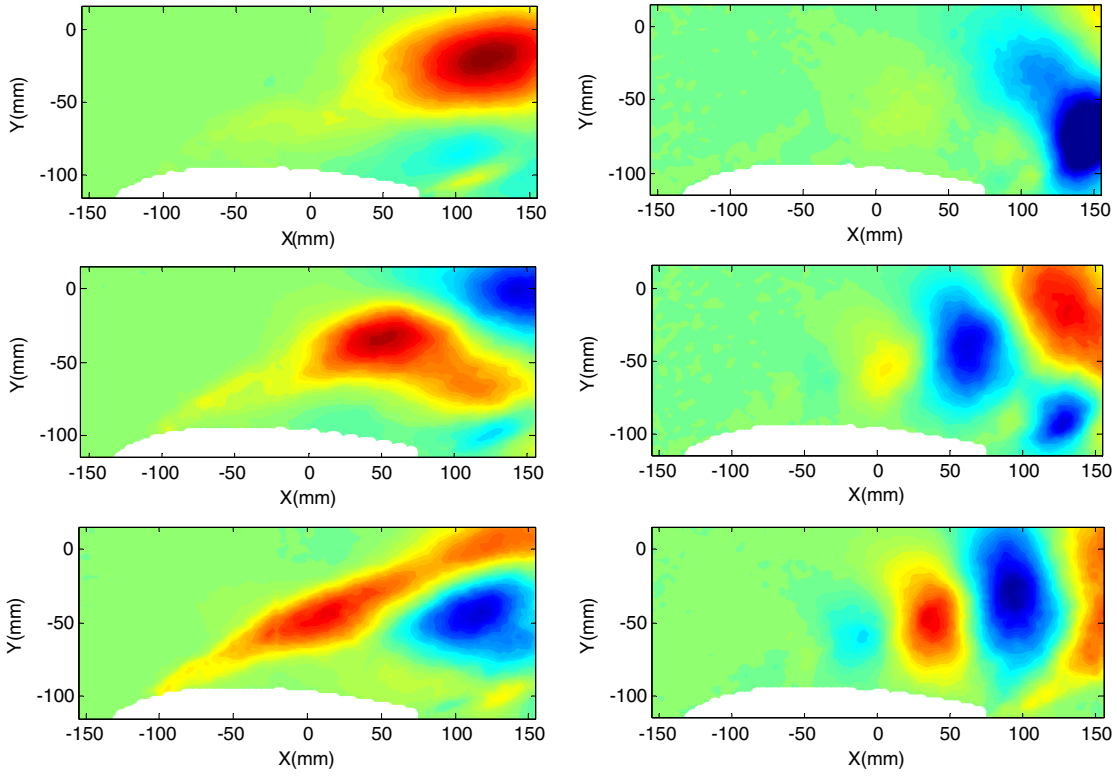
Fig. 11 Convergence of the eigenvalues, using the modified snapshot POD/PCA.

Figure 6 shows the MSE of the U and V components for 1000 snapshots with half of the total samples ($M = 0.5N$). We can see that, for all of the PIV snapshots, the difference between the actual u_i and estimated $\hat{u}_i$ is less than 13% of the original u_i.

B. Performance of Compressive-Sensing Based Reconstructions with Proper Orthogonal Decomposition Bases

In this section, we use the basis obtained from snapshot POD/PCA, as explained in Table 1, as the sparsifying basis instead of the DCT. We use the term "snapshot POD/PCA basis" to refer to it in the rest of this paper. Because the POD/PCA provides an optimal basis for describing the particular physics (in terms of turbulent kinetic energy), the expectation is that the CS approach should improve when POD/PCA is used to determine the sparsifying basis. In this case, the size of the POD/PCA basis coefficient s will be T instead of N as in the DCT basis because the spatial coefficient Φ is an N by T matrix in the snapshot POD/PCA. Therefore, M should be less than T instead of N, as in the DCT basis, to make sure that the size of y is smaller than the size of s.

In Fig. 7, we plot the energy distribution versus the POD/PCA modes. Note that, with the snapshot POD/PCA approach, the total number of modes is equal to the smallest dimension of the original data matrix. In our study, the original velocity data (U and V) is stored in an N by T matrix with $N = 8320$ and $T = 1000$. Thus, the number of total POD/PCA modes is 1000. As can be seen in Fig. 7, approximately 70% of the energy is retrieved with approximately 10% of the modes, and more than 30% of the total energy is contained in the first two modes, which is quite significant in a fully developed turbulent flow.



a) U component **b) V component**

Fig. 12 First three modes in modified snapshot POD/PCA basis.

