Randomized Group-Greedy Method for Large-Scale Sensor Selection Problems

Takayuki Nagata, Keigo Yamada, Kumi Nakai, Yuji Saito, and Taku Nonomura

Abstract—The randomized group-greedy (RGG) method and its customized method for large-scale sensor selection problems are proposed. The randomized greedy sensor selection algorithm is applied straightforwardly to the group-greedy (GG) method, and a customized method is also considered. In the customized method, a part of the compressed sensor candidates is selected using the common greedy method or other low-cost methods. This strategy compensates for the deterioration of the solution due to compressed sensor candidates. The proposed methods are implemented based on the D- and E-optimal design of experiments, and numerical experiments are conducted using randomly generated sensor candidate matrices with potential sensor locations of 10,000–1,000,000. The proposed method can provide better optimization results than those obtained by the original GG method when a similar computational cost is spent as for the original GG method. This is because the group size for the GG method can be increased as a result of the compressed sensor candidates by the randomized algorithm. Similar results were also obtained in the real dataset. The proposed method is effective for the E-optimality criterion, in which the objective function that the optimization by the common greedy method is difficult due to the absence of submodularity of the objective function. The idea of the present method can improve the performance of all optimizations using a greedy algorithm.

Index Terms—Greedy method, linear inverse problem, optimal design of experiment, randomized algorithm, sensor selection.

I. INTRODUCTION

Measurements of physical phenomena are an important topic in various fields. This may involve surface or volume measurements, and in most cases, such measurements are performed by discretely installed point sensors. This kind of situation can be seen in various types of measurements, such as global positioning system [1], [2], acoustic measurements [3], [4], infrastructure health monitoring [5], [6], [7], [8], environment monitoring [9], [10], brain source localization [11], etc. Although each sensor can only measure quantities at a particular location, full-state recovery can be achieved from sparse observations by solving an inverse problem. It is necessary to carefully determine the position of the sensor and maximize the information obtained by sparse observations to perform the required measurement with the minimum number of sensors. This is referred to as the sensor placement/selection problem. The objective of this problem is to determine/select the $p$ sensor locations from the $n$ potential sensor locations.

Sensor placement/selection problems are formulated as combinatorial optimization problems known as NP-hard problems. The exact solution can be obtained by exhaustive search or global optimization techniques, such as branch and bound [12], [13], but these techniques can only be used for the problem of choosing a small number of sensor locations from a small number of potential sensor locations because of the expensive computational cost. Therefore, an approximated method that can find a suboptimal solution with a reasonable computational cost is an interesting topic.
Joshi and Boyd [14] applied the convex relaxation approach to sensor selection problems, and their method can obtain a global optimal solution to the relaxed problem. The computational complexity is proportional to the third power of the number of potential sensor locations. The sparsity-promoting framework and the proximal splitting algorithm were introduced by Lin et al. [15] and Fardad et al. [16]. Their framework can obtain block-sparse feedback and observer gains as well as select actuators and sensors and was extended by Dhingra et al. [17] and Zare and Jovanović [18]. Furthermore, Nagata et al. [19] proposed a sensor selection method for a nondynamical system based on the proximal splitting algorithm, and the method was extended to the problem with strongly correlated measurement noise [20]. Although the computational cost was significantly reduced, it is still expensive for many degree-of-freedom problems, such as a data-driven sensor selection problem. The greedy method has been studied for solving large-scale sensor selection problems [21], [22], [23], [24], and both convex relaxation and greedy methods have been extended for various purposes [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37]. In addition, Inoue et al. [38] proposed a novel sensor optimization method for a high-dimensional system using an annealing machine.

The number of potential sensor locations in problems to which data-driven methods are often applied can reach more than $O(10^4)$. The data-driven sensor selection problem is the problem for selecting the (sub-)optimal measurement points based on the data, such as planar or volumetric measurement or simulation data. In the case of image-based measurement data, for example, each pixel of images corresponds to a potential sensor location. The selected sensors are used for the full-state recovery based on a tailored basis [21] or neural network [39], [40], [41], for example. Particularly, a framework that conducts full-state estimation from the measurement data at several locations is actively studied in fluid dynamics. In the case of linear estimation with sparse observations, the problem becomes a simple linear inverse problem, and the framework has been applied for the pressure field [42], [43], temperature field [44], [45], and velocity field [46], [47], [48]. This kind of framework has also been applied to aeroacoustic source localization [49] and seismic wavefield reconstruction [50]. In addition, not only the static sensor selection method, but also the dynamic sensor selection method was proposed based on graph sampling theory [51].

The greedy method will outperform convex relaxation methods when the problem size is increased [52], [53], [54]. Therefore, compared to convex relaxation methods, greedy methods are more appealing for sensor placement in a centralized context, especially for large-scale problems. In a greedy algorithm, we iteratively find a new sensor location that greatly improves the objective value until the constraint is satisfied. In each step, we determine one new sensor location as a single-sensor problem, and such a strategy may miss some better solutions. This is one of the difficulties of greedy methods. Improvement in the performance of the selected sensor subset is performed by local optimization. For example, the two-opt algorithm, which is a simple local search algorithm, was combined with the sensor selection method via convex optimization [14]. The two-opt solution infers that there is no swapping of selected and unselected sensors that gives a better objective value is obtained.

Jiang et al. [55] considered this issue more radically. They proposed the group-greedy (GG) method, which can obtain a better sensor subset. Their method is inspired by the beam search algorithm in the area of natural language processing [56]. Their method iteratively reserves a group of suboptimal sensor subsets, where “group” indicates the top $L$ sensor subsets in the sense of a certain objective value. In this way, by considering not only the best configuration but also the suboptimal configuration, better results may be obtained than by simply pursuing only the best configuration in each step. Although the computational cost increases as the group size increases, the GG method can obtain the exact solution when the group size is large enough.

The number of sensor locations in problems to which data-driven methods are often applied can reach more than $O(10^4)$, as is often the case in fluid dynamics. In such a case, not only convex optimization, but also improved greedy methods, such as the GG method, have a problem of computational cost, and thus accelerated methods are required. A similar requirement exists when the objective function is expensive to compute. Recently, randomized methods [57], [58] have been applied to convex/nonconvex problems in signal and image processing [59], [60], [61]. For sensor selection problems, the convex relaxation method proposed by Joshi and Boyd [14] has been accelerated by Nonomura et al. [62] using the randomized subspace Newton method [63]. Hashemi et al. [64] proposed a randomized greedy method that selects sensors for state estimation in large-scale linear time-varying dynamical systems based on the study by Mirzaeileiman et al. [65]. They provided a performance guarantee for the proposed algorithm and demonstrated that the randomized method is superior to the common greedy and semidefinite problem relaxation methods in terms of computational time while providing the same or better utility.

In the present study, we propose a randomized GG (RGG) method and its customized method. The sensor selection problem is such that $p$ sensors are selected from among $n$ potential sensors. Sensor subset gives an observation vector of a linear function of latent variables superimposed with independent identically distributed zero-mean Gaussian random noise. The main contributions of the present article are summarized as follows.

