Unsteady Flow Sensing and Estimation via the Gappy Proper Orthogonal Decomposition

Karen Willcox
Aerospace Computational Design Laboratory
Massachusetts Institute of Technology
Cambridge, MA 02139
kwillcox@mit.edu

***Abstract*— The proper orthogonal decomposition (POD) has been widely used in fluid dynamic applications for extracting dominant flow features. The "gappy" POD is an extension to this method that allows the consideration of incomplete data sets. In this paper, the gappy POD is extended to handle unsteady flow reconstruction problems, such as those encountered when limited flow measurement data is available. In addition, a systematic approach for effective sensor placement is formulated within the gappy framework. Two applications are considered. The first aims to reconstruct the unsteady flow field using a small number of surface pressure measurements for a subsonic airfoil undergoing plunging motion. The second considers estimation of POD modal content of a cylinder wake flow for active control purposes. In both cases, using the dominant POD basis vectors and a small number of sensor signals, the gappy approach is found to yield accurate flow reconstruction results.**

I. Introduction

The proper orthogonal decomposition (POD), also known as Karhunen Loéve expansion and principle components analysis, has been widely used for a broad range of applications, including derivation of reduced-order dynamical models [1], image processing [2] and pattern recognition [3]. The POD method computes a set of basis vectors that capture the dominant structures of the system. For example, for fluid dynamic applications, the dominant POD basis vectors correspond to the most energetic flow modes in the system.

Sirovich introduced the method of snapshots [4] as a way to efficiently determine the POD modes for large problems, such as those encountered in computational fluid dynamic (CFD) applications. A set of instantaneous flow solutions, or "snapshots" is obtained from a simulation of the CFD method. These snapshots are then used to compute the POD basis vectors, which yield a representation of the data that is optimal in the sense that, for any given basis size, the two-norm of the error between the original and reconstructed snapshots is minimized. Reduced-order dynamic models can be derived by projecting the CFD model onto the reduced space spanned by the POD modes [5], [6], [1].

The concept of using active control to enhance the stability properties of an unsteady flow has been addressed for several applications, for example in [7], [8], [9], [10]. In order to derive control models that will be effective, it is vital that the relevant unsteady flow dynamics are captured accurately. A high-fidelity CFD code can offer the degree of flow resolution that is required; however, for control design it is imperative that the flow model have a low number of states. The POD provides not only a way to obtain accurate low-order models for control design, but also a systematic means to identify the most dominant flow structures. One approach for flow controller design, demonstrated to yield effective results, is to control the dominant POD modes [9], [11].

To achieve active flow control in practice, the issue of flow sensing must also be addressed. Using a POD-based control approach, a strategy is required to accurately estimate the POD modal content in real time from a limited number of sensor measurements. In addition, the question of where to best place the sensors in order to achieve this estimation must be addressed. In [11], flow control of a cylinder wake was considered. Two sets of POD basis vectors were derived using collections of flow snapshots, obtained from a CFD simulation of the problem and from particle image velocimetry (PIV) measurements of an experimental setup. In each case, a linear stochastic estimator (LSE) was used to provide a mapping from the velocity data (computational or experimental) to the POD modal content. The sensor locations were chosen according to a heuristic procedure that placed them at spatial maxima and minima of each POD mode.

Here, an alternate approach is proposed for estimating the modal coefficients, based on the gappy POD method. This method was developed by Everson and Sirovich [12], and is a modification of the basic POD method that handles incomplete or "gappy" data sets. Given a set of POD modes, an incomplete data vector can be reconstructed by solving a small linear system. Moreover, if the snapshots themselves are damaged or incomplete, an iterative method can be used to derive the POD basis. This method has been successfully applied for reconstruction of images, such as human faces, from partial data in [12]. In [13], the gappy POD was applied for reconstruction of airfoil pressure fields from limited surface measurements. In that work, it was shown that the entire pressure field for subsonic and transonic inviscid flows could be reconstructed using just a handful of POD modes and a small number of surface measurements.

In this paper, the gappy POD method will first be described and extended to handle unsteady flow reconstruction problems. A quantitative metric for placing sensors will then be developed using the gappy formulation. Results will then

be presented for two test cases. The first considers the two-dimensional, linearized Euler equations to analyze a subsonic airfoil operating in unsteady plunging motion. The second example considers a Navier-Stokes simulation of flow about a circular cylinder at low Reynolds number. Finally, conclusions are drawn and directions for future work are discussed.