Figure 8 presents the spatial behavior of the first three POD/PCA modes for the U and V velocity components. We can see large-scale structures, which contain most of the energy of the flow over the airfoil. Also, note that the spatial structures of the two components Φ_U and Φ_V are quite different, reflecting the different characteristics of each component (more discussion on the POD/PCA analysis of this data can be found in [8]).

The original and reconstructed first 100 coefficients (where most of the significant coefficients are located) at a given time snapshot with respect to the snapshot POD/PCA basis for different compression ratios are shown in Fig. 9. We see that the reconstructed $\hat{s}$ and $\hat{r}$ based on CS, taking the snapshot POD/PCA generated basis as the sparsifying basis, gradually become closer to the original s and r with increasing values of M. The reconstructed velocity field approaches the actual velocity field with an increase in the compression ratio as seen in Fig. 10, especially in the flow separation region as shown.

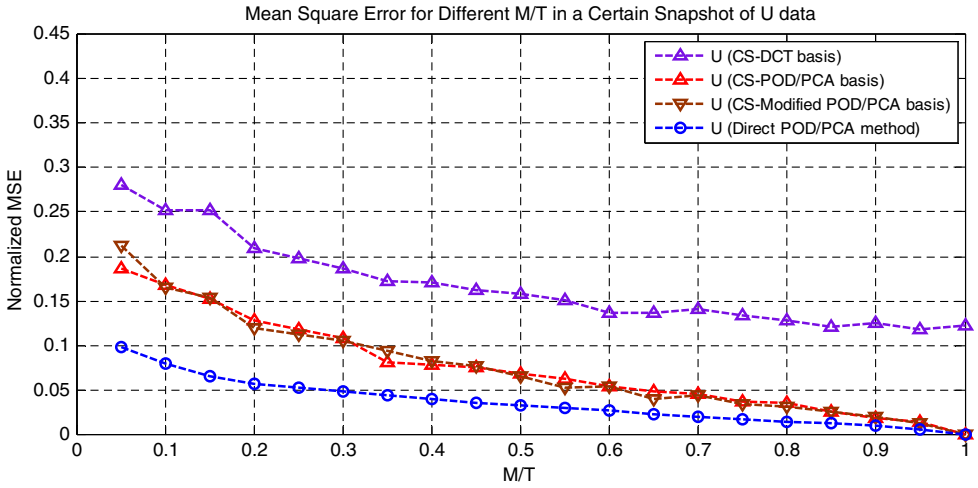
C. Modified Snapshot Proper Orthogonal Decomposition/Principal Component Analysis Basis

As described in the previous section, with CS-based reconstruction using the snapshot POD/PCA basis as the sparsifying basis, we compute the covariance matrix of the data offline, which is done with a certain amount of training data from the original dataset (before taking random projections). To further reduce the complexity of computing the covariance matrix of the data, we consider computing the covariance matrix after application of Eq. (11), i.e., we reduce the number of spatial grid points in a randomly sampled manner. For this, the number of compressive measurements M should be larger than T (here, we set it to $M = 0.2, N = 1664 > T$) to assure the accuracy of the correlation matrix in snapshot POD/PCA; however, for the measurement matrix in the later step, M should be smaller than T, same as that in the original snapshot POD/PCA basis.