1) The computational cost of the GG method is significantly reduced by introducing the randomization technique.
2) The proposed method can solve large-scale problems with almost the same computational time as the common greedy method, and a better sensor subset can be obtained in terms of the objective value.
3) Because of the reduced computational cost, a search with a larger group size (deeper exploration) can be conducted, compared to the GG method.
4) By appending elite sensor candidates obtained by a low-cost method, such as the common greedy method, to the random sketch of the sensor candidates, the performance of the selected sensor subset is further improved in exchange for a slight increase in computational cost.

The sensor optimization method can also be used for actuator optimization, which is a dual problem of sensor optimization. Furthermore, the idea of the present method can improve the performance of all optimizations using greedy methods, not just sensor/actuator optimization. MATLAB code for our algorithm can be found at https://github.com/Aerodynamics-Lab/Elite-and-Randomized-Group-Greedy-Method [66].

II. PROBLEMS AND ALGORITHMS

A. Sensor Selection Problems

The purpose of the present study is to improve the sensor selection method based on the greedy strategy. Particularly, the proposed method was intended to be applied to large-scale sensor selection problems, such as the sensor selection for full-state recovery problems using the data-driven reduced-order model [21]. This problem can be written as a linear inverse problem, and thus, the present work focuses on a linear inverse problem with uncorrelated measurement noise. It should be noted that the present method can be applied to the sensor selection method based on the greedy algorithm considering correlated measurement noise, such as the methods proposed by Yamada et al. [32] and [33] and other problems written as linear inverse problems, such as node selection in a wireless network [67]. Furthermore, the idea of the present method can improve the performance of all optimizations using greedy methods, not just sensor optimization problems.

A snapshot measurement of the full state \( x \in \mathbb{R}^n \) including noise signals \( v \sim N(0, \sigma^2 I) \in \mathbb{R}^p \) by sparse sensors can be expressed as follows:

\[
y = Hx + v
\]

where \( y \in \mathbb{R}^p \) and \( H \in \mathbb{R}^{p \times n} \) are the observation vector and the sensor location matrix, respectively. Here, \( n \) and \( p \) are the numbers of potential sensors and selected sensors, respectively. The sensor location matrix \( H \) is a sparse matrix, and the activated sensor location is indicated as an entry of unity. Each row vector of \( H \) is a unit vector. Equation (1) can be rewritten as follows using the sensor candidate matrix \( U \in \mathbb{R}^{n \times r} \) and the latent variables \( z \in \mathbb{R}^r \):

\[
y = HUz + v = Cz + v.
\]

This system represents the problem of choosing \( p \) sensors for the observation of signals generated by \( r \) latent variables. The matrix \( C \in \mathbb{R}^{p \times r} \) is the measurement matrix, which is the product of the sensor location matrix \( H \) and the sensor candidate matrix \( U \). Because \( n \) corresponds to the degrees of freedom in the spatial direction of the full state [i.e., \( n \geq O(10^6) \)], the sensor candidate matrix \( U \) is usually a tall-and-skinny matrix (i.e., \( n \gg r \)) in a practical problem. The sensor locations are selected based on the optimal design of experiments [68]. The E- and D-optimality criteria that are frequently used are considered

\[
f_E = \begin{cases} 
\lambda_{\min} (CC^T), & p \leq r \\
\lambda_{\min} (C^T C), & p > r 
\end{cases}
\]

\[
f_D = \begin{cases} 
\log(\det(CC^T)), & p \leq r \\
\log(\det(C^T C)), & p > r.
\end{cases}
\]

Maximization of \( f_E \) and \( f_D \) corresponds to the worst case error variance and the minimization of the volume of the confidence ellipsoid, respectively. Sensor selection methods based on the common greedy and GG method using these objective functions are described in [22], [24], and [53]. The objective function \( f_E \) is not a submodular function, and thus, the optimization using the common greedy method does not work well [24]. Therefore, the benefits of introducing the group strategy are great. On the other hand, \( f_D \) has better characteristics for greedy optimization than \( f_E \) [24]. Although the objective function \( f_D \) is also not always submodular, the objective function that is slightly modified one from \( f_D \) is a monotone submodular function in the conditions with independent white noise considered [22] and can be optimized with a good approximation guarantee even by using the common greedy method. Therefore, the advantage of the GG method in the case of D-optimality-based methods is less than those in the case of E-optimality-based methods.

B. GG Method

The GG method was proposed by Jiang et al. [55]. This method can be applied to all methods based on the greedy algorithm, and the performance of these methods in terms of the objective value improves. In the common greedy method, the best location is repeatedly selected in each single-sensor subproblem, and thus, the optimal solution can easily be overlooked. The GG method reserves some other suboptimal configurations other than the current best configuration obtained by the greedy strategy, which may contain a subset of the final optimal solution, as shown in Table I. In each step of the GG method, the \( L \) suboptimal configurations are reserved based on the objective values of each configuration, that is, the suboptimal configurations for the \( p \)th sensors are \( \{S_{p,1}, S_{p,2}, \ldots, S_{p,L}\} \). The group size, which is the number of reserved configurations, is \( L = 2 \) in Table I. In this case, the best two configurations are reserved based on the objective value. The objective value of the obtained sensor configuration at the final step is improved as the group size is increased, but the number of evaluations is also increased, which results in a longer computational time. It should be noted that the result obtained by the GG method is not always better than that obtained by the common greedy method, as shown in Table II. In this example, the GG method could not reach the optimal solution, but the common greedy method reaches the optimal solution. This kind of situation might occur when the group size is not sufficient. Detailed descriptions including the theoretical guarantee can be seen in [37].
TABLE I
PROCEDURE OF SENSOR SELECTION VIA GG METHOD \((L = 2)\)

| \(p\) | Potential choices | Best 2 configurations | Sensor selection |
|------|-------------------|-----------------------|-----------------|
| 1    | \{1\}, \{2\}, \{3\}, \{4\}, \{5\} | \(S_{1,1} = \{1\}, S_{1,2} = \{5\}\) | \(S = S_{1,1} = \{1\}\) |
| 2    | \{1.2\}, \{1.3\}, \{1.4\}, \{1.5\}, \{2.5\}, \{3.5\}, \{4.5\} | \(S_{2,1} = \{1, 2\}, S_{2,2} = \{1, 5\}\) | \(S = S_{2,2} = \{1, 5\}\) |
| 3    | \{1.2.3\}, \{1.2.4\}, \{1.2.5\}, \{1.3.5\}, \{1.4.5\} | \(S_{3,1} = \{1, 3, 5\}, S_{3,2} = \{1, 2, 4\}\) | \(S = S_{3,1} = \{1, 3, 5\}\) |