II. Proper Orthogonal Decomposition

A. Standard POD basis

Before describing the gappy POD procedure, the standard approach to computing POD basis vectors via the method of snapshots is first reviewed [4]. Consider the collection of m flow snapshots, $\left\{U^k\right\}_{k=1}^m$, where U^k is a vector containing the flow solution at a time t_k. The correlation matrix R is formed by computing the inner product between every pair of snapshots

$$R_{ik} = \frac{1}{m}\left(U^i, U^k\right), \tag{1}$$

where $\left(U^i, U^k\right)$ denotes the inner product between U^i and U^k. The eigenvalues λ_i and eigenvectors ψ^i of R are then computed. The j^{th} POD basis vector, Φ^j, is given by a linear combination of snapshots

$$\Phi^j = \sum_{i=1}^{m} \psi_i^j U^i, \tag{2}$$

where ψ_i^j denotes the i^{th} element of the j^{th} eigenvector. The magnitude of the j^{th} eigenvalue, λ_j, describes the relative importance of the j^{th} POD basis vector. This importance is commonly quantified by defining the relative energy content, E_j, for each basis vector j as

$$E_j = \frac{\lambda_j}{\sum_{i=1}^{m} \lambda_i} \tag{3}$$

where the term "energy" refers to a measure in the two-norm.

B. Gappy POD

The gappy POD procedure uses a POD basis to reconstruct missing, or "gappy" data. This procedure was developed in [12] and can be described as follows. The first step is to define a mask vector, which describes for a particular flow vector where data is available and where data is missing. For the flow solution U^k, the corresponding mask vector n^k is defined as follows:

$$\begin{aligned} n_i^k &= 0 \text{ if } U_i^k \text{ is missing} \\ n_i^k &= 1 \text{ if } U_i^k \text{ is known} \end{aligned}$$

where U_i^k denotes the i^{th} element of the vector U^k. Pointwise multiplication is defined as $\left(n^k, U^k\right)_i = n_i^k U_i^k$. Then the gappy inner product is defined as $(u,v)_n = ((n,u),(n,v))$, and the induced norm is $(\|v\|_n)^2 = (v,v)_n$.

Let $\left\{\Phi^k\right\}_{k=1}^m$ be the standard POD basis for the snapshot set $\left\{U^k\right\}_{k=1}^m$, where all snapshots are completely known. Let g be another solution vector that has some elements missing, with corresponding mask vector n. Assume that there is a need to reconstruct the full or "repaired" vector from the incomplete vector g. Assuming that the vector g represents a solution whose behavior can be characterized with the existing snapshot set, the intermediate repaired vector $\tilde{g}$ can be represented in terms of p POD basis functions as follows:

$$\tilde{g} \approx \sum_{i=1}^{p} b_i \Phi^i \tag{4}$$

To compute the POD coefficients b_i, the error, E, between the original and repaired vectors must be minimized. The error is defined as

$$E = \|g - \tilde{g}\|_n^2 \tag{5}$$

using the gappy norm so that only the original existing data elements in g are compared. The coefficients b_i that minimize the error E can be found by differentiating (5) with respect to each of the b_i in turn. This leads to the linear system of equations

$$Mb = f \tag{6}$$

where

$$M_{ij} = \left(\Phi^i, \Phi^j\right)_n \tag{7}$$

and

$$f_i = \left(g, \Phi^i\right)_n \tag{8}$$

Solving equation (6) for b and using (4), the intermediate repaired vector $\tilde{g}$ can be obtained. Finally, the complete g is reconstructed by replacing the missing elements in g by the corresponding repaired elements in $\tilde{g}$, i.e. $g_i = \tilde{g}_i$ if $n_i = 0$.

While not discussed here, we also note that if the original snapshot ensemble has incomplete data, the POD basis vectors can be computed using an iterative gappy approach, as described in [12].

C. Gappy POD for unsteady flows

It is relatively straightforward to extend the gappy POD algorithm for consideration of unsteady flows. Assume we have a sequence of T sensor measurements, $\left\{g^i\right\}_{i=1}^T$, where g^i corresponds to a gappy flow solution at time t^i. At each timestep, one can solve the gappy problem given by equation (6) to determine the corresponding POD basis vector modal content. The matrix M depends only on the POD basis vectors and the mask vector. For a given sensor configuration, the mask vector is fixed and M is thus not a function of time. Its inverse can therefore be precomputed to yield an efficient implementation. The time-dependent gappy problem can be stated as

$$Mb^i = f^i \tag{9}$$

where b^i contains the POD modal coefficients at time t^i. The matrix M and vector f are defined as before in (7) and (8), except now g, and therefore f, varies with time.

III. Sensor placement problem

We consider the problem of placing N sensors in a flow. The sensor locations should be chosen so that the required flow information can be obtained. In particular, in the active flow control framework, we are concerned with using sensor data to determine the POD modal content of the flow.

