

# Sensor placement in arbitrarily restricted region for field estimation based on Gaussian process

Tomoya Nishida\*, Natsuki Ueno\*, Shoichi Koyama\*<sup>†</sup> and Hiroshi Saruwatari\*

\*The University of Tokyo, Graduate School of Information Science and Technology,

7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

<sup>†</sup>JST, PRESTO, 4-1-6 Honcho, Kawaguchi, Saitama 332-0012, Japan

**Abstract**—A sensor placement method that enables us to arbitrarily set a region of candidate positions independently of a region of estimation is proposed. Field estimation aims to estimate and interpolate the physical quantities of fields, e.g., temperature and sound pressure, in an entire region of interest, where Gaussian processes are typically used for modeling. Although a number of sensor placement methods are proposed in the literature, in most of the methods, an optimization criterion is evaluated only at the candidate positions of the sensors. However, a region in which sensors are placed is sometimes restricted in practical applications of field estimation. To overcome this issue, we formulate a cost function on the basis of the expected squared error inside the target region for field estimation, which is derived by Gaussian process regression. We also propose two algorithms, the greedy algorithm and convex relaxation method, to efficiently solve this optimization problem. Numerical simulation results indicated that our proposed method achieves accurate field estimation even when the placement region of sensor candidates is restricted.

**Index Terms**—sensor placement, Gaussian process, field estimation, greedy algorithm, convex optimization

## I. INTRODUCTION

Estimation and interpolation of spatial phenomena with distributed multiple sensors have numerous applications in the analysis and control of various physical fields, such as temperature [1], [2] and acoustic fields [3], [4]. We refer to this type of estimation as *field estimation*. Gaussian processes [5] are among the typical approaches to modeling the physical quantity of a field inside an entire region of interest. Since it is necessary to estimate a field inside a large region in several applications, a large number of sensors are required for accurate field estimation. On the other hand, a small number of sensors are preferable owing to cost and space limitations. Therefore, optimizing sensor placement is also an important task in field estimation, which is referred to as the *sensor placement problem*.

A number of sensor placement and selection methods have been proposed in various contexts. Most of them solve a selection problem from finite candidate positions defined inside the region of interest. Efficient algorithms for solving this combinatorial problem have been derived on the basis of various criteria for linear inverse problems [6]–[12] and Gaussian process regression [13]–[17]. For the linear inverse problems, various criteria for measurement matrices [18], many of which are derived from experimental designs [19], are used. Information-theoretic measures, such as entropy [15] and mutual information [16], are proposed in Gaussian-process-

based sensor placement methods. The expected value of mean squared error (MSE), which is derived as the trace of an error covariance matrix as used in the linear inverse problems [7], [9], can also be employed as a measure in the context of Gaussian process regression [14], [17].

One important issue about current sensor placement methods is that the above-mentioned measures for sensor placement are evaluated only at candidate positions. This means that the sensor placement is selected so that accurate estimation at the candidate positions is obtained. However, particularly in the case of field estimation, the region where the sensors can be placed, the *candidate region*, is not necessarily the exact target region of field estimation, the *estimation region*, in practical applications. The candidate region can be restricted to a part of the estimation region or completely separated from the estimation region. For example, as shown in Fig. 1, sensors can only be placed on the wall whereas the field to be estimated is inside the room.

To overcome this limitation, we propose a Gaussian-process-based sensor placement method that enables us to arbitrarily and independently set the estimation and candidate regions. To the best of our knowledge, such a sensor placement problem has not been investigated, with a few exceptions applied to different problems in the context of wireless sensor networks [20]–[22]. In our proposed method for field estimation, the expected squared error obtained on the basis of Gaussian process regression is evaluated inside the estimation region whereas the sensors are selected from the arbitrarily set candidate region.

This paper is organized as follows. The sensor placement problem for arbitrarily set estimation and candidate regions is defined in Sect. II. In Sect. III, the proposed cost function is formulated on the basis of the expected squared error derived by Gaussian process regression. Two types of algorithm for approximately solving the cost function are proposed in Sect. IV. The experimental results are reported in Sect. V. Finally, Sect. VI concludes this paper.