TABLE II
EXAMPLE OF THE FAILURE CASE OF GG METHOD \((L = 2)\)

| \(p\) | Common greedy | Group greedy | ERGG |
|------|---------------|--------------|------|
| 1    | \{1\}, \{2\}, \{3\}, \{4\}, \{5\} | \{1\}, \{2\}, \{3\}, \{4\}, \{5\} | \{1\}, \{2\}, \{3\}, \{4\}, \{5\} |
| Sensor selection | \(S_G = \{1\}\) | \(S_{GG1,1} = \{1\}, S_{GG1,2} = \{2\}\) | \(S_{ERGG} = S_{GG1,2} = \{2\}\) |
| 2    | \{1.2\}, \{1.3\}, \{1.4\}, \{1.5\} | \{1.2\}, \{1.3\}, \{1.4\}, \{1.5\}, \{2.3\}, \{2.4\}, \{2.5\} | \{1.2\}, \{1.3\}, \{1.4\}, \{1.5\}, \{2.3\}, \{2.4\}, \{2.5\} |
| Sensor selection | \(S_G = \{1, 2\}\) | \(S_{GG1,1} = \{2, 3\}, S_{GG2,2} = \{2, 4\}\) | \(S_{ERGG} = S_{GG1,2} = \{2, 3\}\) |
| 3    | \{1.2.3\}, \{1.2.4\}, \{1.2.5\} | \{1.2.3\}, \{1.2.4\}, \{1.2.5\}, \{2.3.5\}, \{2.4.5\} | \{1.2.3\}, \{1.2.4\}, \{1.2.5\}, \{2.3.5\}, \{2.4.5\} |
| Sensor selection | \(S_G = \{1, 2, 3\}\) (optimal) | \(S_{GG1,1} = \{1, 2, 3\}, S_{GG2,2} = \{1, 2, 4\}\) | \(S_{ERGG} = S_{GG1,2} = \{1, 2, 3\}\) |

C. Proposed Methods

The GG method can obtain better optimization results, but the computational cost becomes a critical issue for large-scale problems. A problem can have more than \(O(10^4)\) potential sensor locations in some applications, such as data-driven sensor selection [21]. The same problem occurs when the computational cost of the objective value is expensive. In the present study, a randomization technique is introduced to reduce the computational cost. In addition, the proposed method can have merit even though the problem size is intermediate because the proposed method can conduct optimization with a larger group size than the original GG method because of the reduced computational cost.

The sparse sampling assumes a low rankness of the sampling target. On the other hand, the number of potential sensor locations is large. In particular, the number of potential sensor locations in problems to which data-driven methods are often applied can reach more than \(O(10^6)\). In such a case, there is a large number of potential sensor locations, but a small number of sensors are enough to solve an inverse problem. Therefore, compression of the sensor candidate matrix by projecting onto the randomized subspace is effective.

In the proposed method, shown in Algorithm 1, the compressed sensor candidate matrix is generated before selecting sensors by the GG method, as shown in Algorithm 2. The RGG method conducts sensor selection by the GG method in the compressed sensor candidates, which is the random subset \(S_c\) of the original sensor candidates \(S_r\). The number of sensor candidates is compressed from |\(S_r| = n\) to |\(S_c| = n_c\) (the compression ratio \(n_c/n\) is 1/10 or 1/100, for example), and the number of evaluations is reduced from \(Ln\) to \(Ln_c\), where \(n_c\) is the size of the compressed sensor candidate generated by random sampling. Here, the compression ratio is one of the hyperparameters of the present method, and \(n_c/n\) of 1/10 or 1/100 are just the values used in the present experiment. If the compression ratio is unity, the proposed method is the same as the GG method. The computational cost of the proposed method becomes small as the compression ratio decreases for the constant group size. The informative locations in the sensor candidates where the objective function can greatly be improved are possibly truncated when generating a compressed sensor candidate matrix by random selection, and the optimization result might be degraded. Therefore, an elite strategy is introduced.

The elite strategy is the crucial technique to compensate for the degradation of the optimization results in the RGG method due to the compressed sensor candidate matrix. Possible informative locations are secured as the elite sensor candidates by the low-cost optimization method, such as the common greedy method. The elite sensor candidates will be appended to the compressed sensor candidate, which is generated by random selection, and the performance degradation due to the compression of the sensor candidate matrix can be minimized. The RGG method with the elite strategy is called the elite-and-RGG (ERGG) method in this article. In this method, the compressed sensor candidate \(S_c\) is \(S_c \sqcup S_e\), where \(S_e\) is the elite sensor candidate, and the sizes of these sets are |\(S_e| = n_c\), |\(S_c| = n_r\), and |\(S_e| = n_e\), respectively. Note that \(n_e\) in the ERGG method was determined so that \(n_c\) becomes the same as that of the RGG method in the present numerical experiments. Although elite sensor candidates can be selected in various ways, those are selected using the common greedy method in the present study to simplify the discussion in the present study.

The computational complexities of the present and previously proposed methods are shown in Table III. For large-scale sensor selection problems, the number of potential sensor locations is much larger than that of the number of latent variables and the number of sensors to be selected, i.e., \(p \approx r \ll n\). The proposed method reduces the computational cost by compressing the number of potential sensor locations. In the present study, the elite sensor candidates for ERGG are selected by the common greedy method, and thus, the additional cost of \(n_c r r^2\) is required. This additional cost depends on the complexity of the method used for selecting the elite sensor candidates.

III. RESULTS AND DISCUSSION

The performance of the proposed methods was evaluated by applying those methods to randomly generated sensor
The number of latent variables was set to $n$, and the number of sensor candidates was set to $p$. The entries follow a normal distribution of $N(0,1)$. The number of sensor candidates was set to $n = 10000$, and the number of latent variables was set to $r = 10$. Since the number of latent variables was $r = 10$, $p < 10$, and $p > 10$ are the underdetermined and overdetermined conditions, respectively. The computations for each condition were conducted 500 times with a different sensor candidate matrix, and the average values of the objective values (E- and D-optimality criteria) and computational time were evaluated. Optimization in randomly generated sensor candidate matrices corresponds to the sensor optimization in the featureless dataset, which represents the average performance on various kinds of datasets. The objective function used in the present study is based on the optimal design of experiment.

### Algorithm 1 ERGG Method [66]

**Input:** $L \in \mathbb{N}, S_c$

**Output:** Indices of the top subset for $p$ sensor locations $S_{\text{selected}}$

1. Set $k \leftarrow 1$, reserved sensor configurations $S_{\text{reserved}} \leftarrow \emptyset$
2. For $k \leq p$ do
   - If $k = 1$ then
     - Set $S_r$ by preliminary random selection from $S_c \setminus S_e$
     - Set combined candidate $S_c \leftarrow S_r \cup S_e$
     - Alg. 2 with $S_{\text{reserved}}, S_c$ and $L$
   - Else
     - For $l \in \{1, 2, \ldots, L\}$ do
       - Set $S_r$ by preliminary random selection from $S_c \setminus S_e$
       - Set combined candidate $S_c \leftarrow S_r \cup S_e$
       - Alg. 2 with $S_{\text{reserved}}, S_c$ and $L$
     - End for
   - Store all $F$ and $T_k$
   - End if
3. End for
4. End for
5. Select best sensor subset $S_{\text{selected}} \leftarrow S_{\text{reserved}}^{(p,1)}$