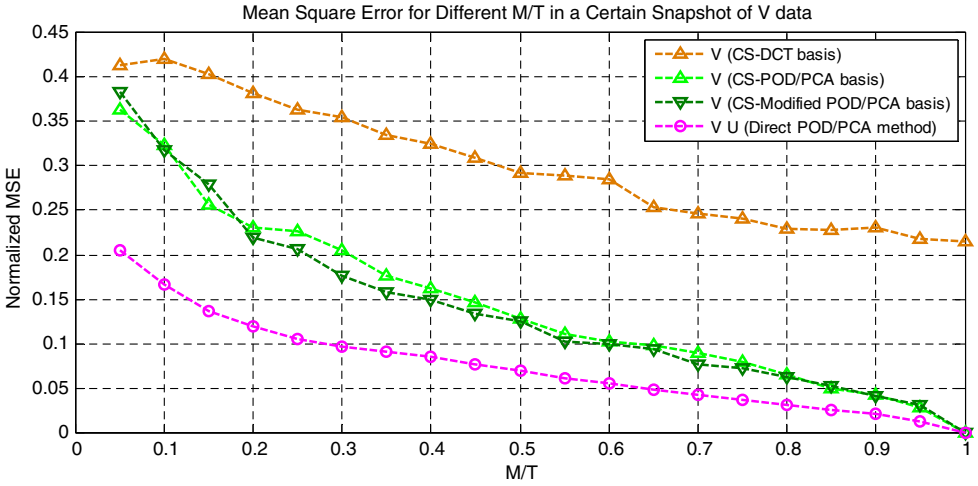
Figure 11 shows the convergence of energy in the modified snapshot POD/PCA basis. We can see that there is not much difference compared to that obtained with the original snapshot POD/PCA basis, even though the velocity data have been compressed spatially using the random projection before computing the correlation.

The first three modes of the modified snapshot POD/PCA basis in U and V components computed in this case are shown in Fig. 12. We see the large-scale structures, having most of the energy in the flow, are very similar to those in Fig. 8, except the small difference in the third mode of the U component. Hence, this might be a useful approach for extracting the snapshot POD/PCA basis while requiring much less storage of the original PIV dataset.

For the modified snapshot POD/PCA basis, the original and reconstructed sparse coefficients for different compression ratios for U and V separately reconstructed $\hat{s}$ and $\hat{r}$ do not match the original s and r as well as the ones in the snapshot POD/PCA basis. However, there is not much difference in terms of the reconstructed $\hat{u}$, $\hat{v}$ compared with the original u, v. The reason is that, although s, r from the modified POD/PCA basis are somewhat different from the original POD/PCA basis, the optimization process still selects the significant coefficients from the POD/PCA domain for reconstruction. As a result, the reconstructed $\hat{s}$ or $\hat{r}$ do not deviate much from the original one. Thus, the error between the reconstructed velocity and the original velocity remains the same, as shown in the error comparison of different



