Greedy Sensor Placement for Weighted Linear-Least Squares Estimation under Correlated Noise

Keigo Yamada, Yuji Saito, Taku Nonomura and Keisuke Asai

Abstract—Optimization for sensor placement has been intensely studied to monitor complex, large scale systems, whereas one needs to overcome its intractable nature of the objective function for the optimization. In this study, a fast algorithm for greedy sensor selection is presented for a linear reduced-ordered reconstruction under the assumption of correlated noise on the sensor signals. The presented algorithm accomplishes the maximization of the determinant of the Fisher information matrix in the linear inverse problem, while this study firstly shows that the objective function with correlated noise is neither submodular nor supermodular. Efficient one-rank computations in the greedy selection procedure are introduced in both of the underdetermined and oversampled problem. Several numerical experiments show the effectiveness of the selection algorithm for its accuracy in the estimation of the states of large dimensional measurement data.

Index Terms—Greedy algorithm, Optimization, Sensor selection

I. INTRODUCTION

OPTIMIZATION for sensor positions is essential when monitoring real-world phenomena. Reduction in the number of measurements is also required for monitoring under resource constraints on sensors and communication energy, and required for processing the signals in real-time. The background of the sensor placement problem varies and one can find them in several monographs; computing, sensor network, health structural monitoring, and source localizing. Especially, the aim of this paper is focused on monitoring phenomena with their data-driven modeling, treated in Refs. [1], [2], [3], [4].

The main challenge of the optimization for such selection problems is its intractability, where problems are often classified to be nondeterministic polynomial-time hard. Therefore, heuristics to find suboptimal solutions were intensely discussed. For example, the selection problem is relaxed and solved via linear convex relaxation method [5], [6], using ADMM [7], [8] or semi-definite programming [9]. Greedy algorithms are also heuristics that seek small subsets maximizing objective functions. A property of the submodularity in those problems encourages the methodologies with greedy algorithms from some mathematical aspects [10], [11], [12].

There have been also several works that include regularization into these optimization along with weighting terms for sensor signals, arising from Bayesian estimation [13], [14] and Kalman filtering [15], [16]. The objective functions were with prior distribution about estimated variables as a kind of hyperparameters. However, there are some difficulties since one must determine the hyperparameters beforehand, which requires another optimization or data-driven assumption. Therefore, the formulation without hyperparameters for the sensor optimization is required for practical use.

In the present paper, the greedy algorithm without hyperparameters for the sensor optimization with a correlated measurement noise is properly formulated and its performance is demonstrated. The proposed algorithm in the manuscript is a sequel of greedy algorithms in Ref. [12], [14] which involves a noise covariance term without any assumption of the prior distribution of estimated variables. The present paper 1) proposes an objective function for the sensor selection adopted to the generalized least-squares estimation, 2) shows that the objective function is neither submodular nor supermodular, and 3) formulates a greedy algorithm that selects sensors for both of underdetermined and overdetermined cases, and reduces its computational complexity by one-rank transformation to the objective function.

II. FORMULATION AND ALGORITHM

This section provides formulation for sensor placement problem tailored for the weighted least-squares estimation. Algorithms for the greedy selection follow them.

A. Sparse Sensing and Sensor Placement Problem

A linear measurement equation with noise is as follows:

\[ y = Cz + v, \]  

where \( y, v \in \mathbb{R}^p, C \in \mathbb{R}^{p \times r} \) and \( z \in \mathbb{R}^r \). Inverse operation of Eq. (1) gives estimator of \( z \),

\[ \hat{z} = \begin{cases} C^T (CC^T)^{-1} y & (p \leq r) \\ (C^T R^{-1} C)^{-1} C^T R^{-1} y & (p > r) \end{cases} \]  

where a positive-definite weighting matrix \( R = E[ vv^T ] \in \mathbb{R}^{p \times p} \) represents the noise covariance over selected sensors. The present study focuses on the above formulation excluding any prior distribution of the latent variables. Note that Eq. (2) gives the minimal norm solution while Eq. (2a) is the generalized least squares estimation. The matrix \( C^T R^{-1} C \) corresponds to the Fisher information matrix.