## II. PROBLEM STATEMENT

We assume a Gaussian process for the probabilistic model of a field in the domain  $D$  [5]. In field estimation,  $D$  is typically set as  $\mathbb{R}^3$  or  $\mathbb{R}^2$ . The physical quantity of the field is denoted by  $x(\mathbf{r}) \in \mathbb{R}$  with the position vector  $\mathbf{r} \in D$ . In the Gaussian process model, for any set of discrete positions  $\mathcal{A} = \{\mathbf{r}_1, \dots, \mathbf{r}_J\} \subset D$ , the vector  $\mathbf{x}_{\mathcal{A}} = [x(\mathbf{r}_1), \dots, x(\mathbf{r}_J)]^T \in$

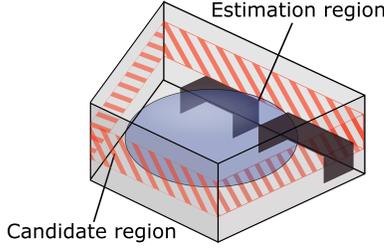


Fig. 1. Example of the case that the candidate region of sensor placement is separated from the estimation region. Sensors can only be placed on the wall whereas the field to be estimated is inside the room.

$\mathbb{R}^J$  is assumed to follow a multivariate Gaussian distribution as

$$\mathbf{x}_{\mathcal{A}} \sim \mathcal{N}(\mathbf{u}_{\mathcal{A}}, \mathbf{K}_{\mathcal{A},\mathcal{A}}), \quad (1)$$

where  $\mathbf{u}_{\mathcal{A}} = [u(\mathbf{r}_1), \dots, u(\mathbf{r}_J)]^T \in \mathbb{R}^J$  with the mean function  $u(\cdot)$ , and  $\mathbf{K}_{\mathcal{A},\mathcal{A}} \in \mathbb{R}^{J \times J}$  is the Gram matrix whose  $(i, j)$ th element is represented as  $[\mathbf{K}_{\mathcal{A},\mathcal{A}}]_{i,j} = \kappa(\mathbf{r}_i, \mathbf{r}_j)$  with the positive definite kernel function  $\kappa(\cdot, \cdot)$  and  $\mathbf{r}_i, \mathbf{r}_j \in \mathcal{A}$ . In general, the kernel function is learned from observed data [23], [24] or determined by *a priori* information on properties of the observations [25]. A typically used kernel function is the Gaussian kernel defined as [5]

$$\kappa(\mathbf{r}_i, \mathbf{r}_j) = \exp\left(-\frac{1}{2\gamma^2}\|\mathbf{r}_i - \mathbf{r}_j\|^2\right), \quad (2)$$

where  $\gamma > 0$  is a constant parameter to determine the spread of the function.

We consider estimating the field  $x(\mathbf{r})$  with  $K$  sensors at the set of positions  $\mathcal{S}$ . The discrete set of estimation positions is defined as  $\mathcal{E} = \{\mathbf{p}_1, \dots, \mathbf{p}_M\} \subset D$ , which is normally set by discretizing the region of interest inside  $D$ . The sensor locations  $\mathcal{S}$  are taken from the discrete set of candidate positions  $\mathcal{C} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\} \subset D$ , i.e.,  $\mathcal{S} \subset \mathcal{C}$ . Thus,  $|\mathcal{S}| = K$ ,  $|\mathcal{E}| = M$ , and  $|\mathcal{C}| = N > K$ . The sets of estimation and candidate positions,  $\mathcal{E}$  and  $\mathcal{C}$ , respectively, can be designed arbitrarily and independently inside  $D$ . The objective is to choose the optimal  $\mathcal{S}$  from  $\mathcal{C}$  so that the field on  $\mathcal{E}$  is accurately estimated. Note that  $\mathcal{E}$  and  $\mathcal{C}$  are set to be identical in prior works of the sensor placement problems, such as [14], [15].

