Data-driven Vector-measurement-sensor Selection based on Greedy Algorithm for Particle-image-velocimetry Measurement

Yuji Saito∗, Taku Nonomura†, Koki Nankai‡, Keisuke Asai§

Tohoku University, Sendai, Miyagi 980-8579, Japan

Yasuo Sasaki¶ and Daisuke Tsubakino∥

Nagoya University, Nagoya, Aichi, 464-8603, Japan

I. Introduction

Recently, reduced-order modelling for fluid analysis and flow control gathers a lot of attentions. With regard to reduced-order modelling, proper orthogonal decomposition (POD) is one of the effective methods to decompose the high-dimension fluid data into several significant modes of flow fields. Here, POD is a data-driven method which gives us the most significant and relevant structure in the data, and it exactly corresponds to principal component analysis and KL decomposition, where the decomposed modes are orthogonal each other. The POD analysis for discrete data matrix can be carried out by applying the singular value decomposition, as is often the case in the engineering fields. Although several advanced data-driven methods, dynamic mode decomposition, empirical mode decomposition, and others which include efforts by the present authors, this research is only based on POD which is the most basic data-driven method for reduced-order modelling.

If the data, such as flow fields, can be effectively expressed by limited numbers of POD modes, the limited sensor placed at appropriate positions gives us the approximated full state information. This effective observation might be one of the keys for flow control and flow prediction.
idea has been adopted by Manohar et al.\textsuperscript{6} and the sparse-sensor-placement algorithm has been developed and discussed. The idea here is expressed by the following equation:

\begin{align}
\mathbf{y} &= \mathbf{HUx} \\
&= \mathbf{Cx}
\end{align}

Here, $\mathbf{y} \in \mathbb{R}^p$, $\mathbf{x} \in \mathbb{R}^r$, $\mathbf{H} \in \mathbb{R}^{p \times n}$ and $\mathbf{U} \in \mathbb{R}^{n \times r}$ are the observation vector, the POD mode amplitude vector, the sparse sensor location matrix, and spatial POD modes, respectively. In addition, $p$, $n$, and $r$ are numbers of sensor location, degree of freedom of the spatial POD modes, and the rank for truncated POD, respectively. The problem above is considered to be one of the sensor selection problems when the POD mode $\mathbf{U}$ is assumed to be a sensor-candidate matrix. The graphical image of the equation above is shown in Fig. 1. Thus far, this sensor selection problem has been solved by convex approximation and greedy algorithm, where greedy algorithm was shown to be much faster than the convex approximation algorithms. The greedy algorithm is based on QR-discrete-empirical-interpolation method\textsuperscript{7, 8} when the number of the sensors is the same as that of POD mode and its extension for the least square problem when the number of the sensor is greater than that of POD modes. Both convex approximation and greedy algorithm work pretty well for the sensor selection problems. Now, the present authors attempt to apply those methods to vector-measurement

![Graphical image for sensor matrix H on Eq. (1)](image)

sensor, such as two components of velocity of particle image velocimetry, or simultaneous velocity, pressure and temperature measurements used in weather forecasting. The extension of the
vector-measurement-sensor selection of convex approximation has already been addressed in the original paper,\textsuperscript{9} while one of the greedy algorithm has not been conducted. The sensor selection of very high dimension with such a constraint should be resolved in reasonable time scale when the real-time measurement and flow-control or flow prediction would be conducted. Therefore, the extension of greedy sparse sensor selection method with the vector measurement sensors is straightforwardly proposed and it is applied to test and PIV data to reconstruct the full state based on the information given by the sparse vector-measurement sensors.

II. Material and Methods

II.A. PIV Measurement

The PIV measurement for acquiring time-resolved data of flow fields around an airfoil was conducted previously.\textsuperscript{10} Here, the experimental data are briefly explained. The wind tunnel testing was conducted in Tohoku-university Basic Aerodynamic Research Wind Tunnel (T-BART) with a closed test section of 300 mm × 300 mm cross-section. The airfoil of the test model has an NACA0015 profile the chord length and span width of which are 100 mm and 300 mm, respectively. The freestream velocity $U_\infty$ and attack angle of the airfoil $\alpha$ were set to be 10 m/s and 16 degree, respectively. The time-resolved PIV measurement was conducted with the double-pulse laser. In PIV measurement, the time between pulses, the sampling rate, the particle image resolution and the total number of image pairs were $100 \mu s$, 5000 Hz, $1024 \times 1024$ pixels, and $N = 1000$, respectively. The tracer particles were 50 % aqueous solution of glycerin with estimated diameter of a few micrometers. The particle images were acquired by using the double pulse laser (LDY-300PIV, Litron) and a high-speed camera (SA-X2, Photron) which were synchronised each other.