Sensing phenomena via a few sensors leverages the sparsity in the collected high-dimensional data. For this purpose, decomposition on the data matrix is conducted and the data matrix is represented by a small number of bases [1], [18]. This
Because this metric takes positive values if \( r \) minimizing the determinant of the Fisher information matrix. In these literature, one can find two ways to construct \( R \) and the measurement matrix \( C \). Measurements from sensors \( Y \) into spatial modes \( X \), then compressed to a reduced-order representation using principle \( r \) modes:

\[
X = U_X \Sigma_X V_X^T = U_X \Sigma_X V_X^T + U_N \Sigma_N V_N^T \approx U_X \Sigma_X V_X^T, \tag{3}
\]

where \( X, U_X \in \mathbb{R}^{n \times m} \) and \( \Sigma_X, V_X \in \mathbb{R}^{m \times m} \), and \( U \in \mathbb{R}^{n \times r} \), \( \Sigma \in \mathbb{R}^{r \times r} \) and \( V \in \mathbb{R}^{m \times r} \), respectively.

Then, a sensor selection matrix \( H \in \mathbb{R}^{p \times n} \) is introduced for \( p \) selections of sensors. The position of unity in the \( i \)-th row of \( H \) is associated with the \( i \)-th sensor position, while the rest of the row are zeros. Measurements from sensors \( Y \in \mathbb{R}^{p \times m} \) and the measurement matrix \( C \in \mathbb{R}^{p \times r} \) are constructed:

\[
Y = HX, \quad C = HU.
\]

Associating the variance of estimation error with an error ellipsoid, its volume is minimized by selecting sensors \[5\].

\[
\text{vol} \left( E \left( \begin{bmatrix} z - \tilde{z} \\ \tilde{z} \end{bmatrix} \right)^T \right) = e^r \det (C^T R^{-1} C)^{-1}. \tag{4}
\]

Because this metric takes positive values if \( p > r \), the objective function takes its inverse:

\[
\text{argmax}_H \det (C^T R^{-1} C). \tag{5}
\]

It should be noted that this objective is equivalent to maximizing the determinant of the Fisher information matrix. According to the derivation in Ref. \[17\], the objective function for \( p \leq r \) is derived in Eq. \[6\]:

\[
\det (C^T R^{-1} C + \delta I) = \delta^r \det (\delta^{-1} C^T R^{-1} C + I) = \delta^{r-p} \det (R^{-1} C C^T + \delta I)
\]

\[
\therefore \quad \text{argmax}_H \det (C^T R^{-1} C) \quad (\delta \rightarrow 0) \quad \text{argmax}_H \det (R^{-1} C C^T). \tag{6}
\]

Finally, the objective function is formulated as in Eq. \[7\]:

\[
\text{argmax}_H f = \begin{cases} 
\text{argmax}_H \det (R^{-1}) \det (C C^T) & (p \leq r) \\
\text{argmax}_H \det (C^T R^{-1} C) & (p > r).
\end{cases} \tag{7a, 7b}
\]

It recovers the sensor selection for the uncorrelated noise \[2\], \[17\], i.e. assuming \( R \) is the identity matrix.

Some studies have suggested concrete methods for building the noise covariance matrix in the context of sensor placement. In these literature, one can find two ways to construct \( R \): one is the use of kernel functions that is familiar with signal processing fields \[13\] and the other is a data-driven modeling of the noise \[14\]. Note that a model of noise in the latter case is a byproduct of the data-driven reduced-order modeling of phenomena as \( U_N \Sigma_N V_N^T \) of Eq. \[3\]. The product of matrices represents the measurement noise at every sensor location, therefore the noise covariance matrix is \( R = (HU_N) \Sigma_N (HU_N)^T \) \[14\], Section 2]. The data-driven method is used in the results of section \[11\].