This problem can be cast in the gappy framework. We assume that POD basis vectors have been computed, which are representative of the flows under consideration. A particular set of sensor measurements then corresponds to a gappy solution vector: the solution is known at the sensor locations and unknown for other parts of the flow, i.e. $n_j = 1$ if location j is a sensor. Using the gappy POD procedure outlined above, the modal content, b_k^i for POD mode k at timestep i, can be determined by solving equation (9). If desired, the full flowfield can then be constructed using (4). In [13], it was shown that this approach worked well for reconstructing steady flow pressure fields when considering a number of sensors distributed evenly around the surface of an airfoil.

Within the gappy framework, we now consider the problem of where best to place the sensors, i.e. given N sensors and l possible locations, how does one select the locations that will enable the POD modal coefficients to be determined most accurately? Consider equations (6), (7) and (8). If all data are available, i.e. $n_j = 1$ for all j, then M is the identity matrix, and the modal coefficients can be calculated exactly. Consider removing available data, and computing the gappy inner products between basis vectors to form the entries of M. In general, the POD basis vectors are no longer orthogonal when inner products are considered in the gappy sense, and M therefore becomes fully populated. As this orthogonality is lost, so is the ability to exactly identify the modal content. The sensor locations, and correspondingly the non-zero entries in the mask vector, should therefore be chosen to preserve orthogonality between the POD basis vectors, when calculated using the gappy inner product. One should also ensure that the diagonal entries of M are not too small (this would correspond to choosing sensor locations where a POD basis vector value is close to zero). Mathematically, one way to achieve these goals is to minimize the condition number of M. The sensor location problem is therefore stated as:

$$\begin{aligned} & \min \ \kappa(M) \\ \text{s.t.} \quad & n_j \in \{0,1\}, \ j = 1,2,..,l \\ & \sum_{j=1}^{l} n_j = N \end{aligned} \tag{10}$$

where $\kappa(M)$ is the condition number of M.

A solution to this combinatorial optimization problem may be obtained using a greedy algorithm as follows.

(i) Consider placing the first sensor: loop over all possible placement points, evaluate M for each point, and choose the point that minimizes $\kappa(M)$.

(ii) With the first sensor location set, loop over all possible remaining placement points. For each point, update the mask vector, evaluate M, and choose the point that minimizes $\kappa(M)$.

(iii) Repeat step (ii) for all remaining sensor locations.

Note that this approach does not yield an optimal solution, but can be implemented efficiently.

IV. Results

Results will be presented for two cases: unsteady plunging motion of a subsonic airfoil and low Reynolds number flow over a circular cylinder. Both the sensor placement algorithm and the time-dependent gappy reconstruction of POD modal content will be considered.

A. Subsonic airfoil plunging motion

The first case considered is a NACA 0012 airfoil operating in unsteady plunging motion about a steady-state condition that has a freestream Mach number of 0.755 and an angle of attack of 0°. The CFD method is described in [14], and uses a finite-volume formulation on an unstructured grid to solve the two-dimensional linearized Euler equations. The CFD mesh used has 3482 grid points, which corresponds to 13,928 flow perturbation unknowns. POD basis vectors were calculated using a frequency domain method of snapshots approach [15], [14]. The reduced frequency, k_c, is defined as

$$k_c = \frac{\omega c}{\bar{U}} \tag{11}$$

where ω is the frequency of the airfoil plunging motion, c is the airfoil chord, and $\bar{U}$ is the freestream velocity. In order to compute the POD basis vectors, flow snapshots were evaluated at 21 evenly spaced reduced frequencies between zero and one.

1) Gappy sensor placement: The full POD basis vectors contain all flow variables; however, only pressure measurements on the airfoil surface were considered as possible sensor locations. For the grid used, there were 115 possible sensor locations. The first four POD basis vectors capture 99.99% of the energy in the snapshot ensemble, thus $p = 4$ modes were considered when formulating the gappy problem. The greedy algorithm was used to determine the sensor locations while attempting to minimize the condition number of the 4×4 M matrix. The resulting sensor locations are shown in Figure 1 for $N = 5$, 10 and 20. In the figure, the dots correspond to a possible sensor location (this is a CFD grid point on the airfoil surface), while the crosses correspond to the chosen sensor locations. The condition number of the gappy matrix M is 70.96, 12.60, and 9.82 for $N = 5, 10$ and 20, respectively.