II.B. Previous Greedy Algorithm for Scalar Measurement Sensors

In the greedy algorithm based on QR decomposition for scalar measurement problem, $i$-th sensor is chosen where

$$i = \arg \max \|v_i\|^2_2. \quad (3)$$

Here, $v_i = [V_{i,1} V_{i,2} \ldots V_{i,r}]$, and $V_{ij} = V = U$ and $V_{ij} = V = UU^T$ for $p = r$ and $p > r$ sensor conditions, respectively. Given $i$ index, the $V$ matrix is pivoted and QR decomposition is conducted. After that, the next sensor is chosen for remaining matrix. The algorithms for $p = r$ and $p > r$ sensors are QR-discrete-empirical-interpolation method (QDEIM)\textsuperscript{8} and optimized sparse sensor placement as an extension of QDEIM, respectively.

The both optimizations are considered to conduct the maximization of determinant of the $C$
or $C^TC$ matrix to stably solve $x$ vector. Here, $C^TC$ is used when the dimension of observation $y$ is larger than that of the state $x$ in the least square method or in the pseudo inverse matrix multiplication. For this purpose, the select of the sensor position is based on maximizing the norm of the corresponding row vector of the sensor-candidate matrix. Although the round-off error increases, this procedure can be written by Gram-Schmidt orthogonalization with choosing the rows of the largest norm. This replacement of QR decomposition by Gram-Schmidt procedure clearly shows what we do in this algorithm as shown in Algorithm 1.

**Algorithm 1** Greedy algorithm for sparse scalar measurement sensor placement

```markdown
procedure Greedy Sparse Scalar Measurement Sensor Placement Algorithm $(a, b)$
    Set sensor-candidate matrix $U$.
    if $p = r$ then
        $V = U$
    else
        $V = UU^T$
    end if
    $k ← 1$
    for $k = 1, \ldots, p$ do
        $v_i = [V_{i,1}V_{i,2} \ldots V_{i,r}]$
        $i ← \arg \max_i \|v_i\|_2^2$
        $V ← V - Vv_i^T/\|v_i\|_2^2$
        $H_{k,i} = 1$
        $k ← k + 1$
    end for
    return $H$
end procedure
```

II.C. Proposed Greedy Algorithm for Vector Measurement Sensors

In the vector measurement, we consider the following equation:

$$y = C \begin{bmatrix} H & 0 & 0 & 0 \\ 0 & H & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & H \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_s \end{bmatrix} x$$

$$= H_s U x$$

$$= C x$$

Here, $U_k \in \mathbb{R}^{n_s \times r}$ is a $k$th vector component of a sensor candidate matrix, $H \in \mathbb{R}^{p \times \frac{n_s}{r}}$ is the sensor location matrix for each vector components. Here, $s$ is the number of components of the measure-
ment vector. Again, $p$, $n$, and $r$ are the numbers of sensor location, degree of freedom of the spatial POD modes including different vector component, and the rank for truncated POD, respectively. This arrangement of data is intentionally chosen with considering the situation when we apply the data matrix of $X = [X_u^T \ X_v^T]^T$ and use the spatial POD modes of $X$ as a sensor-candidate matrix, where $X_u^T$ and $X_v^T$ are data matrices of $x$-velocity component and $y$-velocity component in PIV data, respectively. This arrangement does not matter for the Gram-Schmidt procedure, but we recommend to reorder the data as the data of the same vector-measurement sensor are gathered in successive rows when this algorithm is further straightforwardly extended by block-pivoting and block-QR algorithm for eliminating round-off error. Unfortunately, the latter extension is not addressed in this short note, for brevity.

In the sense of Eq. [5] the vector-measurement sensor placement problem can be seen as the restriction on the choice of sensor of $H_s$ matrix. Similar to the scalar version, the next sensor can be chosen to maximize the determinant of $C$ or $C^T C$ matrix. Because the multiple ($s$) rows of $U$ matrix are chosen simultaneously by selecting one point in the vector version, the hypervolume of selected row vectors are maximized, instead of the norm of the row vector in the scalar version. This hypervolume (square root of $J_i$ in the algorithm) can be simply computed by the Gram-Schmidt-like procedure by multiplying the norm of one row and removing its component from other selected rows in order. After choosing the sensors, we can subtract the components of selected row vectors from the sensor-candidate matrix, and then proceed to next sensor placement selection. The algorithm is summarised in Algorithm 2.