### Algorithm 2 L-Best Greedy Search for the $k$th Sensor

**Input:** $S_r^{(k-1),l} = \{i_1, i_2, \ldots, i_{k-1}\}, S_c, \text{ L } \in \mathbb{N}$

**Output:** $F \in \mathbb{R}^L, T_k \in \mathbb{R}^{L \times k}$

1. Calculate objective values $f(S_r^{(k-1),l} \cup i)$ for all $i \in S_c \setminus S_r^{(k-1),l}$
2. $F \leftarrow \text{Best } L \text{ objective values of } f$
3. $T_k \leftarrow \text{Corresponding } L \text{ subsets of sensor location}$

### Table III

| Method                        | Complexity         |
|-------------------------------|--------------------|
| Common greedy method          | $O(pmn^2)$         |
| Group-greedy method           | $O((Lpm)^2)$       |
| Randomized group-greedy method| $O((Lpm, r^2)$      |
| Elite and randomized group-greedy method | $O((Lpm, r^2 + npn^2)$ |

This statistical criterion is directly related to the error in the estimation of the linear inverse problem [68]. Better results in the objective function lead to better estimation errors in the linear inverse problem [24].

### A. Performance of Common Greedy and GG Method

The performances of the common greedy [22], [24] and GG methods [55], which are the basis of the proposed method, are compared with random selection and convex relaxation method [14] based on the Newton method. The objective function of all methods except for the random selection is the D-optimality criterion $f_D$. The objective values $f_D$ normalized by the value obtained by the common greedy method $f_D$ is shown in Fig. 1. It notes the sensor selection method based on the greedy method proposed by Manohar et al. [21] (QR-based method) is mathematically the same as the D-optimality-based common greedy (D-G) method at $r < p$, and the performance of the D-G method is better than that of the QR-based method at $r > p$ due to mathematical issues. The objective value obtained by random selection was also calculated as the mean value of the objective values obtained by 100 trials for each sensor candidate matrix.

The objective value obtained by the random selection is much lower than that of the optimized ones. The objective values obtained by the convex relaxation method and the common greedy and GG methods are similar to each other, except for around $p = 10$. Although the optimization method based on the greedy strategy can only obtain a suboptimal solution, the objective value obtained by the D-G method is higher or almost the same as that obtained by the convex relaxation method. The convex relaxation method can obtain a better solution than the D-G method when the number of selected sensors is large, but the difference in the obtained objective value is small. Although the increment of the objective value is small due to the characteristics of the objective function, the objective value obtained by the D-optimality-based GG (D-GG) method is larger than or almost the same as that obtained by the D-G method and the convex relaxation method. In addition, the computational time of the convex relaxation method is much longer than that of the greedy method [22]. From these results, the sensor selection
method based on the greedy method is a favorable choice, especially when the number of sensor candidates is large.

B. Comparison With Previously Proposed Greedy Methods

The performance of the proposed methods was compared with that of the common greedy [22], [24] and GG methods [55]. Numerical experiments with two parameter settings were conducted. For the first case, the parameters for the RGG and ERGG methods were set so that the number of evaluations would be the same as for the common greedy method \((L_{RGG}n_c = L_{ERGG}n_c = n)\). In this case, the size of the compressed sensor candidate matrix was set to \(n_c = 1000\), and the group size for the GG, RGG, and ERGG methods was set to \(L = 10\). For the second case, the parameters for the RGG and ERGG methods were set so that the number of evaluations would be the same as that for the GG method at \(L_{GG} = 10\) \((L_{RGG}n_c = L_{ERGG}n_c = L_{GG}n)\). In this case, the size of the compressed sensor candidate matrix and the group size for the RGG and ERGG methods were set to \(n_c = 1000\) and \(L = 100\), respectively. The number of elite sensor candidates for the ERGG method was fixed at \(n_c = 100\) (i.e., \(n_c = 900\)). The influence of the group size and the number of compressed sensor candidates on the performance of the proposed methods will be discussed in Section III-C.

1) Comparison at Constant Group Size (\(L = 10\)): Objective values obtained by each method are compared in Fig. 2. The vertical axis of Fig. 2(a) and (b) is normalized with the common greedy methods 

\[
\frac{f_E}{f_E, E-G} \quad \text{and} \quad \frac{f_D}{f_D, D-G}
\]

The error bars indicate the standard deviation in 500 times trials with different sensor candidate matrices.

The performance of the GG method is the best under most conditions of \(p \leq 50\). Note that the number of evaluations for the GG method is \(L_{GG}n = 100,000\), and the number of evaluations for the RGG and ERGG is ten times less than that of the GG methods. The standard deviation for the \(E\)-optimality-based method is small in underdetermined conditions, but it becomes large in overdetermined conditions. On the other hand, the standard deviation for the \(D\)-optimality-based method is large in underdetermined conditions, but it becomes small in overdetermined conditions. There is no large difference in the standard deviation for each algorithm, and thus, the error bars are removed in the subsequent figures for a clear presentation.

The objective values obtained by the RGG method are decreased compared to those obtained by the common greedy methods due to the compressed sensor candidate matrix. However, the performance of the RGG method is improved by introducing the elite sensor candidates (ERGG method), and the objective values obtained by the RGG and ERGG methods are asymptotic to those obtained by the GG method for larger \(p\).

There are several differences in the trend of the objective values between the \(E\)-optimality-based and \(D\)-optimality-based methods. The performance of the \(E\)-RGG method is superior to that of the \(E\)-G method in overdetermined conditions, but the objective value obtained by the \(D\)-RGG method never exceeds the value obtained by the \(D\)-G method for \(p \leq 50\). This is because \(f_D\) have better characteristics for greedy optimization than the \(E\)-optimality-based method [24]. The same as \(f_E\), the objective function \(f_D\) is also not always submodular but the objective function that is slightly modified one from \(f_E\) is a monotone submodular function in the conditions with independent white noise considered [22] and can be optimized with a good approximation guarantee, even if the common greedy method is used. The objective value obtained by the \(D\)-RGG method does not exceed that obtained by the \(D\)-G method for \(p \leq 50\), and thus, the \(D\)-RGG method is not effective. However, degradation of the solution due to the compressed sensor candidate matrix is compensated in the \(D\)-ERGG method by introducing the elite sensor candidate matrix, and the objective values obtained by the \(D\)-ERGG method are approximately the same as those obtained by the \(D\)-GG method.

The objective function \(f_E\) is not a submodular function [24], and thus, optimization using the common greedy method does not work well. Therefore, the effectiveness of the group strategy is significant [55], and the performance of the \(E\)-RGG method in overdetermined conditions is better than that of the \(E\)-G method, although the sensor candidate matrix is compressed. The differences in relative performance between the \(E\)-G and \(E\)-RGG methods in underdetermined and overdetermined conditions are caused by the nature of the \(E\)-G method. The performance of the \(E\)-G method is relatively high in the underdetermined conditions, but it gets rapidly worse in overdetermined conditions [24]. Hence, the performance of
the E-RGG method is inferior in underdetermined conditions due to the compressed sensor candidate matrix and is superior in overdetermined conditions because of the benefit brought by the group strategy, even though the sensor candidate matrix is compressed. The objective values obtained by the randomized methods are asymptotic to those obtained by the E-GG method at $p > 20$, the performance of the E-ERGG method is higher. However, unlike the D-optimality-based method, the E-ERGG method is significantly inferior to the E-GG method in underdetermined conditions. This is because the performance of the E-G method used to select elite sensor candidates is poor due to the lack of submodularity of the objective function $f_E$.