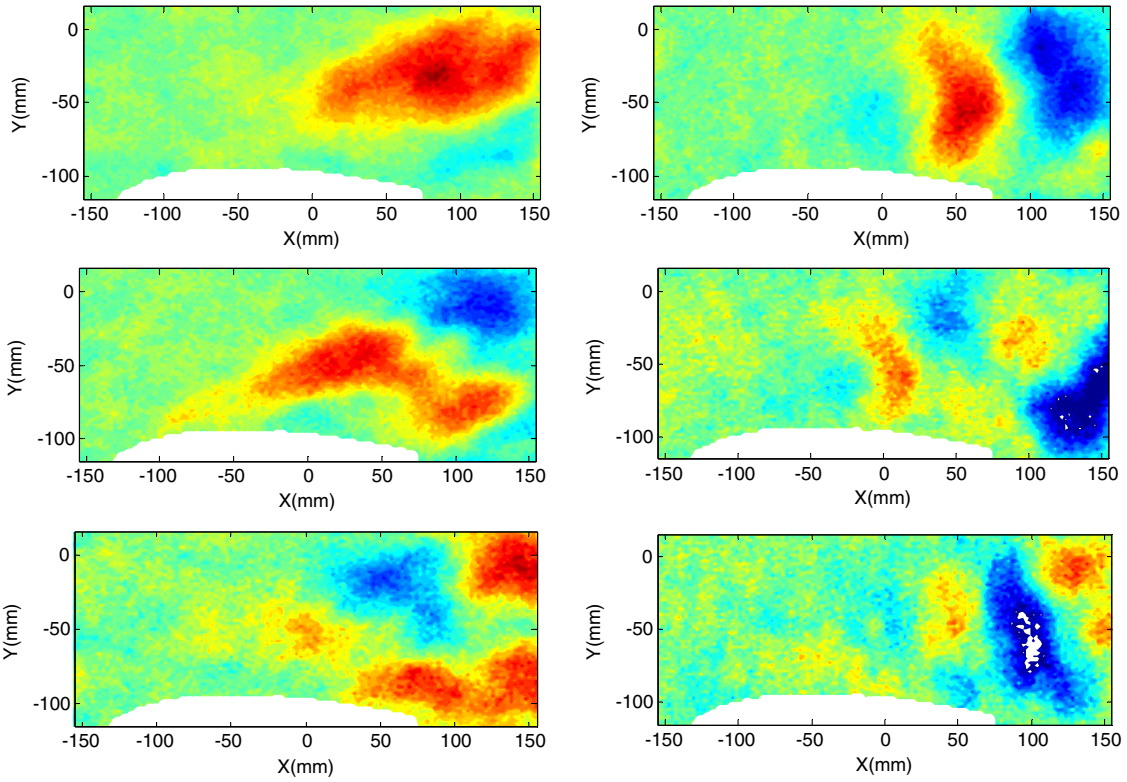
a) U component



b) V component

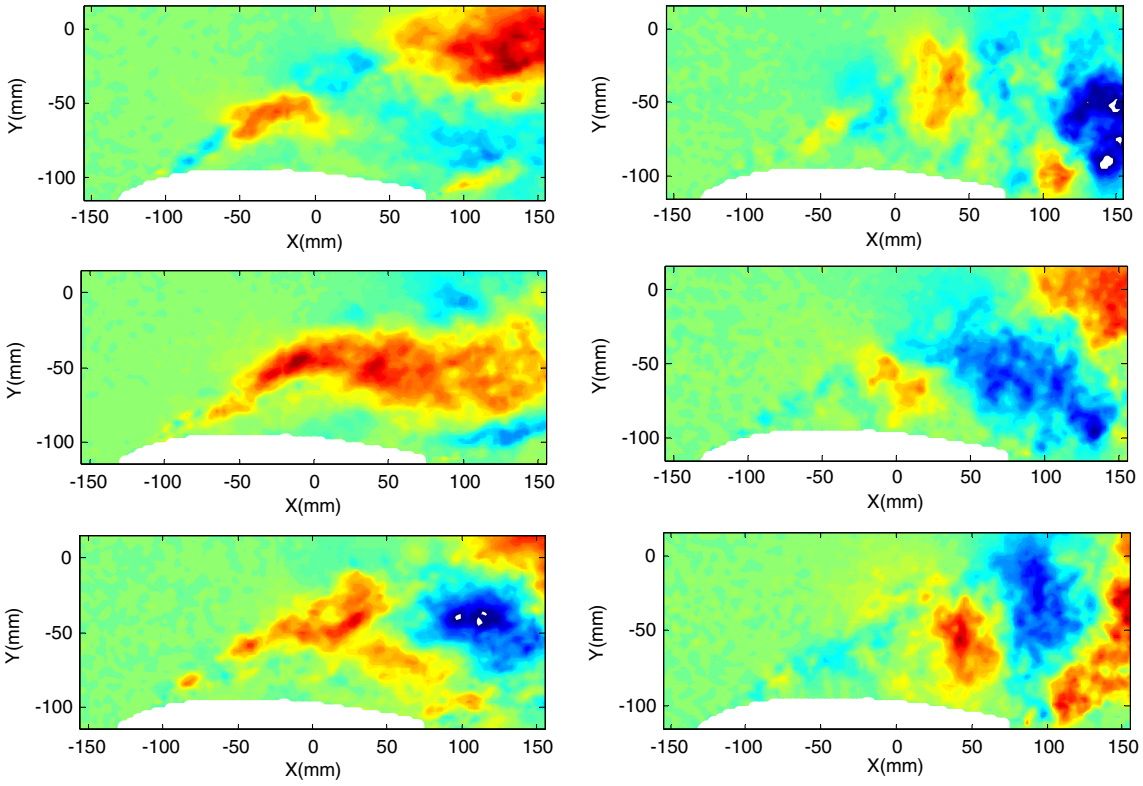
Fig. 13 Normalized MSE comparison of CS-based reconstruction with DCT basis, original snapshot POD/PCA basis, modified snapshot POD/PCA basis, and direct snapshot POD/PCA method.

schemes in Fig. 13. While taking M/T as the independent variable, we clearly see that the POD/PCA-based approaches (CS-POD/PCA, CS-modified POD/PCA, and direct POD/PCA method) behave better with faster convergence rates because we take a lower-dimension T instead of the number of grid points N to solve the eigenvalue problem and pick out the significant structures in the velocity field more efficiently. Although the CS-based reconstruction does not reach the same level as the POD/PCA-based reconstruction, we still consider it to be a useful method to analyze the physical data for its universality, which depends only on the number of grid points for the DCT basis. The CS-based



a) U component b) V component

Fig. 14 First three modes of POD spatial eigenfunctions with CS-based reconstruction using 20 snapshots.



a) U component b) V component

Fig. 15 First three modes of POD spatial eigenfunctions with direct POD/PCA using 20 snapshots.