### III. Results and Discussions

The reduced-order PIV data are reconstructed by sparse sensors that are chosen by several method. Here, the PIV data for flows around airfoils are adopted, and the truncated POD with $r = 10$ modes are preconditioned. Here, at least five sensors which has two velocity components for each are required for the reconstruction of the PIV data of $r = 10$. The extension of the convex approximation method for the vector sensor placement are addressed in the original paper, and it is adopted in present test case. This method is called a convex (vector) method, for brevity. The greedy method for the scalar sensor placement is applied only for the $u$ field, but both $u$ and $v$ information are employed when the data is reconstructed. Because the $v$ information is not used for the sensor placement at all, it is not expected to work well. This method is called a greedy (scalar) method for brevity. The proposed greedy method for vector sensor placement is called a greedy (vector) method for brevity.

Figure 2 shows the selected sensor position (black open circle) over the $u$-velocity flow field at the same PIV-data time and POD mode amplitude histories of $r = 1, 2$ and 3 obtained by the noisy data (blue plus line) and the sparse-sensor reconstruction (black close circle line) at $p = 5$.
Algorithm 2 Greedy algorithm for sparse vector-measurement sensor placement

procedure Greedy Sparse Vector Measurement Sensor Placement Algorithm \((a, b)\)

Set sensor-candidate matrix \(U_j\) for \(j\)th vector component measurement.

\[
U = \begin{bmatrix}
U_1^T & U_2^T & \ldots & U_r^T
\end{bmatrix}^T
\]

if \(sp = r\) then
\[V = U\]
else
\[V = UU^T\]
end if

for \(k = 1, \ldots, p\) do
for \(i = 1, \ldots, n\) do
\[v_i = \begin{bmatrix}
V_{i,1} & V_{i,2} & \ldots & V_{i,r}
\end{bmatrix}\]
end for
for \(i = 1, \ldots, \frac{n}{s}\) do
\[J_i = 1\]
\[\tilde{V} = V\]
for \(j = 1, \ldots, s\) do
\[J_i = J_i \| (\bar{v}_{i+\frac{n}{s}(j-1)} \|_2^{2}\]
\[\tilde{V} \leftarrow \tilde{V} - \tilde{V}(\bar{v}_{i+\frac{n}{s}(j-1)} \bar{v}_{i+\frac{n}{s}(j-1)} / \| \bar{v}_{i+\frac{n}{s}(j-1)} \|_2^{2})\]
end for
\[i = \arg \max J_i\]
\[H_{i,j} = 1\]
for \(j = 1, \ldots, s\) do
\[v_{i+\frac{n}{s}(j-1)} = \begin{bmatrix}
V_{i+\frac{n}{s}(j-1),1} & V_{i+\frac{n}{s}(j-1),2} & \ldots & V_{i+\frac{n}{s}(j-1),r}
\end{bmatrix}\]
\[V \leftarrow V - V(v_{i+\frac{n}{s}(j-1)} v_{i+\frac{n}{s}(j-1)} / \| v_{i+\frac{n}{s}(j-1)} \|_2^{2})\]
end for
end for
return \(H\)
end procedure
and \( r = 10 \). For all the cases, the reconstruction is based on the linear least squares method and it is written with the \( C \) matrix as follows:

\[
x_{i}^{\text{reconst}} = (C^T C)^{-1} C^T y_i
\]  

(7)

Figures 2(a), (b), (c) and (d) are original data, the reconstructed data based on the sensor chosen by the convex approximation method for vector sensor placement, those based on the sensor chosen by the present greedy method for vector sensor placement, and those based on the sensor chosen by the previous greedy method for scalar sensor placement respectively. The smaller difference between the POD mode amplitude history of the noisy data and the reconstructed based on the sparse sensors corresponds to the more precise reconstruction of the flow field by the sparse sensor chosen by the method above. Although there is no qualitatively clear difference in the selected positions of each sparse sensor by convex (vector), greedy (vector) and greedy (scalar) methods, each POD mode amplitude is quite different. Here, the estimation error \( \epsilon \) is introduced for quantitative evaluation of flow field reconstruction.

\[
\epsilon = \sqrt{\sum_{i=1}^{r} \sum_{j=1}^{N} \frac{e_{i,j}^2}{N}}
\]  

(8)