A comparison of the computational time is shown in Fig. 3. The computational time for the GG method rapidly increases as the number of sensors to be selected increases. On the other hand, the computational time for the RGG method is the same as that for the common greedy method because the number of evaluations is the same. Although the computational time is the same as that for the common greedy method, the RGG method can obtain a better solution in overdetermined conditions when the objective function is the E-optimality criterion. The increase in the computational time for the ERGG method with respect to the number of sensors to be selected is on the same level as that for the common greedy method, but the cost for selecting the elite sensor candidates is appended. By tolerating this additional cost, a significant performance improvement can be achieved, and the computational time is shorter than that of the GG method when the number of sensors to be selected is large.

2) Comparison at Equivalent Number of Evaluations: Comparisons of the objective value when the number of evaluations in the RGG and ERGG methods is the same as in the GG method are presented in Fig. 4. The numbers of evaluations for the GG, RGG, and ERGG methods are $L_{GGn} = L_{RGGn} = L_{ERGGn}$. At $p \leq 15$, the performance of the E-RGG method is degraded compared to the E-GG method, although the number of evaluations is the same. This is because the informative locations are missed due to compression of the sensor candidates in the case of the E-RGG method. Degradation of the solution due to the compressed sensor candidate matrix can be reduced by using the E-ERGG method. However, the common greedy method was used for selecting the elite sensor candidates in the present study, and thus the objective value is still degraded compared with that for the E-GG method. On the other hand, the objective value obtained by the E-RGG method is superior to that obtained by the E-GG method at $p \geq 15$, even though the computational cost is the same. The objective value obtained by the E-ERGG method is further improved in exchange for a slight increase in the computational time.

The objective value obtained by the D-RGG method is smaller than not only that obtained by the D-GG method but also that obtained by the D-G method, despite the same number of evaluations as the D-GG method. Although the difference in the objective values obtained by the D-GG and D-ERGG methods is small, the objective value obtained by the D-ERGG method is larger than that for the D-GG method, particularly in overdetermined conditions.

### C. Influence of Hyperparameters on Performance

The proposed method has three parameters, which are the group size $L$, the size of the compressed sensor candidate matrix $n_c$, and the number of elite sensor candidates $n_e$. The effect of the group size on the obtained objective value is shown in Fig. 5. The number of evaluations for the GG, RGG, and ERGG methods is the same ($L_{GGn} = L_{RGGn} = L_{ERGGn}$) for the same line colors in Fig. 5. The compression ratio for the RGG and ERGG methods was set to $n_c/n = 0.1$, where $n = 10000$.

In the case of E-optimality-based methods, there is no large impact of the group size on the objective value obtained in each method at $p = 3$, but the objective values are different between the E-GG, E-RGG, and E-ERGG methods. The influence of the group size on the objective value appears at larger $p$ in underdetermined conditions. Although the objective values are different in each method, the increase in the objective value by increasing $p$ is almost the same when the number of evaluations is the same ($L_{GGn} = L_{RGGn} = L_{ERGGn}$). As discussed in Section III-B, the performance of the E-RGG method is degraded compared to the E-GG method due to the compressed sensor candidate matrix. In particular, although the number of evaluations for the E-RGG method is ten times greater than that for the E-GG method, the objective...
value obtained by the E-RGG method with $L = 500$ is smaller than that obtained by the E-GG method with $L = 5$. Therefore, the compression ratio is more important than the number of evaluations and the group size in underdetermined conditions.

For overdetermined conditions, the influence of the group size on the objective value depends on the employed algorithm. In the case of the E-GG method, there is no large effect of the group size on the increase in the objective value. On the other hand, the increase in the objective values obtained by the E-RGG and E-ERGG methods is larger than that obtained by the E-GG method. The increase in the objective values obtained by the E-RGG and E-ERGG methods is similar, that is, the difference in the objective value in the underdetermined condition remains even at larger $p$ when the group size is the same.

Although the objective value obtained by the E-RGG method is quite smaller than that obtained by the E-GG method under underdetermined conditions, the objective values at $p > 15$ obtained by the E-RGG and E-ERGG methods are larger than that obtained by the E-GG method when the number of evaluations is the same. In particular, the objective values at larger $p$ obtained by the E-RGG and E-ERGG methods with $L = 50$ are better than that obtained by the E-GG method with $L = 50$, even though the number of evaluations for the E-GG method is ten times larger than that for the E-RGG and E-ERGG methods. Hence, it is considered that the group size is more important than the size of the sensor candidate matrix, unlike underdetermined conditions.

This indicates that the combination of the selected sensor locations is more important than the selected location itself in overdetermined conditions. Consequently, the objective values obtained by the E-RGG and E-ERGG methods are improved by increasing the group size. This point will be discussed in detail in Fig. 6. For the large-scale problem, the group size is limited due to computational cost, and thus, the proposed methods that can increase the group size by reducing the computational cost using the randomized algorithm are effective.

In the case of the D-optimality-based methods, the effect of the group size on the objective value is smaller than that of the E-optimality-based methods. Even if the group size for the D-RGG method increases until $L = 500$, the obtained objective value is smaller than that obtained by the D-GG method with $L = 5$. Hence, the RGG method is not effective for the optimization of the D-optimality criterion.

Since the EGG and ERGG methods use the compressed sensor candidate matrix, the number of evaluations is a function of both the group size and the compressed sensor candidate matrix $n_c$. Fig. 6 shows the influence of the size of the compressed sensor candidate matrix on the obtained
Fig. 6. Effects of the compression ratio $n_c/n$ on objective values of the E-RGG and E-ERGG methods. The number of evaluations for the E-RGG and E-ERGG methods are the same as that for the E-G method ($L_{\text{RGG}}n_c = L_{\text{ERGG}}n_c = n$).

Objective value. It should be noted that the group size was set so that the computational cost for the RGG and ERGG methods would be the same as that for the common greedy method (i.e., $L_{\text{RGG}}n_c = L_{\text{ERGG}}n_c = n$). In the case of the E-RGG method, the obtained objective value is smaller than that obtained by the E-G method in underdetermined conditions and becomes small as the size of the compressed sensor candidates $n_c$ decreases. As the number of selected sensors increases, the objective value obtained by the E-RGG method becomes larger than that obtained by the E-G method. Furthermore, the objective value obtained with smaller $n_c$ (i.e., larger group size) is larger, and the trend is different for a smaller $p$. This trend indicates that the deeper search, i.e., larger group size, is effective in overdetermined conditions, even if $n_c$ is small. For the E-ERGG method, the objective value in underdetermined conditions is better than those obtained by the E-G and E-RGG methods because the locations that are possibly informative are secured by the elite strategy, and there is a small influence of $n_c$. The increase in objective values by increasing the number of sensors for the E-RGG and E-ERGG methods in overdetermined conditions is similar when $n_c$ is the same. The difference between the objective values for these two methods is due to the influence of the difference in the sensor subset selected in underdetermined conditions. The trend for the D-optimality-based methods is similar to that for the E-optimality-based methods, but the D-RGG method cannot be superior to the D-G method.