Under the selected sensor locations  $\mathcal{S}$ , the sensor measurements  $\mathbf{y}_{\mathcal{S}} \in \mathbb{R}^K$  are represented as

$$\mathbf{y}_{\mathcal{S}} = \mathbf{x}_{\mathcal{S}} + \mathbf{v}_{\mathcal{S}}, \quad (3)$$

where  $\mathbf{v}_{\mathcal{S}}$  is the Gaussian noise of mean  $\mathbf{0}$  and covariance  $\Sigma_{\mathcal{S},\mathcal{S}}$ , i.e.,  $\mathbf{v}_{\mathcal{S}} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathcal{S},\mathcal{S}})$ . The covariance matrix  $\Sigma_{\mathcal{S},\mathcal{S}}$  is the submatrix of  $\Sigma_{\mathcal{C},\mathcal{C}}$ , i.e., the covariance matrix of the noise at the candidate positions  $\mathcal{C}$ .

### III. PROPOSED COST FUNCTION BASED ON EXPECTED SQUARED ERROR

Now, we formulate the cost function for the sensor placement problem on the basis of the expected squared error. The joint probability distribution of the sensor measurements at  $\mathcal{S}$  and the field at  $\mathcal{E}$  can be written as

$$\begin{bmatrix} \mathbf{x}_{\mathcal{E}} \\ \mathbf{y}_{\mathcal{S}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{u}_{\mathcal{E}} \\ \mathbf{u}_{\mathcal{S}} \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{\mathcal{E},\mathcal{E}} & \mathbf{K}_{\mathcal{E},\mathcal{S}} \\ \mathbf{K}_{\mathcal{S},\mathcal{E}} & \mathbf{K}_{\mathcal{S},\mathcal{S}} + \Sigma_{\mathcal{S},\mathcal{S}} \end{bmatrix}\right), \quad (4)$$

where  $\mathbf{K}_{\mathcal{S},\mathcal{E}} = \mathbf{K}_{\mathcal{E},\mathcal{S}}^T \in \mathbb{R}^{K \times M}$  are the cross-covariance matrices between  $\mathbf{x}_{\mathcal{S}}$  and  $\mathbf{x}_{\mathcal{E}}$ , calculated using the kernel function. The posterior distribution of  $\mathbf{x}_{\mathcal{E}}$  given  $\mathbf{y}_{\mathcal{S}}$  also becomes Gaussian as

$$p(\mathbf{x}_{\mathcal{E}}|\mathbf{y}_{\mathcal{S}}) = \mathcal{N}(\hat{\mathbf{x}}_{\mathcal{E}}, \Lambda_{\mathcal{E}}), \quad (5)$$

where the posterior mean  $\hat{\mathbf{x}}_{\mathcal{E}}$  and the covariance  $\Lambda_{\mathcal{E}}$  are respectively obtained as

$$\hat{\mathbf{x}}_{\mathcal{E}} = \mathbf{u}_{\mathcal{E}} + \mathbf{K}_{\mathcal{E},\mathcal{S}}(\mathbf{K}_{\mathcal{S},\mathcal{S}} + \Sigma_{\mathcal{S},\mathcal{S}})^{-1}(\mathbf{y}_{\mathcal{S}} - \mathbf{u}_{\mathcal{S}}), \quad (6)$$

$$\Lambda_{\mathcal{E}} = \mathbf{K}_{\mathcal{E},\mathcal{E}} - \mathbf{K}_{\mathcal{E},\mathcal{S}}(\mathbf{K}_{\mathcal{S},\mathcal{S}} + \Sigma_{\mathcal{S},\mathcal{S}})^{-1}\mathbf{K}_{\mathcal{S},\mathcal{E}}. \quad (7)$$

This means that the MAP estimate of the field at  $\mathcal{E}$  is  $\hat{\mathbf{x}}_{\mathcal{E}}$ . It is worth noting that (6) is equivalent to the kernel ridge regression [26], which is often used for field interpolation.