Table 2 Modal correlation of the POD spatial eigenfunctions, at the same mode number

	CS filtered POD (full snapshots POD)		20 snapshots POD (full snapshots POD)	
Modal correlation	*U* component	*V* component	*U* component	*V* component
First mode/first mode	0.9062	0.3143	0.7611	0.7658
Second mode/second mode	0.7558	0.2191	0.6295	0.7377
Third mode/third mode	0.3067	0.4733	0.8401	0.8534
Fourth mode/fourth mode	0.5112	0.0598	0.3693	0.6287
Fifth mode/fifth mode	0.2952	0.4402	0.0603	0.0080

method taking POD/PCA as the sparsifying basis also provides a novel view of combining CS and POD/PCA and taking advantage of both. Unlike the DCT basis, the POD/PCA basis or modified POD/PCA basis requires the whole physical field information; however, the total computational cost is dramatically reduced with additional offline computation more related to the velocity field and flow characteristics in terms of the turbulence in the separation region.

D. Extracting Proper Orthogonal Decomposition Modes from Compressed Snapshots with Limited Time Support

Here, we present a methodology for POD/PCA mode extraction based on CS preprocessing (filtering) using only the DCT basis. The CS preprocessing (CS-POD/PCA) approach is exploited to extract the significant POD/PCA spatial eigenfunctions under conditions with limited time support. The idea here is to use the CS approach with the DCT basis to filter the PIV snapshots before applying POD/PCA to deal with real-world applications where only limited time support PIV data (or large-eddy simulation and direct numerical simulation datasets without great time support) are available but yet POD/PCA modes are desired for various reasons. Situations where this might occur for PIV datasets include full-scale PIV measurements: for large wind turbines in the field, on aircraft wings, and in large wind tunnels where seeding can be difficult, etc. Klenk and Heidemann[‡‡] have proposed a similar approach to calculate principal components based on CS for high-dimensional heart sound data. For the first attempt, CS is applied with a compression ratio of 10% to 20 snapshots individually. A compression ratio of 10% is chosen based on the previous analysis as observed in Fig. 4. Once the compression and reconstruction are implemented, the snapshot POD/PCA is then performed on those 20 preprocessed snapshots.

The first three modes of POD spatial eigenfunctions using the reconstructed data from CS are shown in Fig. 14. When comparing these results to the direct snapshot POD/PCA (without CS) on the same 20 snapshots shown in Fig. 15, we see that the eigenfunctions are substantially different. When comparing both of these results with the full time support (1000 snapshots) POD/PCA results previously applied as shown in Fig. 8, we see that the results for the CS filtered case shown in Fig. 14 compare quite well for the first couple of eigenfunctions extracted for the streamwise velocity and substantially better than those from the 20 mode direct POD/PCA approach. For the *V* component, however, the CS preprocessing method and subsequently extracted POD/PCA do not compare well with the full time support data.

Table 2 presents a quantitative comparison between the POD/PCA modes obtained from the CS filtered 20 snapshot data and the unfiltered 20 snapshot data using the modes from the full time support as a reference. As noted in the earlier discussions of Figs. 14 and 15, the CS filtered POD approach performs well for the *U* component but not for the *V* component. This is consistent with what was described in Fig. 3, where the DCT coefficients of the streamwise velocity are more sparse than those of the wall-normal velocity. In future work, we will perform vector CS filtering and anticipate that the additional information extracted from the strong *UV* correlations will improve the CS filtered POD/PCA results for the wall-normal velocity component. The convergence of the POD eigenvalues for the three different approaches is given in Fig. 16. The energy within the streamwise component converges faster than the wall-normal component, as expected. The CS filtered approach has a similar performance with the full 1000 snapshots for the first three modes with respect to energy convergence. Overall, these initial results appear promising, and future work will explore ways to improve the performance of the CS filtered POD/PCA approach and develop a more clear understanding of its limitations.