### B. Objective function viewed from estimation error covariance

One can also derive Eq. \[7\] from minimizing the determinant of the error covariance matrix, by following the formula in Ref. \[19\]. “Whitening” is required before conducting the least-squares estimation on the measurement \( y \):

\[
R^{-1/2} y = R^{-1/2} C z + R^{-1/2} v \Rightarrow y_n = C_n z + v_n, \tag{8}
\]

where terms with subscription \( n \) in Eq. \[8\] have been corrected by the noise weighting term \( R^{-1/2} \), and \( v_n \) becomes white noise therefore \( E \left[ v_n v_n^T \right] = \sigma_n I \) is a constant matrix. Then, the linear least-squares process to the whitened measurement vector is carried out and \( \tilde{z} \) is obtained:

\[
\tilde{z} = C_n^+ y_n = \begin{cases} 
C_n^T (C_n C_n^T)^{-1} y_n & (p \leq r) \\
(C_n^T C_n)^{-1} C_n^T y_n & (p > r).
\end{cases} \tag{9a, 9b}
\]

Interestingly, Eq. \[9a\] recovers to the pseudo-inverse operation which is used for the nonweighted linear least-squares estimation (Eq. \[25\]). However, the sensor selection criteria should be different from that for the nonweighted linear least-squares estimation, as discussed below and in the previous section.

The coordinate transformation \( z \rightarrow \xi = V_1^T z \) allows us to evaluate the set of sensors along the observable space of \( C_n \):

\[
E \left[ \xi - \hat{\xi} \right] (\xi - \hat{\xi})^T = \begin{cases} 
\sigma_n^2 U_C^T (C_n C_n^T)^{-1} \hat{U}_C & (p \leq r) \\
\sigma_n^2 V_C^T (C_n^T C_n)^{-1} \hat{V}_C & (p > r).
\end{cases} \tag{10}
\]

where

\[
C_n = \begin{bmatrix} \hat{U}_C & \hat{V}_C \end{bmatrix} \begin{bmatrix} \hat{\Sigma}_C & 0 \\ 0 & \hat{V}_C^T \end{bmatrix} \begin{bmatrix} \hat{\Sigma}_C & 0 \\ 0 & \hat{V}_C^T \end{bmatrix}^{-1} \tag{11}
\]

Here, Eq. \[11\] is singular value decomposition on \( C_n \). Finally, one can take various metrics of the projected covariance matrix Eq. \[10\] as Ref. \[4\], \[5\], \[19\], which are equal to Eq. \[7\] if the determinant is considered.

### C. Greedy algorithm

Algorithm \[1\] shows the procedure implemented in the computation conducted in section \[11\] which implicitly exploits the one-rank determinant lemma as \[13\], \[17\], \[14\]. The equations are converted by the lemma as shown hereafter. First, consider the objective function when there are less sensors deployed than the number of the state variables:

\[
\det \left( R_{k(i)}^{-1} C_{k(i)} C_{k(i)}^T \right) = \det \left( R_{k(i)}^{-1} \right) \det \left( C_{k(i)} C_{k(i)}^T \right)
\]

\[
= \left( u(i) u(i)^T - u(i) C_{k(i)}^{-1} C_{k(i)}^{-1} u(i)^T \right) \left( R_{k(i)}^{-1} C_{k(i)} C_{k(i)}^T \right)^{-1}
\]

\[
= \left( t_{k(i)} - s_{k(i)} R_{k(i)}^{-1} s_{k(i)}^T \right) \left( R_{k(i)}^{-1} C_{k(i)} C_{k(i)}^T \right)^{-1}, \tag{12}
\]

where the subscript \( i \) represents the component produced by the \( i \)-th sensor candidate and \( R_{k(i)} \) is constructed as:

\[
R_{k(i)} = \begin{bmatrix} R_{k(i)} & s_{k(i)}^T \\ s_{k(i)} & t_{k(i)} \end{bmatrix} \tag{13}
\]
Algorithm 1 Determinant-based greedy algorithm with a noise covariance matrix

Set sensor candidate indices and selected indices

\[ S = \{1, \ldots, n\}, S_0 = \emptyset \]

for \( k = 1, \ldots, r, \ldots, p \) do

if \( k \leq r \) then

\[ i_k = \arg\max_{i \in S \setminus S_{k-1}} \det \left( R_{k(i)}^{-1} C_{k(i)} C_{k(i)}^T \right) \]

else

\[ i_k = \arg\max_{i \in S \setminus S_{k-1}} \det \left( C_{k(i)}^T R_{k(i)}^{-1} C_{k(i)} \right) \]

end if

Update sensor selection matrix, measurement matrix and noise-covariance matrix

\[ H_k = \left[ H_{k-1} \epsilon_{ik} \right]^T, C_k = H_k U, \]

\[ R_k = H_k U_N \Sigma_N H_k^T \]