The expected value of  $\|\hat{\mathbf{x}}_{\mathcal{E}} - \mathbf{x}_{\mathcal{E}}\|^2$  is given by the trace of the posterior covariance matrix and reformulated by the matrix inversion lemma as

$$\begin{aligned} E[\|\hat{\mathbf{x}}_{\mathcal{E}} - \mathbf{x}_{\mathcal{E}}\|^2] &= \text{tr}\left(\mathbf{K}_{\mathcal{E},\mathcal{E}} - \mathbf{K}_{\mathcal{E},\mathcal{S}}(\mathbf{K}_{\mathcal{S},\mathcal{S}} + \Sigma_{\mathcal{S},\mathcal{S}})^{-1}\mathbf{K}_{\mathcal{S},\mathcal{E}}\right) \\ &= \text{tr}\left(\mathbf{K}_{\mathcal{E},\mathcal{E}} - \mathbf{K}_{\mathcal{E},\mathcal{C}}\mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1}\mathbf{K}_{\mathcal{C},\mathcal{E}} + \right. \\ &\quad \left. \mathbf{K}_{\mathcal{C},\mathcal{E}}\mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1}\left(\mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} + \Phi_{\mathcal{S}}^T\Sigma_{\mathcal{S},\mathcal{S}}^{-1}\Phi_{\mathcal{S}}\right)^{-1}\mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1}\mathbf{K}_{\mathcal{C},\mathcal{E}}\right). \end{aligned} \quad (8)$$

Here,  $\Phi_{\mathcal{S}} \in \{0,1\}^{K \times N}$  is defined as the binary-valued matrix that extracts the rows of the sensor indexes of  $\mathcal{S}$  by multiplication. For instance, the rows  $\mathcal{S}$  are extracted from the rows  $\mathcal{C}$  of  $\mathbf{K}_{\mathcal{C},\mathcal{E}}$  as  $\mathbf{K}_{\mathcal{S},\mathcal{E}} = \Phi_{\mathcal{S}}\mathbf{K}_{\mathcal{C},\mathcal{E}}$ . Each row of  $\Phi_{\mathcal{S}}$  is a unit vector with values of 1 for the sensor index and 0 for the other indexes.

Finally, the optimization problem for finding  $K$  sensor locations from the candidates  $\mathcal{C}$  that minimize the expected squared error at  $\mathcal{E}$  is formulated as

$$\begin{aligned} \underset{\mathcal{S} \subset \mathcal{C}}{\text{minimize}} \quad & J(\mathcal{S}) := \text{tr}\left(\left(\mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} + \Phi_{\mathcal{S}}^T\Sigma_{\mathcal{S},\mathcal{S}}^{-1}\Phi_{\mathcal{S}}\right)^{-1}\bar{\mathbf{K}}\right), \\ \text{subject to} \quad & |\mathcal{S}| = K, \end{aligned} \quad (9)$$

where  $\bar{\mathbf{K}} = \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1}\mathbf{K}_{\mathcal{C},\mathcal{E}}\mathbf{K}_{\mathcal{E},\mathcal{C}}\mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1}$ . The terms irrelevant to the optimization with respect to  $\mathcal{S}$  are omitted. Since this optimization problem is a combinatorial problem and  $\mathcal{C}$  is generally obtained by finely discretizing a part of the domain  $D$  in the field estimation, exhaustive search is impractical. We derive two algorithms to efficiently solve (9) in the next section.

We briefly discuss the relationship between the cost function obtained by our proposed method and that obtained in prior works. In the current method based on MSE [14], the estimation and candidate positions must be identical, i.e.,  $\mathcal{E} = \mathcal{C}$ . In the case of  $\mathcal{E} = \mathcal{C}$ , the proposed cost function (9) corresponds to that in [14]; therefore, the proposed method is a natural extension of the MSE-based method to the case that  $\mathcal{E}$  and  $\mathcal{C}$  are not necessarily identical. In addition, the proposed method for  $\mathcal{E} = \mathcal{C}$  corresponds to the MSE-based sensor placement method for finite dimensional linear inverse problems [7] by setting the measurement matrix of the linear measurement model as an identity matrix.