It can be seen from the figure that as the number of sensor locations is increased, the positions show some clustering in particularly sensitive regions of the flow, such as near the airfoil leading edge. In the three cases considered, it can be seen that the algorithm chooses to place the sensors in particular regions of the airfoil surface. The choice of these regions can be understood by plotting the spatial variation of the POD basis vectors. Figure 2 shows the values of the first four basis vectors along the airfoil surface. The basis vectors are symmetric with respect to the airfoil top and bottom surfaces. The figure shows increasing spatial frequency

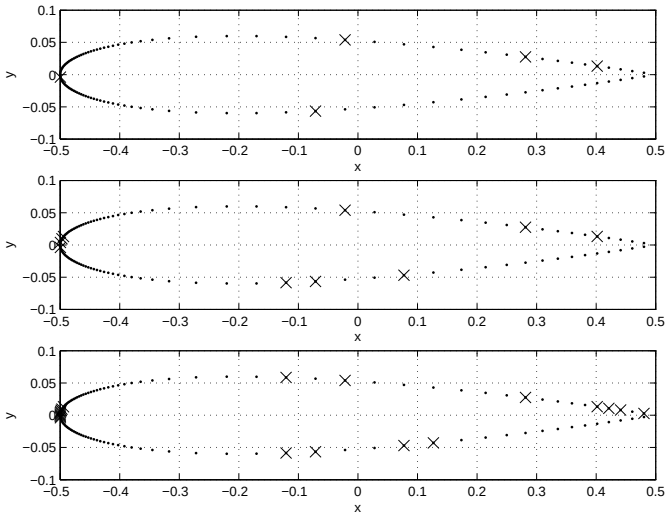

Fig. 1. Resultant positions for sensors using $p = 4$ basis vectors. Dots represent available CFD mesh surface points, crosses represent the chosen sensor locations. From top: 5 sensors, 10 sensor, 20 sensors.

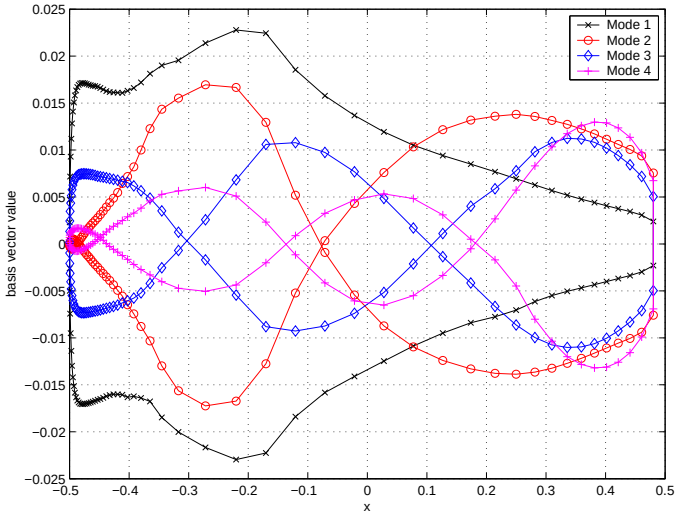

Fig. 2. POD basis vector values along the airfoil surface. Note that the vectors are symmetric with respect to top and bottom surfaces and that spatial frequency increases with increasing mode number.

of subsequent POD vectors, which is commonly observed in flow applications.

By comparing Figures 1 and 2, it can be seen that the locations of the sensors are chosen in regions that correspond roughly to local optima of the POD basis vector variation. This observation corresponds well to the heuristic sensor placement procedure suggested in [11]. In that work, sensors were placed at local spatial optima for each mode. The locations determined by the gappy methodology proposed here do not correspond exactly to optima because the POD basis vectors are not pure harmonics, and the optimization criteria is based on maintaining orthogonality between modes. For example, it can be seen in Figure 2 that $x = -0.25$ is close to a maxima for modes one, two and four. Data at this location may not provide a strong differentiation between the three modes, therefore the algorithm chooses not to place a sensor there. It can also be seen that the algorithm chooses to place a number of sensors at the airfoil leading edge, which is a region where the POD basis vectors vary rapidly.

2) Gappy flow reconstruction: The question of interest is how well, with the selected sensor configuration, the modal content of the time-dependent flow solution can be predicted. For the NACA 0012 airfoil, an unsteady simulation was run that used a forcing input that varied temporally as a Gaussian pulse, i.e. the plunging input, h, is given by

$$h(t) = \bar{h} e^{-\alpha(t-t_0)^2} \tag{12}$$

where $\bar{h}$ is the amplitude of the pulse, t_0 determines the time at which the peak input occurs, and the parameter α determines the sharpness of the pulse and, therefore, the range of frequencies excited in the system. The case considered used $\bar{h} = 1$, $t_0 = 40$, and $\alpha = 0.01$. This value of α yields an input whose significant frequency content lies within the POD snapshot sample range.

At each timestep in the simulation, the actual POD modal content was computed and compared to that predicted using the gappy reconstruction from just the sensor data. The results for the first four POD modes are shown in Figure 3 for $N = 10$. The solid lines show the exact modal amplitude while the symbols are the results calculated using data from the ten sensors. It can be seen that a very good match is obtained for all four modes. The corresponding errors between the actual and predicted values are plotted in Figure 4. For the first two modes, the error can be seen to be very small in magnitude relative to the actual values shown in Figure 3. For modes three and four, the error is larger; however, the prediction is still very good.