Update selected sensor indices

\[ S_k \leftarrow S_{k-1} \cup i_k \]

end for

Here, \( S_{k(i)} \) and \( t_{k(i)} \) are noise covariance between the selected sensors by the previous steps and the \( i \)th candidate, and noise variance at the \( i \)th candidate, respectively. Expensive computations of taking the determinant are avoided by separating components of the obtained sensors from the objective function in the current selection step:

\[
\therefore i_k = \arg\max_{i \in S \setminus S_{k-1}} \det \left( R_{k(i)}^{-1} C_{k(i)} C_{k(i)}^T \right) \]

\[ = \arg\max_{i \in S \setminus S_{k-1}} \frac{u_{(i)} \left( I - C_{k-1}^{-1} \left( C_{k-1} C_{k-1}^T \right)^{-1} C_{k-1} \right) u_{(i)}^T}{t_{k(i)} - s_{k(i)} R_{k-1}^{-1} s_{k(i)}^T}, \tag{14} \]

and then an unit vector \( \epsilon_{ik} \in \mathbb{R}^{1 \times n} \) which the \( i_k \)-th component is unity goes into the \( k \)-th row of \( H \). Note that Eq. (14) corresponds to maximization of the difference when an arbitrary sensor is added to the sensor set of the previous step. The numerator of Eq. (14) is the L2 norm of vector and the denominator is positive, since the covariance matrix \( R \) is assumed to be positive definite then the determinant of Eq. (13) is positive. These facts ensure the monotone increase in the objective function by adding sensors in a greedy manner. Subsequently, the objective function is modified for the case in which more sensors have already been determined than the number of the state variables. The details of expansion are found in Ref. [14]:

\[
i_k = \arg\max_{i \in S \setminus S_{k-1}} \det \left( C_{k(i)}^T R_{k(i)}^{-1} C_{k(i)} \right) \]

\[ = \arg\max_{i \in S \setminus S_{k-1}} \frac{v_{(i)} \left( C_{k-1}^{-1} R_{k-1}^{-1} C_{k-1} \right)^{-1} v_{(i)}^T}{t_{k(i)} - s_{k(i)} R_{k-1}^{-1} s_{k(i)}^T}, \tag{15} \]

where \( v_{(i)} = s_{k(i)} R_{k-1}^{-1} C_{k-1} - u_{(i)} \). Eq. (15) is positive, and the objective functions Eq. (7) are monotone increase in both cases of \( p < r \) and \( p \geq r \).

The objective function is not shown to have submodularity if the measurement noise has strong correlation with each other.

The following example provides a non-submodular and non-supermodular case for Eq. (7). For simplicity, the spatial modes \( U \in \mathbb{R}^3 \) and the noise covariance \( \tilde{R} \in \mathbb{R}^{3 \times 3} \) are set as follows. Here, noise components \( i = 2, 3 \) are strongly correlated, while those for \( i = 1 \) are relatively independent.

\[ U = \begin{pmatrix} 0.1 & 1 & -0.1 & 0.1 \\ 1 & -0.1 & 0.8 & 0.7 \\ 1 & 0.1 & 0.7 & 2 \end{pmatrix}, \]

With these matrices, values of the determinant function Eq. (7) are

\[ g_{(1,2)} - g_{(1)} = 1.2913 > g_{(1,2,3)} - g_{(1,3)} = 0.8038, \]

\[ g_{(1,3)} - g_{(3)} = 0.0025 < g_{(1,2,3)} - g_{(2,3)} = 0.0450, \]

where \( g_s (s \in \{1,2,3\}) \) refers to the value of the function for the selected sets of sensors. This example immediately shows the objective function Eq. (7b) has neither submodularity nor supermodularity, while the submodularity was confirmed for the case with the equally distributed uncorrelated measurement noise [17]. Thus, solving the sensor selection problem Eq. (7) with a greedy method has generally no apparent performance guarantee based on submodularity or supermodularity.

III. RESULTS

This section concerns some experiments that validate the algorithm. First, data-matrices are constructed from randomly generated orthonormal basis, NOAA SST dataset [20] follows to show results in a practical application.