---

**Algorithm 1** Greedy algorithm for solving (9)

---

**Require:**  $\mathcal{S} = \emptyset$  and  $\mathbf{A}_{\mathcal{S}} = \mathbf{K}_{\mathcal{C},\mathcal{C}}$

- 1: **for**  $k = 1, 2, \dots, K$  **do**
  - 2: Find  $\mathbf{q}_j \in \mathcal{C} \setminus \mathcal{S}$  that maximize  $J(\mathcal{S}) - J(\mathcal{S} \cup \{\mathbf{q}_j\})$  in (11)
  - 3: Update sensor selection  $\mathcal{S} \leftarrow \mathcal{S} \cup \{\mathbf{q}_j\}$
  - 4: Update  $\mathbf{A}_{\mathcal{S}}$  by (10), and then  $\mathbf{A}_{\mathcal{S}} \bar{\mathbf{K}} \mathbf{A}_{\mathcal{S}}$
  - 5: **end for**
- 

#### IV. ALGORITHMS FOR PROPOSED COST FUNCTION

We derive two computationally efficient algorithms to obtain an approximate solution for the optimization problem (9): greedy algorithm and convex relaxation method. These algorithms are basically derived as extensions of the algorithms proposed in [7].

##### A. Greedy algorithm

In the greedy algorithm, the sensor location that minimizes the cost function is sequentially selected one by one. We here propose a method to efficiently calculate the cost function at each iteration.

At each iteration of the greedy algorithm, the decrease in the cost function by adding an additional sensor at  $\mathbf{q}_j \in \mathcal{C} \setminus \mathcal{S}$  to the current selection  $\mathcal{S}$ , i.e.,  $J(\mathcal{S}) - J(\mathcal{S} \cup \{\mathbf{q}_j\})$ , is evaluated. We here describe the cost function in (9) as  $J(\mathcal{S}) = \text{tr}(\mathbf{A}_{\mathcal{S}} \bar{\mathbf{K}})$ , where  $\mathbf{A}_{\mathcal{S}}$  is defined as  $\mathbf{A}_{\mathcal{S}} := (\mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} + \Phi_{\mathcal{S}}^T \Sigma_{\mathcal{S},\mathcal{S}}^{-1} \Phi_{\mathcal{S}})^{-1}$ . Then,  $\mathbf{A}_{\mathcal{S} \cup \{\mathbf{q}_j\}}$  is obtained using  $\mathbf{A}_{\mathcal{S}}$  as

$$\mathbf{A}_{\mathcal{S} \cup \{\mathbf{q}_j\}} = \mathbf{A}_{\mathcal{S}} - \frac{t_j^{-1} \mathbf{A}_{\mathcal{S}} \gamma_j \gamma_j^T \mathbf{A}_{\mathcal{S}}}{1 + t_j^{-1} \gamma_j^T \mathbf{A}_{\mathcal{S}} \gamma_j}, \quad (10)$$

where  $\gamma_j = \mathbf{e}_j - \Phi_{\mathcal{S}} \Sigma_{\mathcal{S},\mathcal{S}}^{-1} \Sigma_{\mathcal{S},j}$  and  $t_j = [\Sigma_{\mathcal{C},\mathcal{C}}]_{j,j} - \Sigma_{j,\mathcal{S}} \Sigma_{\mathcal{S},\mathcal{S}}^{-1} \Sigma_{\mathcal{S},j}$ . Here,  $\mathbf{e}_j$  is the unit vector whose  $j$ th element is 1 and the others are 0, and  $\Sigma_{\mathcal{S},j} = \Sigma_{j,\mathcal{S}}^T$  is the covariance vector between the noise at the selected sensor positions  $\mathcal{S}$  and that of the  $j$ th candidate position, extracted from  $\Sigma_{\mathcal{C},\mathcal{C}}$ . Thus, the decrease in the cost function is obtained as

$$J(\mathcal{S}) - J(\mathcal{S} \cup \{\mathbf{q}_j\}) = \frac{\gamma_j^T \mathbf{A}_{\mathcal{S}} \bar{\mathbf{K}} \mathbf{A}_{\mathcal{S}} \gamma_j}{t_j + \gamma_j^T \mathbf{A}_{\mathcal{S}} \gamma_j}. \quad (11)$$