B. Cylinder wake flow

The second example analyzes the problem described in [11] of flow prediction in a cylinder wake at a Reynolds number of 100. CFD simulation data from that study was used to form an ensemble of snapshots. This data was obtained using a direct numerical simulation of the Navier Stokes equations on an unstructured grid with the COBALT solver. The cylinder has an incoming flow Mach number of 0.1. The Strouhal number of the wake shedding was computed to be 0.163 [11].

From these snapshots, POD basis vectors were calculated for the vorticity flow component. The relative energy content, defined by (3), corresponding to the first twelve POD basis vectors is given in Table I. It can be seen that the modes occur roughly in pairs (due to the periodic nature of the shedding) and that twelve modes are required to capture 99.9% of the flow vorticity energy.

1) Gappy sensor placement: The sensor placement algorithm described earlier was applied to this problem. As discussed in [11], it is desirable to place the sensors in location that experiences relatively large modal amplitudes.

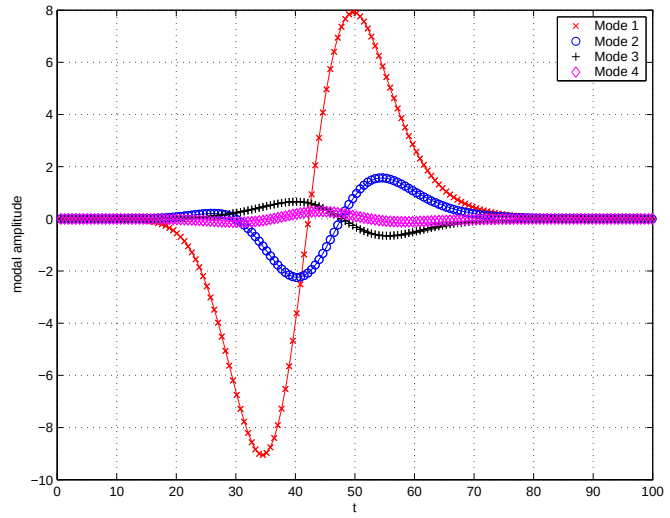

Fig. 3. POD modal amplitude as a function of time for pulse plunge input. Solid lines denote the actual amplitudes; symbols denote the amplitude predicted using gappy POD using $p = 4$, $N = 10$.

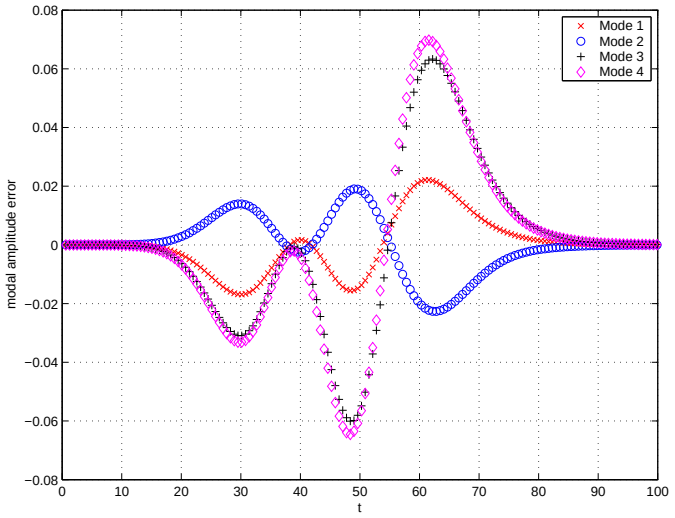

Fig. 4. Error in POD modal amplitude as a function of time for pulse plunge input using $N = 10$ sensors. Plotted is predicted minus actual amplitude for the first four modes.

The optimization formulation based only on minimizing the condition number of M was found to yield reasonable results; however a much better solution could be obtained by modifying the approach as follows. The first sensor was placed by considering the sum of the diagonal elements of M minus the sum of the off-diagonal elements of M. The first sensor location was chosen so as to maximize this quantity, yielding a larger emphasis on amplitude size. Subsequent sensors were then placed according to the condition number minimization described previously. The modified algorithm is therefore as follows.

(i) Consider placing the first sensor: loop over all possible placement points, evaluate M for each point, and choose the point that maximizes the summation of diagonal minus off-diagonal entries of M.

(ii) With the first sensor location set, loop over all possible remaining placement points. For each point, update the mask vector, evaluate M, and choose the point that minimizes $\kappa(M)$.

(iii) Repeat step (ii) for all remaining sensor locations.