[‡‡]Klenk, S., and Heidemann, G., "A New Method for Principal Component Analysis of High-Dimensional Data Using Compressive Sensing," 2010, http://www.sfb716.uni-stuttgart.de/uploads/tx_vispublications/pca.pdf [retrieved April 2010].

V. Conclusions

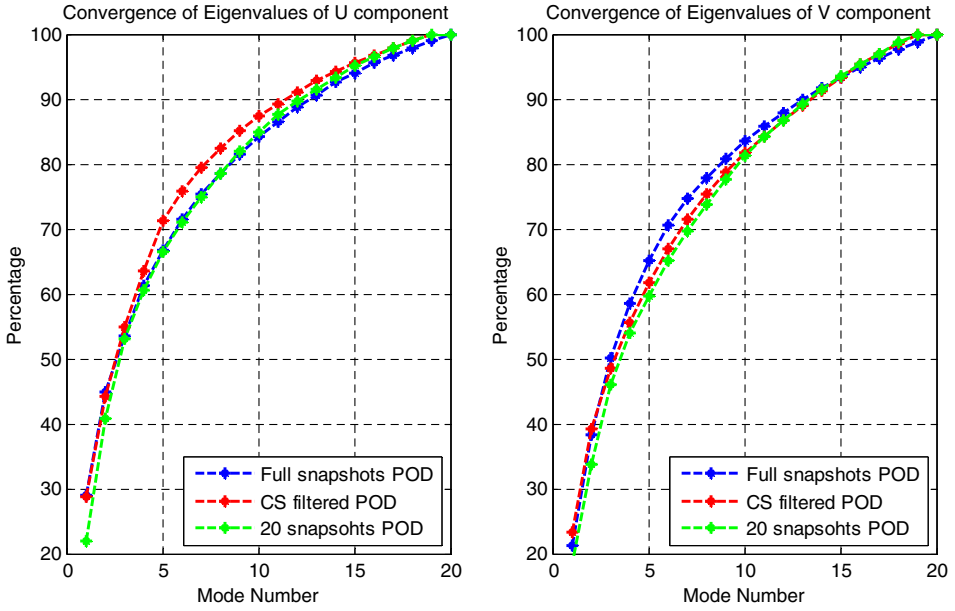
In this study, the feasibility and effectiveness of using compressive sensing (CS) for compression and reconstruction of large airfoil datasets is shown. With the CS scheme, using an orthogonal basis such as DCT in which the signals are sparse, reconstruction can be achieved relatively accurately through a small number of samples obtained via random projections. CS is an effective compression process without the time support as required for many traditional compression schemes such as POD/PCA. The CS-based sampling scheme is universal, and it can also be expected to be readily applied to other types of data acquisition.

The reconstruction performance is illustrated with CS taking different bases in which the data are sparse and compared the performance with the traditional snapshot POD/PCA-based reconstruction. When DCT is used as the sparsifying basis, acceptable performance with CS is achieved compared to the POD/PCA-based reconstruction. The reconstruction performance with CS is further improved by taking the POD/PCA basis as the sparsifying basis, which, however, has to be computed offline with an additional computational cost compared to the DCT. Examination of the two reconstruction approaches shows that the snapshot POD/PCA offers a more efficient reconstruction, taking only 10% of the computational time as compared to the DCT. The modified snapshot POD/PCA provides a unique technique for compressing the spatial data using random projections to decrease the number of spatial grid points, resulting in a reduced dimension of the eigenvalue problem for extraction of the snapshot POD/PCA basis.

The POD/PCA basis extraction with CS preprocessing and limited time support has been shown to work well for the first two streamwise velocity eigenfunctions but not for the wall-normal velocity eigenfunctions. This is due in part to the wall-normal velocity data not being as sparse in the DCT basis as the streamwise velocity component. Future work will include application of vector CS filtering to use the additional information available from the strong UV correlations along with exploring optimal compression ratios and number of snapshots to give the best results.

Acknowledgments

The authors gratefully acknowledge the support of the U.S. Air Force Office of Scientific Research through a Phase 1 Small Business Technology Transfer Program (award number FA9550-12-C-0002) and CASE: The Center for Advanced Systems and Engineering, a NYSTAR center for advanced technology at Syracuse University.



a) U component **b) V component**

Fig. 16 Convergence of the eigenvalues for full snapshots POD (first 20 modes), CS filtered POD, and 20 snapshots POD.