Here, \( e_{i,j} = x_{i,j} - x_{i,j}^{\text{reconst}} \). Figure 3 shows the relationship between the estimation error \( \epsilon \) and the number of sensors \( p \), obtained by convex (vector) (gray close circle), greedy (scalar) (black open circle), and the present greedy (vector) (blue close circle) at \( r = 10 \). The estimation error decreases as the number of sensors increases. In the case of the number of sensors \( p = 5 \), the estimation error of convex (vector), greedy (vector) and greedy (scalar) algorithms are 144, 73, and 301, respectively. Although the greedy (vector) algorithm is more effective than the other algorithms in a small number of sensors \( p < 8 \), the estimation error of greedy (scalar) is smaller than that of greedy (vector) in \( p > 8 \) as shown in Fig. 3. This might be because the greedy (vector) algorithm cannot find good sensor positions after \( p > 8 \) due to elimination of the \( v \) components of the chosen sensors from the \( u \)-components of the sensor candidate matrix, whereas \( u \) components of the present PIV data are better than \( v \) components from the view points of the signal-to-noise ratio. Meanwhile, the greedy (scalar) algorithm can choose better sensor position after \( p > 8 \) because it only choose the sensors based only on \( u \) components which has better signal-to-noise ratio than \( v \) components. Also, the estimation error of the convex (vector) algorithm is smaller than that of the greedy (vector) algorithm in \( p > 8 \). However, because the calculation cost of convex is much bigger than greedy (vector) algorithm, the greedy (vector) algorithm is more effective for sparse sensor placement than the convex (vector) algorithm especially in the case of a small number of sensors.
Figure 2. Sensor positions and POD mode amplitude histories $u$-velocity flow field at the same PIV data time at $p = 5$ and $r = 10$: (a) original data, (b) reconstructed data by convex (vector) method, (c) those by greedy (vector) method, (d) those by greedy (scalar) method.

Figure 3. The relationship between number of sensors and estimation error at $r = 10$. 

- Convex (vector)
- Greedy (scalar)
- Greedy (vector)
IV. Conclusions

The greedy method extended for the vector-measurement sensor problem, such as the sensor placement problem in PIV data, is introduced and investigated in this paper. The sensor selection problem is solved by a convex approximation method for vector sensor and greedy algorithms for scalar and vector sensors, where greedy algorithms are shown to be much faster than the convex approximation algorithms. The calculation results show that estimation error of the vectored greedy is smaller than other methods in a small number of sensors ($p < 8$). However, the difference between estimation errors obtained by each method is small as a number of sensors increases. Therefore, the vectored greedy is illustrated to be more effective sparse sensing than other methods in a small number of sensor.

Acknowledgements

The second author T.N. is grateful for support of the grant JPMJPR1678 of JST Presto.

References

1. Berkooz, G., Holmes, P., and Lumley, L. J., “The proper orthogonal decomposition in the analysis of turbulent flows,” *Annual Review of Fluid Mechanics*, Vol. 25, No. 1971, 1993, pp. 539–575.
2. Schmid, P. J., “Dynamic mode decomposition of numerical and experimental data,” *Journal of Fluid Mechanics*, Vol. 656, No. July 2010, 2010, pp. 5–28.
3. Kutz, J. N., Brunton, S. L., Brunton, B. W., and Proctor, J. L., *Dynamic mode decomposition: data-driven modeling of complex systems*, Vol. 149, SIAM, 2016.
4. Nonomura, T., Shibata, H., and Takaki, R., “Dynamic mode decomposition using a Kalman filter for parameter estimation,” *AIP Advances*, Vol. 8, 2018, pp. 105106.
5. Nonomura, T., Shibata, H., and Takaki, R., “Extended-Kalman-filter-based dynamic mode decomposition for simultaneous system identification and denoising,” *PloS one*, Vol. 14, No. 2, 2019, pp. e0209836.
6. Manohar, K., Brunton, B. W., Kutz, J. N., and Brunton, S. L., “Data-Driven Sparse Sensor Placement for Reconstruction: Demonstrating the Benefits of Exploiting Known Patterns,” *IEEE Control Systems Magazine*, Vol. 38, No. 3, June 2018, pp. 63–86.
7. Chaturantabut, S. and Sorensen, D. C., “Nonlinear model reduction via discrete empirical interpolation,” *SIAM Journal on Scientific Computing*, Vol. 32, No. 5, 2010, pp. 2737–2764.
8. Drmac, Z. and Gugercin, S., “A new selection operator for the discrete empirical interpolation method—improved a priori error bound and extensions,” *SIAM Journal on Scientific Computing*, Vol. 38, No. 2, 2016, pp. A631–A648.
9. Joshi, S. and Boyd, S., “Sensor selection via convex optimization,” *IEEE Transactions on Signal Processing*, Vol. 57, No. 2, 2009, pp. 451–462.
10. Nankai, K., Ozawa, Y., Nonomura, T., and Asai, K., “Linear Reduced-order Model based on PIV Data of Flow Field around Airfoil,” *Transactions of Japan Society of Space and Aeronautical Science (to appear)*, 2019.