Mode number, i	E_i (%)	$\sum_{j=1}^{i} E_j$ (%)
1	46.17	46.17
2	37.81	83.99
3	5.47	89.46
4	4.94	94.40
5	2.31	96.71
6	2.26	98.97
7	.35	99.31
8	.34	99.65
9	.10	99.75
10	.10	99.85
11	.03	99.88
12	.03	99.91

TABLE I

PERCENTAGE ENERGY AND CUMULATIVE ENERGY CONTENT FOR VORTICITY POD MODES FOR CYLINDER FLOW.

2) Gappy flow reconstruction: In [11], the time histories of the POD modal content were estimated using a linear combination of sensor signals. The coefficients of the linear combination were determined using a least squares fit to the known POD modal amplitudes of the original snapshot simulation. The sensing locations were determined using a heuristic criterion of placing sensors in areas of high modal activity, thus using a total of twelve sensors to estimate the first four POD modes: two sensors each for modes one and two, and four sensors each for modes three and four.

The modified gappy sensor placement algorithm described above was implemented for this problem and then the gappy reconstruction approach was used to determine the time-dependent POD modal content. The first case considered was with $p = 4$ POD modes and $N = 12$ sensors. Figure 5 shows the actual and predicted POD modal coefficients using the gappy approach. While the overall trends are captured, it can been seen that there is some error in the prediction, particularly in the peak areas. The RMS error of the prediction for the j^{th} mode is defined as

$$e_j = \frac{\sqrt{\sum_{i=1}^{m} \left(b_j^i - \tilde{b}_j^i \right)^2}}{\sqrt{\sum_{i=1}^{m} \left(\tilde{b}_j^i \right)^2}} \tag{13}$$

where b_j^i and $\tilde{b}_j^i$ are respectively the estimated and actual coefficient for POD mode j at timestep i, $m = 70$ is the number of timesteps considered, and the error is normalized by the RMS value of the actual modal coefficient. The RMS errors for reconstruction of the first four modes are given in Table II. The table also shows that the condition number of

N	p	e_1	e_2	e_3	e_4	$\kappa(M)$
12 (greedy)	4	11.84%	23.48%	32.72%	12.69%	1.82
12 (modal max/min)	4	6.12%	5.74%	8.03%	9.32%	1.65
12 (modal max/min)	LSE	2.1%	0.6%	7.1%	2.9%	-
12 (modal max/min)	12	3.63%	1.98%	48.56%	7.76%	1.64×10^4
20 (greedy)	12	1.75%	1.46%	5.94%	5.23%	46.42

TABLE II

RMS ERRORS OF GAPPY RECONSTRUCTION FOR VARIOUS SENSOR CONFIGURATIONS AND NUMBERS OF MODES. THE LSE DATA IS TAKEN FROM [11].

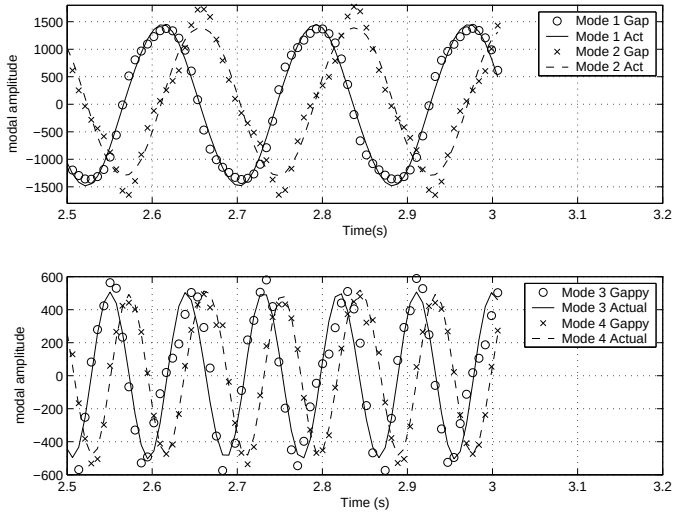

Fig. 5. Actual (lines) and estimated (points) POD modal coefficients using twelve sensors and four POD modes.

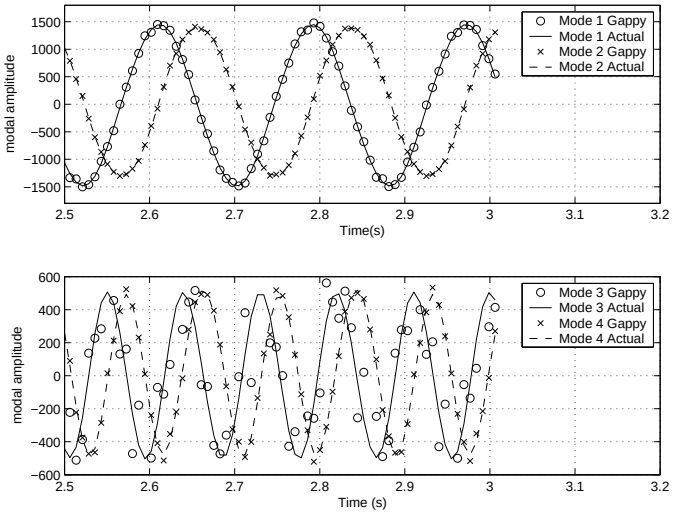

Fig. 6. Actual (lines) and estimated (points) POD modal coefficients using twelve sensors and twelve POD modes.

the matrix M in this case was 1.82.