D. Effect of Number of Sensor Candidates

The effect of the number of sensor candidates on the objective value and computational time is shown in Figs. 7 and 8. Although the computational complexity of the proposed method is in square order of the number of the latent state variable, the number of potential sensor locations is much larger than that of the latent variables ($n \gg r$). Therefore, the effects of the number of sensor candidates are investigated. The group size and the compression ratio were set to $L = 10$ and $n_c/n = 1/100$, respectively, and thus, the number of evaluations for the E-RGG and E-ERGG methods was $1/10$ of the E-G method. Here, the number of elite sensor candidates for the E-ERGG method was fixed at $n_e = 100$. The number of selected sensors was $p = 50$.

Overall, the advantage of the GG strategy in the objective value is kept even if the number of potential sensor locations is $n = O(10^6)$. The computational time for selecting $p = 50$ sensors by the E-GG method is longer than 1000 s. Although the objective value obtained by the E-ERGG method is larger than or similar to the value obtained by the E-GG method, the E-ERGG method has an advantage in terms of computational time. On the other hand, the E-RGG method can obtain the result with $1/100$ computational time of the E-GG method. In this case, the objective value obtained by the E-RGG method is smaller than that obtained by the E-GG method, but the E-RGG method can obtain a better objective value than the E-G method with similar computational time. Therefore, the proposed methods are very effective for large-scale problems.

IV. APPLICATION TO REAL DATASET

The proposed method was applied to the optimization of sensor locations for sparse reconstruction based on a data-driven basis [21]. The measurement points for the sparse reconstruction were selected by the proposed methods. The adopted dataset was the time-series streamwise velocity field of flow over an NACA0015 airfoil. This dataset was obtained experimentally in a wind tunnel experiment with particle
The data matrix, which consists of \( m \) snapshots with spatial dimension \( n \) (\( X = \{x_1, x_2, \ldots, x_m\} \)), is decomposed into the left singular matrix \( U \in \mathbb{R}^{n \times m} \), which shows the spatial modes; the diagonal matrix of singular values \( S \in \mathbb{R}^{m \times m} \); and the right singular matrix \( V \in \mathbb{R}^{m \times m} \), which shows temporal modes, and the rank-\( r \) reduced-order modeling of a data matrix is given as follows:

\[
X \approx U_{1:r} S_{1:r} V_{1:r}^T = U_{1:r} Z.
\]

Therefore, the measurement matrix \( C \) in (2) is \( HU_{1:r} \), and the latent variables are the coefficients of the spatial modes \( Z \).

In the present experiment, the rank of the reduced-order model was set to \( r = 10 \). The used dataset consists of 1000 snapshots with 9353 points of the velocity data (originally the dataset has 14641 points, and 5288 points were masked). The time-series data with 1000 snapshots were split into ten segments with 100 snapshots, and ten-fold cross-validation was conducted. The training data \( X_{\text{train}} \) and the test data \( X_{\text{test}} \) were 90% and 10% of the original data, respectively. The obtained objective values and the reconstruction error were averaged.

Fig. 9(a) and (b) show the snapshot of the original velocity field and low-dimensionalized velocity field, respectively. Because of reduced-order modeling, detailed characteristics of the velocity field are truncated in the snapshot of the low-dimensionalized velocity field, but the main features are retained. The gray and black regions are airfoil and masked regions, respectively. Examples of the locations of selected sensors and reconstructed velocity fields are displayed in Fig. 9(c)–(g). The number of selected sensor locations is \( p = 20 \). Reconstruction with randomly selected sensors cannot reproduce the velocity field properly, but the sensors selected by the greedy methods can reproduce the main features of the velocity field. The sensors are located behind the airfoil, particularly in the shear layer and wake, which are fluctuating regions in time.

Fig. 10(a) illustrates the relationship between the number of sensors and the objective value obtained. The objective values were normalized by that obtained by the E-G method. The group sizes for the E-GG method and the E-RGG and E-ERGG methods were set to \( L_{\text{GG}} = 5 \) and \( L_{\text{RGG}} = L_{\text{ERGG}} = 50 \), respectively. The size of the compressed sensor candidate was set at \( n_c = 1000 \) (\( n_c/n \approx 1/10 \) and \( L_{\text{RGG}} n_c = n \)) for the E-RGG and E-ERGG methods, and the size of the elite sensor candidate was \( n_e = 100 \). The error bars indicate the standard deviation obtained in the ten-fold cross-validation. The trend in the objective value is similar to that obtained in the experiment with the randomly generated sensor candidate matrices. The proposed methods show better performance than the common greedy and GG methods.

Fig. 10(b) shows the effect of the number of sensors on the reconstruction error. The definition of the reconstruction error is as follows:

\[
\epsilon_{\text{reconst}} = \frac{||X_{\text{test}} - U_{\text{train},:r} \tilde{Z}||_F}{||X_{\text{test}}||_F}
\]

where \( \tilde{Z} \) is the estimated mode coefficients.

In the present study, the estimation was conducted by the least-square estimation, which assumes uniform independent Gaussian noise \( v \sim \mathcal{N}(0, \sigma^2 I) \) as observation noise.

\[
\tilde{Z} = C^T Y = \begin{cases} 
C^T (C C^T)^{-1} Y, & p \leq r \\
(C C^T)^{-1} C Y, & p > r
\end{cases}
\]

Here, \( Y \in \mathbb{R}^{n \times m} \) is the matrix that consists of the observation vectors of the test data with \( m \)-snapshots \( HX_{\text{test}} = Y = \{y_1, y_2, \ldots, y_m\} \), and the notation \( c^\dagger \) indicates the pseudo-inverse operation.
The present study proposed the RGG method and its customized method. The E-optimality-based method and D-optimality-based method were implemented, and the performance of the proposed methods was compared with that of the common greedy method and the GG method. The performance evaluation was conducted by applying the methods to a randomly generated dataset in which the entries follow a normal distribution and have $10^4$–$10^6$ potential sensor locations. In addition, the effectiveness of the proposed method was confirmed by the experiment using the real dataset.

The results of the numerical experiment showed that the proposed method can obtain similar or better performance compared to the GG method, while significantly reducing the computational cost. In particular, the proposed method was effective for the optimization of the E-optimality criterion, in which the objective function that the optimization by the common greedy method is difficult due to the absence of submodularity.

When the number of evaluations of the objective function is the same as that for the common greedy method, the RGG method is inferior to the common greedy method in terms of the objective value in underdetermined conditions but superior in overdetermined conditions. By introducing the elite strategy, in exchange for a slight increase in computational cost, the common greedy method is surpassed in all conditions from the underdetermined to the overdetermined conditions. Furthermore, although the number of evaluations in the ERGG method is lower than that for the GG method, it is possible to obtain the same level of performance in overdetermined conditions as that for the GG method.