References

[1] Donoho, D. L., "Compressed Sensing," *IEEE Transactions on Information Theory*, Vol. 52, No. 4, April 2006, pp. 1289–1306. doi:10.1109/TIT.2006.871582

[2] Candès, E. J., Romberg, J. K., and Tao, T., "Stable Signal Recovery from Incomplete and Inaccurate Measurements," *Communications on Pure and Applied Mathematics*, Vol. 59, No. 8, 2006, Aug. 2006, pp. 1207–1223.

[3] Carlson, H., Glauser, M., Higuchi, H., and Young, M., "POD Based Experimental Flow Control on a NACA-4412 Airfoil," AIAA Paper 2004-0575, 2004.

[4] Holmes, P. J., Lumley, J. L., and Berkooz, G., "Part 3. Proper Orthogonal Decomposition," *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge Univ. Press, Cambridge, England, U.K., 1996, pp. 86–128.

[5] Taylor, J. A., and Glauser, M. N., "Toward Practical Flow Sensing and Control via POD and LSE Based Low-Dimensional Tools," *Journal of Fluids Engineering*, Vol. 126, No. 3, 2004, pp. 337–345; also American Soc. of Mechanical Engineers Paper FEDSM2002-31416, New York, 2002. doi:10.1115/1.1760540

[6] Low, K. R., Berger, Z. P., Kostka, S., El-Hadidi, B., Gogineni, S., and Glauser, M. N., "A Low-Dimensional Approach to Closed-Loop Control of a Mach 0.6 Jet," *Experiments in Fluids*, Vol. 54, No. 4, 2013, pp. 1–17. doi:10.1007/s00348-013-1484-8

[7] Tinney, C., Ukeiley, L., and Glauser, M., "Low-Dimensional Characteristics of a Transonic Jet. Part 2: Estimate and Far-Field Prediction," *Journal of Fluid Mechanics*, Vol. 615, Nov. 2008, pp. 53–92. doi:10.1017/S0022112008003601

[8] Pinier, J. T., Ausseur, J. M., Glauser, M. N., and Higuchi, H., "Proportional Closed-Loop Feedback Control of Flow Separation," *AIAA Journal*, Vol. 45, No. 1, Jan. 2007, pp. 181–190. doi:10.2514/1.23465

[9] Bai, Z., Wimalajeewa, T., Berger, Z., Glauser, M., and Varshney, P., "Physics Based Compressive Sensing Approach Applied to Airfoil Data Collection and Analysis," *Bulletin of the American Physical Society*, Vol. 57, 2012.

[10] Glauser, M., Varshney, P., Wimalajeewa, T., Bai, Z., Berger, Z., and Wang, G., "Application of Compressive Sensing to NACA 4412 PIV Data," *51st AIAA Aerospace Sciences Meeting*, AIAA Paper 2013-0772, Jan. 2013.

[11] Bourguignon, J. L., Tropp, J. A., Sharma, A. S., and Mckeon, B. J., "Compact Representation of Wall-Bounded Turbulence Using Compressive Sampling," *Physics of Fluids*, Vol. 26, No. 1, 2014, Paper 015109. doi:10.1063/1.4862303

[12] Tayler, A. B., Holland, D. J., Sederman, A. J., and Gladden, L. F., "Exploring the Origins of Turbulence in Multiphase Flow Using Compressed Sensing MRI," *Physical Review Letters*, Vol. 108, No. 26, 2012, Paper 264505. doi:10.1103/PhysRevLett.108.264505

[13] Bright, I., Lin, G., and Kutz, J. N., "Compressive Sensing Based Machine Learning Strategy for Characterizing the Flow Around a Cylinder with Limited Pressure Measurements," *Physics of Fluids*, Vol. 25, No. 12, 2013, Paper 127102. doi:10.1063/1.4836815

[14] Schneiders, J. F. G., Dwight, R. P., and Scarano, F., "Time-Supersampling of 3D-PIV Measurements with Vortex-in-Cell Simulation," *Experiments in Fluids*, Vol. 55, No. 3, 2014, pp. 1–15. doi:10.1007/s00348-014-1692-x

[15] Huang, X., "Compressive Sensing and Reconstruction in Measurements with an Aerospace Application," *AIAA Journal*, Vol. 51, No. 4, 2013, pp. 1011–1016. doi:10.2514/1.J052227