A. Randomly generated data matrix

Generalized results are shown in this subsection. Problem setting considered here is as follows; data matrix \( X \) is constructed as \( X = U \Sigma X V_X^T \), where \( U_X \) and \( V_X \) are 5000 \times 100 and 100 \times 100 orthonormal matrices generated from appropriate sized matrices containing numbers from standard normal distribution, and \( \Sigma_X \) is diagonal matrix with \( \text{diag}(\Sigma_X) = (1 \ 0.99 \ldots (101 - j)/100 \ldots 0.01) \), respectively. The algorithms for the sensor selection treat these matrices after dividing first 10 columns as \( U \), \( V \) and the rest as \( U_N \), \( V_N \), then first 10 diagonal components as \( \Sigma \) and the rest as \( \Sigma_N \). Note that the measure \( e \) in term of “reconstruction
error" is expressed as \( e = \| X - U \tilde{Z} \|_F / \| X \|_F \). Here, series of the estimation \( \tilde{z} \) of least squares estimation or generalized estimation Eq. (2a) are concatenated as \( \tilde{Z} \), and \( \| \cdot \|_F \) represents the Frobenius norm of \( \cdot \). Figure 1 shows the result of the reconstruction with the estimate through \( p \) sensors and the \( r \) dimensional reduced-order model Eq. (3). Here, DG and DG/NC in the legend refer to “determinant-based Greedy algorithm” in Ref. [17] and the presented algorithm considering “noise covariance” in the measurement, respectively, and LS and GLS to “linear least-squares estimation” and “generalized linear least-squares estimation” using noise covariance, respectively. Note that the plots for \( p \leq r \) are calculated by the same estimator Eq. (2a), and therefore, the estimation with a small number of sensors for both GLS and LS are equal. Firstly, the GLS estimation reduces the reconstruction error in over-determined cases for sensors of both algorithms. Secondly, the more the number of sensors are deployed, the less the reduction thanks to the GLS estimation becomes. This is partly because the measurement through a large number of sensors suppresses the outlier resulted from the correlated measurement noise. If a much larger number of sensors is available than the number of the estimated variables, importance of consideration on the correlation in the measurement noise might diminish.

B. NOAA-SST

Here, we try the strategy to pursue the sensor selection using large dimensional climate data. The brief description for the NOAA-SST data are summarized in Table I.

| Label | Temporal Coverage | Spatial Coverage |
|-------|-------------------|-----------------|
|       | NOAA Optimum Interpolation (OI) SST V2 [20] | 1.0 degree latitude \( \times \) 1.0 degree longitude global grid (\( n = 44219 \) measurement on the ocean) |

| Number of sensors | Reconstruction Error |
|-------------------|----------------------|
| 10                | 0.25                 |
| 20                | 0.3                  |
| 30                | 0.35                 |
| 40                | 0.4                  |
| 50                | 0.45                 |

Table I: Description for SST data

Fig. 2: Comparison of estimation error in the generalized least-squares configuration and the least-squares of SST data

Fig. 3: Twenty sensors selected by the greedy algorithms (red: presented Eq. (7a), black: previous [17] on color map representing root mean square of noise component

DG algorithm [17] owing to the noise covariance matrix in the sensor selection procedure. There are several differences in the trend compared to Fig. 1 e.g. the contribution from the presented algorithm is more significant than that from the GLS estimation. The reason is perhaps the similarity in the location where the reduced-ordered phenomena and the measurement noise greatly fluctuates. The presented algorithm that involves noise covariance evaluates the positions with less measurement noise, therefore accurate estimation is acquired even with the linear least-squares estimation. In these figures, the positions of sensors are represented by the circle marks on the colored map which illustrates the fluctuation of measurement noise. Thus, the diagonal components of \( R \) that correspond to the red-colored region in Fig. 3 are relatively large. The difference in the sensor positions is remarkable, since the presented algorithm spreads sensors to avoid neighboring sensors that might be affected by the correlated measurement noise.

IV. CONCLUSIONS

A greedy algorithm of the sensor placement for the generalized least-squares estimation is provided. A covariance matrix generated in the reduced-order modeling is applied to build the weighting matrix in the estimation. A specialized one-rank lemma involving the covariance matrix realises simple transformation from the true optimization into greedy scalar evaluation. In addition, the objective function is illustrated to be neither submodular nor supermodular, for the first time. Numerical tests using two kinds of dataset assess the presented determinant-based optimization method. The presented algorithm gives less noisy sensors and results in the stable estimation under the measurement noise from truncated modes of the reduced-order modeling.