Since the computational cost of the proposed method is low, optimization with a larger group size is possible when the computational cost of the RGG method is the same as that of the GG method. In particular, the optimization results obtained by the proposed methods are better than those obtained by the GG method when the number of sensors is large in overdetermined conditions. The performance of the proposed method is further improved by including elite sensor candidates into the compressed sensor candidates, and a similar performance to that of the GG method can be achieved, even if the condition is underdetermined. The characteristics of the proposed method for the E-optimality-based method in the real dataset were similar to those obtained in the randomly generated dataset.

The degradation of performance of the D-optimality-based method due to compressed sensor candidates selected purely randomly is significant compared with the E-optimality-based method. This is because the D-optimality criterion has better characteristics than the E-optimality criterion for greedy optimization, and thus, the proposed method using the randomized algorithm is not effective. However, the involvement of elite sensor candidates enables us to obtain similar or better performance compared with the GG method while reducing the computational cost. The proposed method is effective for problems involving a large number of sensor candidates, such as data-driven sensor selection problems. In addition, the proposed method is also effective when the computational cost of the objective value is expensive.

The sensor optimization method can also be used for actuator optimization, which is a dual problem of sensor optimization. Furthermore, the idea of the present method can improve the performance of all optimizations using greedy methods, not just sensor/actuator optimization.

**References**

[1] M. Kihara and T. Okada, “A satellite selection method and accuracy for the global positioning system,” *Navigation*, vol. 31, no. 1, pp. 8–20, Mar. 1984.
[2] M. S. Phatak, “Recursive method for optimum GPS satellite selection,” IEEE Trans. Aerosp. Electron. Syst., vol. 37, no. 2, pp. 751–754, Apr. 2001.

[3] K. Dogancay and H. Hmam, “On optimal sensor placement for time-difference-of-arrival localization utilizing uncertainty minimization,” in Proc. 17th Eur. Signal Process. Conf., 2009, pp. 1156–1160.

[4] R. Macho-Pedrosa, F. Domingo-Perez, J. Velasco, C. Losada-Gutierrez, and J. Macias-Guara, “Optimal microphone placement for indoor acoustic localization using evolutionary optimization,” in Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN), Oct. 2016, pp. 1–8.

[5] K. Worden and A. P. Burrows, “Optimal sensor placement for fault detection,” Eng. Struct., vol. 23, no. 8, pp. 885–901, Aug. 2001.

[6] T.-H. Yi, H.-N. Li, and M. Gu, “Optimal sensor placement for structural health monitoring based on multiple optimal placement strategies,” Struct. Design Tall Special Buildings, vol. 20, no. 7, pp. 881–900, 2011.

[7] S. M. Hoseyni, F. Di Maio, and E. Zio, “Subset simulation for optimal sensors positioning based on value of information,” Proc. Inst. Mech. Eng., O, J. Risk Rel., vol. 174, pp. 18006012X11814342, Aug. 2022.

[8] A. C. Tapia and A. R. Messina, “Constrained sensor placement and state reconstruction in power systems from partial system observations,” Int. J. Electr. Power Energy Syst., vol. 146, Mar. 2023, Art. no. 108720.

[9] W. Du, Z. Xing, M. Li, B. He, L. H. C. Chua, and H. Miao, “Optimal sensor placement and measurement of wind for water quality studies in urban reservoirs,” in Proc. 15th Int. Symp. Inf. Process. Sensor Netw., 2014, pp. 167–178.

[10] M. M. Kelp, S. Lin, J. N. Kutz, and L. J. Mickley, “A new approach for determining optimal placement of PM$_2.5$ air quality sensors: Case study for the contiguous United States,” Environ. Res. Lett., vol. 17, no. 3, Mar. 2022, Art. no. 034034.

[11] W.-J. Yeo, S. Taulu, and J. N. Kutz, “Efficient magnetometer sensor array selection for signal reconstruction and brain source localization,” 2022, arXiv:2205.10025.

[12] W. J. Welch, “Branch-and-bound search for experimental designs based on D optimality and other criteria,” Technometrics, vol. 24, no. 1, pp. 41–48, 1982.

[13] E. L. Lawler and D. E. Wood, “Branch-and-bound methods: A survey,” Oper. Res., vol. 14, no. 4, pp. 699–719, Jul/Aug. 1966.

[14] S. Joshi and S. Boyd, “Sensor selection via convex optimization,” IEEE Trans. Signal Process., vol. 57, no. 2, pp. 451–462, Feb. 2009.

[15] F. Lin, M. Fardad, and M. R. Jovanovic, “Design of optimal sparse feedback gains via the alternating direction method of multipliers,” IEEE Trans. Autom. Control, vol. 58, no. 9, pp. 2426–2431, Sep. 2013.

[16] M. Fardad, F. Lin, and M. R. Jovanovic, “Sparsity-promoting optimal control for a class of distributed systems,” in Proc. Amer. Control Conf., Jun. 2011, pp. 2050–2055.

[17] N. K. Dhingra, M. R. Jovanovic, and Z.-Q. Luo, “An ADMM algorithm for optimal sensor and actuator selection,” in Proc. 53rd IEEE Conf. Decis. Control, Dec. 2014, pp. 4039–4044.

[18] A. Zare and M. R. Jovanovic, “Optimal sensor selection via proximal optimization algorithms,” in Proc. IEEE Conf. Decis. Control (CDC), Dec. 2018, pp. 6514–6518.

[19] T. Nagata, T. Nonomura, K. Nakai, K. Yamada, Y. Saito, and S. Ono, “Data-driven sparse sensor selection based on A-optimal design of experiment with ADMM,” IEEE Sensors J., vol. 21, no. 13, pp. 15248–15257, Jul. 2021.

[20] T. Nagata, K. Yamada, T. Nonomura, K. Nakai, Y. Saito, and S. Ono, “Data-driven sensor selection method based on proximal optimization for high-dimensional data with correlated measurement noise,” IEEE Trans. Signal Process., vol. 70, pp. 5251–5264, 2022.

[21] 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 Syst. Lett., vol. 3, no. 19, pp. 1–13, Nov. 2019.

[22] Y. Saito et al., “Determination-based fast greedy sensor selection algorithm,” IEEE Access, vol. 9, pp. 68535–68551, 2021.

[23] Y. Saito et al., “Data-driven vector-measurement-sensor selection based on greedy algorithm,” IEEE Sensors Lett., vol. 4, no. 7, pp. 1–4, Jul. 2020.

[24] K. Nakai, 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.

[25] E. Clark, T. Askham, S. L. Brunton, and J. N. Kutz, “Greedy sensor placement with cost constraints,” IEEE Sensors J., vol. 19, no. 7, pp. 2642–2656, Apr. 2019.
[47] N. Kanda, K. Nakai, Y. Saito, T. Nonomura, and K. Asai, “Feasibility study on real-time observation of flow velocity field using sparse processing particle image velocimetry,” Trans. Jpn. Soc. Aeronaut. Space Sci., vol. 64, no. 4, pp. 242–245, 2021.