[16] Mathelin, L., Pastur, L., and Maître, O. L., "A Compressed-Sensing Approach for Closed-Loop Optimal Control of Nonlinear Systems," *Theoretical and Computational Fluid Dynamics*, Vol. 26, Nos. 1–4, 2012, pp. 319–337. doi:10.1007/s00162-011-0235-9

[17] Tu, J. H., Rowley, C. W., Kutz, J. N., and Shang, J. K., "Spectral Analysis of Fluid Flows Using Sub-Nyquist-Rate PIV Data," *Experiments in Fluids*, Vol. 55, No. 9, 2014, pp. 1–13. doi:10.1007/s00348-014-1805-6

[18] Jovanovic, M. R., Schmid, P. J., and Nichols, J. W., "Low-Rank and Sparse Dynamic Mode Decomposition," *Center for Turbulence Research Annual Research Briefs*, Center for Turbulence Research, Stanford Univ., Stanford, CA, 2012, pp. 139–152.

[19] Brunton, S. L., Proctor, J. L., and Kutz, J. N., "Compressive Sampling and Dynamic Mode Decomposition," 2013, http://arxiv.org/pdf/1312.5186.pdf [retrieved April 2014].

[20] Baraniuk, R. G., "Compressive Sensing," *IEEE Signal Processing Magazine*, Vol. 24, No. 4, 2007, pp. 118–120, 124.

[21] Candes, E. J., and Wakin, M. B., "An Introduction to Compressive Sampling," *IEEE Signal Processing Magazine*, Vol. 25, No. 2, March 2008, pp. 21–30. doi:10.1109/MSP.2007.914731

[22] Tropp, J. A., and Gilbert, A. C., "Signal Recovery from Random Measurements via Orthogonal Matching Pursuit," *IEEE Transactions on Information Theory*, Vol. 53, No. 12, Dec. 2007, pp. 4655–4666. doi:10.1109/tit.2007.909108

[23] Lustig, M., Donoho, D., and Pauly, J. M., "Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging," *Magnetic Resonance in Medical Sciences*, Vol. 58, No. 6, 2007, pp. 1182–1195. doi:10.1002/Mrm.21391

[24] Lumley, J. L., "The Structure of Inhomogeneous Turbulent Flows," *Atmospheric Turbulence and Radio Wave Propagation*, edited by Yaglom, A. M., and Tatarsky, V. I., Nauka, Moscow, 1967, pp. 166–178.
[25] Sirovich, L., "Turbulence and the Dynamics of Coherent Structures. Parts 1, 2, and 3," *Quarterly of Applied Mathematics*, Vol. 45, No. 3, 1987, pp. 561–590.
[26] Fernndez, M., Aridgides, T., and Sadjadi, F., "Practical Optimal Processing in Hyperdimensional Spaces via Domain-Reducing Mappings," *Proceedings of SPIE: Automatic Target Recognition 21*, Vol. 8049, May 2011, Paper 80490R. doi:10.1117/12.885260
[27] Candès, E. J., "The Restricted Isometry Property and Its Implications for Compressed Sensing," *Comptes Rendus Mathematique*, Vol. 346, Nos. 9–10, 2008, pp. 589–592. doi:10.1016/j.crma.2008.03.014
[28] Davenport, M. A., "Random Observations on Random Observations: Sparse Signal Acquisition and Processing," Ph.D. Dissertation, Rice Univ., Houston, TX, Aug. 2010.
[29] Wimalajeewa, T., and Varshney, P. K., "Performance Bounds for Sparsity Pattern Recovery with Quantized Noisy Random Projections," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 6, No. 1, Feb. 2012, pp. 43–57. doi:10.1109/JSTSP.2011.2175700
[30] Baraniuk, R., Davenport, M., Devore, R., and Wakin, M., "A Simple Proof of the Restricted Isometry Property for Random Matrices," *Constructive Approximation*, Vol. 28, No. 3, 2008, pp. 253–263. doi:10.1007/s00365-007-9003-x
[31] Donoho, D. L., and Tsaig, Y., "Fast Solution of L1-Norm Minimization Problems When the Solution May Be Sparse," *IEEE Transactions on Information Theory*, Vol. 54, No. 11, Nov. 2008, pp. 4789–4812. doi:10.1109/tit.2008.929958

A. Naguib
Associate Editor