ACKNOWLEDGEMENTS

This work was partially supported by JST ACT-X (JPMJAX20AD) and JST CREST (JPMJCR1763), Japan.
REFERENCES

[1] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2019.

[2] K. Manohar, B. W. Brunton, J. N. Kutz, and S. L. Brunton, “Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns,” *IEEE Control Systems Magazine*, vol. 38, no. 3, pp. 63–86, 2018.

[3] K. Nankai, Y. Ozawa, T. Nonomura, and K. Asai, “Linear reduced-order model based on piv data of flow field around airfoil,” *Transactions of the Japan Society for Aeronautical and Space Sciences*, vol. 62, no. 4, pp. 227–235, 2019.

[4] Y. Saito, T. Nonomura, K. Nankai, K. Yamada, K. Asai, Y. Sasaki, and D. Tsubakino, “Data-driven vector-measurement-sensor selection based on greedy algorithm,” *IEEE Sensors Letters*, vol. 4, 2020.

[5] S. Joshi and S. Boyd, “Sensor selection via convex optimization,” *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.

[6] K. Nankai, Y. Ozawa, T. Nonomura, and K. Asai, “Linear reduced-order model based on piv data of flow field around airfoil,” *Transactions of the Japan Society for Aeronautical and Space Sciences*, vol. 62, no. 4, pp. 227–235, 2019.

[7] N. K. Dhingra, M. R. Jovanović, and Z.-Q. Luo, “An admm algorithm for optimal sensor and actuator selection,” in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 4039–4044.

[8] T. Nagata, T. Nonomura, K. Nakai, K. Yamada, Y. Saito, and S. Ono, “Data-driven sparse sensor placement based on a-optimal design of experiment with admm,” *arXiv preprint arXiv:2010.09329*, 2020.

[9] A. Zare and M. R. Jovanović, “Optimal sensor selection via proximal optimization algorithms,” in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 6514–6518.

[10] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, “An analysis of approximations for maximizing submodular set functions,” *Mathematical programming*, vol. 14, no. 1, pp. 265–294, 1978.

[11] U. Feige, V. S. Mirrokni, and J. Vondrák, “Maximizing non-monotone submodular functions,” *SIAM Journal on Computing*, vol. 40, no. 4, pp. 1133–1153, 2011.

[12] D. Golovin and A. Krause, “Adaptive submodularity: Theory and applications in active learning and stochastic optimization,” *Journal of Artificial Intelligence Research*, vol. 42, pp. 427–486, 2011.

[13] S. Liu, S. P. Chepuri, M. Fardad, E. Maazade, G. Leus, and P. K. Varshney, “Sensor selection for estimation with correlated measurement noise,” *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3509–3522, 2016.

[14] K. Yamada, Y. Saito, K. Nankai, T. Nonomura, K. Asai, and D. Tsubakino, “Fast greedy optimization of sensor selection in measurement with correlated noise,” *Mechanical Systems and Signal Processing*, vol. 158, p. 107619, 2021.

[15] M. Shamaiah, S. Banerjee, and H. Vikalo, “Greedy sensor selection: Leveraging submodularity,” in *49th IEEE conference on decision and control (CDC)*. IEEE, 2010, pp. 2572–2577.

[16] L. Ye, S. Roy, and S. Sundaram, “On the complexity and approximability of optimal sensor selection for kalman filtering,” in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 5049–5054.

[17] Y. Saito, T. Nonomura, K. Yamada, K. Nakai, T. Nagata, K. Asai, Y. Sasaki, and D. Tsubakino, “Determinant-based fast greedy sensor selection algorithm,” *IEEE Access*, 2021.

[18] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.

[19] K. Nankai, K. Yamada, T. Nagata, Y. Saito, and T. Nonomura, “Effect of objective function on data-driven greedy sparse sensor optimization,” *IEEE Access*, vol. 9, pp. 46731–46743, 2021.

[20] NOAA/OAR/ESRL, “Noaa optimal interpolation (oi) sea surface temperature (sst) v2.” [Online]. Available: [https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html](https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html)