The sensor configuration suggested by Cohen et al. [11] was also considered, obtained by selecting sensing locations at spatial minima and maxima of the POD modes. The resulting RMS errors from the gappy reconstruction are given in Table II. It can be seen that this configuration resulted in an improved condition number of 1.65, demonstrating the inability of the greedy algorithm to find the true optimum. Since the cylinder flow case exhibits strong periodicity, the POD modes resemble Fourier modes, and the heuristic approach is expected to yield accurate results. The RMS errors for this case are also given in Table II and can be seen to be much lower than for the previous configuration. This result suggests that the condition number criterion is appropriate for sensor location choice, but that a better optimization algorithm is required. For comparison, the RMS errors using the least squares approach with this sensor configuration are also given in Table II.

There are two possible sources of error for the results presented in Figure 5. First, there may be insufficient sensors to accurately reconstruct the modal information using gappy POD. Secondly, an insufficient number of POD modes may be used in the reconstruction. Recall that the gappy procedure chooses the coefficients so as to minimize the gappy norm between the available and the reconstructed data. If the neglected higher POD modes have significant contribution to the sensor signals, then the reconstruction will be inaccurate. As Table I shows, the first four modes capture only 94.4% of the flow energy. In order to more accurately represent the flow, the number of POD modes considered was increased to $p = 12$. Using the twelve sensors located at modal optima, the reconstructed modal coefficients were calculated and are shown in Figure 6. As can be seen from the figure, and from the corresponding RMS errors in Table II, the reconstructions for modes 1,2 and 4 are now excellent; however, the estimated response for mode 3 shows significant oscillations and a large reconstruction error. This error is most likely due to poor numerical conditioning of the system, which has a condition number of 16,400.

In order to reduce the condition number of the system and thus reduce the reconstruction error, more sensors must be added. The greedy algorithm was used to place $N = 20$ sensors while attempting to minimize the condition number of a gappy matrix with $p = 12$ POD modes, yielding an improved condition number of 46.42. The modal amplitudes are plotted in Figure 7 and, along with the RMS errors in Table II, show an excellent prediction.

3) Discussion and Comparison of Approaches: As can be seen in Table II, the RMS errors for the gappy estimation of the first four modes obtained with $N = 20$ sensors and $p = 12$ modes are of similar magnitude to those obtained using a least

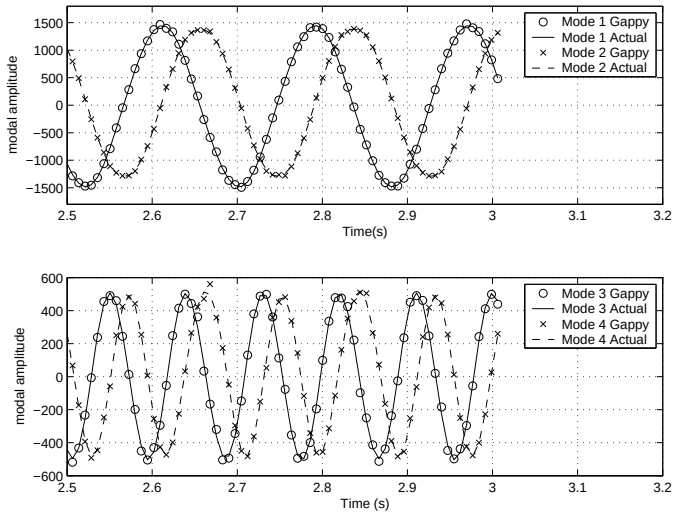

Fig. 7. Actual (lines) and estimated (points) POD modal coefficients using twenty sensors and twelve POD modes.

squares approach with $N = 12$. In the latter method, the RMS error represents a measure of the quality of the least squares fit for the known evolution of the modal coefficients. It is not obvious how accurate the prediction would be for flows where the relative modal content is different to that used for calibration. The gappy estimation procedure does not require knowledge of the actual modal coefficients, and should yield accurate results for a range of different flows, provided the flow can be resolved with a sufficient number of available basis vectors.