[48] N. Kanda et al., “Proof-of-concept study of sparse processing particle image velocimetry for real time flow observation,” Experim. Fluids, vol. 63, no. 9, p. 143, Sep. 2022.

[49] S. Kaneko et al., “Data-driven sparse sampling for reconstruction of acoustic-wave characteristics used in aeroacoustic beamforming,” Appl. Sci., vol. 11, no. 5, p. 4216, May 2021.

[50] T. Goga et al., “Seismic wavefield reconstruction based on compressed sensing using data-driven reduced-order model,” Geophys. J. Int., vol. 233, no. 1, pp. 33–50, Nov. 2022.

[51] S. Nomura, J. Hara, and Y. Tanaka, “Dynamic sensor placement for linear inverse problems,” IEEE Trans. Signal Process., vol. 62, no. 5, pp. 1135–1146, Mar. 2014.

[52] C. Jiang, Y. C. Soh, and H. Li, “Sensor placement by maximal projection on minimum eigenspace for linear inverse problems,” IEEE Trans. Signal Process., vol. 64, no. 21, pp. 5595–5610, Nov. 2016.

[53] M. Shamaiah, S. Banerjee, and H. Vikalo, “Greedy sensor selection: Leveraging submodularity,” in Proc. 49th IEEE Conf. Decis. Control (CDC), Dec. 2010, pp. 2572–2577.

[54] C. Jiang, Z. Chen, R. Su, and Y. C. Soh, “Group greedy method for sensor placement,” IEEE Trans. Signal Process., vol. 67, no. 9, pp. 2249–2262, May 2019.

[55] C. Tillmann and H. Ney, “Word reordering and a dynamic programming beam search algorithm for statistical machine translation,” Comput. Linguistics, vol. 29, no. 1, pp. 97–133, Mar. 2003.

[56] R. Motwani and P. Raghavan, Randomized Algorithms. Cambridge, U.K.: Cambridge Univ. Press, 1995.

[57] N. Halko, P. G. Martinsson, and J. A. Tropp, “Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions,” SIAM Rev., vol. 53, no. 2, pp. 217–288, 2011.

[58] V. Cevher, S. Becker, and M. Schmidt, “Convex optimization for big data: Scalable, randomized, and parallel algorithms for big data analytics,” IEEE Trans. Signal Process., vol. 61, no. 5, pp. 32–43, Sep. 2014.

[59] Y. Huang and D. P. Palomar, “Randomized algorithms for optimal solutions of double-sided QCQP with applications in signal processing,” IEEE Trans. Signal Process., vol. 62, no. 5, pp. 1093–1108, Mar. 2014.

[60] S. Ono, “Efficient constrained signal reconstruction by randomized epigraphical projection,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), May 2019, pp. 4993–4997.

[61] T. Nonomura, S. Ono, K. Nakai, and Y. Saito, “Randomized subspace Newton convex method applied to data-driven sensor selection problem,” IEEE Signal Process. Lett., vol. 28, pp. 284–288, 2021.

[62] R. Gower, D. Koralev, F. Lieder, and P. Richtarik, “RSN: Randomized subspace Newton,” in Proc. Adv. Neural Inf. Process. Syst., H. Wallach, H. Larochelle, A. Beygelzimer, F. D. Alché-Buc, E. Fox, and R. Garnett, Eds. MIT Press, 2019, pp. 616–625. [Online]. Available: http://papers.nips.cc/paper/8351-rsn-randomized-subspace-newton.pdf

[63] A. Hashemi, M. Ghasemi, H. Vikalo, and U. Topcu, “Randomized greedy sensor selection: Leveraging weak submodularity,” IEEE Trans. Autom. Control, vol. 66, no. 1, pp. 199–212, Jan. 2021.

[64] B. Mirzasoleiman, A. Badanidiyuru, A. Karbasi, I. Vondrak, and A. Krause, “Lazier than lazy greedy,” in Proc. AAAI Conf. Artif. Intell., 2015, vol. 29, no. 1, pp. 1–7.

[65] T. Nagata. (2023). Elite-and-Randomized-Group-Greedy-Method. [Online]. Available: https://github.com/Aerodynamics-Lab/Elite-and-Randomized-Group-Greedy-Method

[66] V. Kekatos, G. B. Giannakis, and B. Wollenberg, “Optimal placement of phasor measurement units via convex relaxation,” IEEE Trans. Power Syst., vol. 27, no. 3, pp. 1521–1530, Aug. 2012.

[67] A. Atkinson, A. Donev, and R. Tobias, Optimum Experimental Designs, With SAS, vol. 34. Oxford, U.K.: Oxford Univ. Press, 2007.

[68] T. Nonomura, K. Nankai, Y. Iwasaki, A. Komuro, and K. Asai, “Quantitative evaluation of predictability of linear reduced-order model based on particle-image-velocimetry data of separated flow field around airfoil,” Experim. Fluids, vol. 62, no. 5, p. 112, May 2021.

[69] N. Onomura, K. Nankai, Y. Iwasaki, A. Komuro, and K. Asai. (2021). Airfoil PIV Data for Linear Rom. [Online]. Available: https://github.com/Aerodynamics-Lab/Airfoil-PIV-data-for-linear-ROM

[70] C. Eckart and G. Young, “The approximation of one matrix by another of lower rank,” Psychometrika, vol. 1, no. 3, pp. 211–218, 1936.

Takayuki Nagata received the B.S. and M.S. degrees in mechanical and aerospace engineering from Tokai University, Hiratsuka, Japan, in 2015 and 2017, respectively, and the Ph.D. degree in aerospace engineering from Tohoku University, Sendai, Japan, in 2020.

From 2018 to 2020, he was a Research Fellow with the Japan Society for the Promotion of Science (JSPS), Tohoku University. He is currently a Project Assistant Professor with Tohoku University.

Keigo Yamada received the B.S. degree in physics from Tohoku University, Sendai, Japan, in 2019, where he is currently pursuing the Ph.D. degree with the Department of Aerospace Engineering.

He is a Research Fellow with the Japan Society for the Promotion of Science (JSPS).

Kumi Nakai received the Ph.D. degree in mechanical systems engineering from the Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2020.

From 2017 to 2020, she was a Research Fellow with the Japan Society for the Promotion of Science (JSPS). Tokyo University of Agriculture and Technology. She is currently a Researcher with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan.

Yui Saito received the B.S. degree in mechanical engineering from Hokkaido University, Sapporo, Japan, in 2014, and the M.S. and Ph.D. degrees in mechanical space engineering from Hokkaido University, in 2016 and 2018, respectively.

From 2017 to 2018, he was a Research Fellow of the Japan Society for the Promotion of Science (JSPS) at Hokkaido University. He is currently an Assistant Professor at Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai, Japan.

Takuo Nonomura received the B.S. degree in mechanical and aerospace engineering from Nagoya University, Nagoya, Japan, in 2003, and the Ph.D. degree in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2008.

He is currently an Associate Professor with the Department of Aerospace Engineering, Tohoku University, Sendai, Japan.