For the cylinder flow, which exhibits strong periodicity, the heuristic approach to sensor placement results in a well-conditioned gappy problem and accurate reconstructions for the first four modes. In this case, the greedy algorithm did not find a twelve-sensor configuration that yielded a better condition number. However, the greedy algorithm provides a satisfactory alternative for cases where the heuristic approach may not be appropriate. For example, if the POD basis vectors do not exhibit a strong sinusoidal structure, location choices may not be obvious. The locations of modal optima may also not be physically available for sensing (for example, in the cylinder case it is assumed that velocity measurements are available anywhere in the flow). Finally, the heuristic approach often may not extend to estimation of higher modes. If the higher modes exhibit sinusoidal behavior, then an increasing number of sensors per mode will be required to capture all optima. If the higher modes lose their structure (as is often the case for flow applications) then it will not be clear where to choose the sensor locations. The gappy formulation easily handles an increasing number of modes, although as the results demonstrated, a sufficient number of measurement points must be provided in order to reduce the condition number of the gappy matrix M.

V. Conclusions

The gappy POD methodology provides a natural framework to directly estimate POD modal content from limited flow measurements and has been shown to work effectively for two examples. With limited sensor measurements, modal content for the dominant POD modes can be estimated accurately for an unsteady flow, provided a sufficient number of modes are used to resolve the flow. When used in conjunction with a heuristic approach of placing sensors at POD spatial optima, the gappy reconstruction yields excellent results. Alternatively, the sensor placement problem can be formulated mathematically in the gappy framework using a condition number criterion. While the resulting optimization problem cannot be solved exactly, an approximate solution method leads to results that support the intuitive approach of placing sensors at modal optima. The more formal approach extends to cases where the heuristic criterion might be difficult to apply. Further investigation into an improved optimization solution method is required.

In another area of future work, an iterative gappy procedure can be used to derive a set of POD basis vectors using incomplete snapshot data. This would, for example, enable the CFD and PIV measurements to be combined when determining the POD basis vectors. It would also enable effective handling of imperfect PIV measurement data, a situation often encountered in practice.

Acknowledgment

The author would like to thank Dr. Cohen, Dr. Siegel, and Dr. McLaughlin of the U.S. Air Force Academy for providing the CFD data for the cylinder flow example, and for helpful discussions regarding the research.

References

[1] P. Holmes, J.L. Lumley, and G. Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press, Cambridge, UK, 1996.

[2] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. Jounal of the Optical Society of America A, 4(3):519-524, 1987.

[3] K. Fukanaga. *Introduction to Statistical Pattern Recognition.* Academic Press, New York, 1972.

[4] L. Sirovich. Turbulence and the Dynamics of Coherent Structures. Part 1 : Coherent Structures. *Quarterly of Applied Mathematics*, 45(3):561–571, October 1987.

[5] E.H. Dowell, K.C. Hall, J.P. Thomas, R. Florea, B.I. Epureanu, and J. Heeg. Reduced Order Models in Unsteady Aerodynamics. AIAA Paper 99-1261, 1999.

[6] M.C. Romanowski. Reduced Order Unsteady Aerodynamic and Aeroelastic Models using Karhunen-Loéve Eigenmodes. AIAA Paper 96-194, 1996.

[7] J.E. Ffowcs Williams and X. Huang. Active Stabilization of Compressor Surge. *Journal of Fluid Mechanics*, 204:245–262, 1989.

[8] J.D. Paduano, A.H. Epstein, L. Valavani, J.P. Longley, E.M. Greitzer, and G.R. Guenette. Active Control of Rotating Stall in a Low-Speed Axial Compressor. *Journal of Turbomachinery*, 115:48–56, 1993.

[9] M. Samimy *et al.* Development of closed-loop control for cavity flows. AIAA Paper 2003-4258, 2003.

[10] G. Lassaux. High-Fidelity Reduced-Order Aerodynamic Models: Application to Active Control of Engine Inlets. Master's thesis, Dept. of Aeronautics and Astronautics, MIT, June 2002.

[11] K. Cohen, S. Siegel, and T. McLaughlin. Sensor placement based on proper orthogonal decomposition modeling of a cylinder wake. AIAA Paper 2003-4259, 2003.
[12] R. Everson and L. Sirovich. The Karhunen-Loeve Procedure for Gappy Data. J.Opt.Soc.Am., 12: 1657-1664, 1995.
[13] B.-T. Tan, M. Damodaran, and K. Willcox. Proper orthogonal decomposition extensions for parametric applications in transonic aerodynamics. AIAA Paper 2003-4213, 2003.
[14] K.E. Willcox. *Reduced-Order Aerodynamic Models for Aeroelastic Control of Turbomachines*. PhD thesis, Dept. of Aeronautics and Astronautics, MIT, February 2000.
[15] T. Kim. Frequency-Domain Karhunen-Loeve Method and Its Application to Linear Dynamic Systems. *AIAA Journal*, 36(11):2117–2123, 1998.