The matrix multiplication of  $\mathbf{A}_{\mathcal{S}} \bar{\mathbf{K}} \mathbf{A}_{\mathcal{S}} \in \mathbb{R}^{N \times N}$  can be carried out before computing  $J(\mathcal{S}) - J(\mathcal{S} \cup \{\mathbf{q}_j\})$  for each  $\mathbf{q}_j \in \mathcal{C} \setminus \mathcal{S}$ . Although the direct computation of the cost function requires the computational cost of  $O(N^3)$  for each  $\mathbf{q}_j \in \mathcal{C} \setminus \mathcal{S}$ , it can be reduced to  $O(N^2)$  by using (11), which is equivalent to the computational cost for the greedy sensor selection algorithm for the finite dimensional linear inverse problem [7]. The greedy algorithm is summarized in Algorithm 1.

##### B. Convex relaxation

The convex relaxation of the optimization problem (9) is derived by two steps. First, we reformulate the cost function with a vector variable instead of  $\Phi_{\mathcal{S}}$ . Second, a semidefinite programming (SDP) is derived by its convex relaxation.

According to the procedure described in [7], the noise covariance matrix  $\Sigma_{\mathcal{C},\mathcal{C}}$  ( $\succ 0$ ) can be decomposed as

$$\Sigma_{\mathcal{C},\mathcal{C}} = \beta \mathbf{I} + \mathbf{P}, \quad (12)$$

where  $0 < \beta < \sigma_{\min}(\Sigma_{\mathcal{C},\mathcal{C}})$  with  $\sigma_{\min}(\Sigma_{\mathcal{C},\mathcal{C}})$  being the smallest singular value of  $\Sigma_{\mathcal{C},\mathcal{C}}$ ,  $\mathbf{P} \succ 0$ , and  $\mathbf{I}$  is the identity matrix. The matrix  $\mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} + \Phi_{\mathcal{S}}^T \Sigma_{\mathcal{S},\mathcal{S}}^{-1} \Phi_{\mathcal{S}}$  that appears in  $J(\mathcal{S})$  is reformulated as

$$\begin{aligned} & \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} + \Phi_{\mathcal{S}}^T \Sigma_{\mathcal{S},\mathcal{S}}^{-1} \Phi_{\mathcal{S}} \\ &= \mathbf{C} - \mathbf{P}^{-1} (\mathbf{P}^{-1} + \beta^{-1} \text{diag}(\mathbf{w}))^{-1} \mathbf{P}^{-1}, \end{aligned} \quad (13)$$

where  $\mathbf{C}$  and  $\mathbf{w}$  are defined as  $\mathbf{C} := \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} + \mathbf{P}^{-1}$  and  $\mathbf{w} \in \{0, 1\}^N$ . The vector variable  $\mathbf{w}$  is the binary-valued vector, whose elements of the indexes corresponding to  $\mathcal{S}$  are 1 and the others are 0, to represent the sensor selection instead of  $\Phi_{\mathcal{S}}$ . The cost function  $J(\mathcal{S})$  is rewritten as

$$\begin{aligned} J(\mathcal{S}) &= \text{tr} \left( \left( \mathbf{C} - \mathbf{P}^{-1} (\mathbf{P}^{-1} + \beta^{-1} \text{diag}(\mathbf{w}))^{-1} \mathbf{P}^{-1} \right)^{-1} \bar{\mathbf{K}} \right) \\ &:= J(\mathbf{w}). \end{aligned} \quad (14)$$

Thus, the cost function  $J(\mathbf{w})$  is minimized with respect to  $\mathbf{w}$ .

Next, we derive an SDP for minimizing  $J(\mathbf{w})$  with convex relaxation similarly to the method proposed in [7]. Two matrix variables  $\mathbf{V} \in \mathbb{R}^{N \times N}$  and  $\mathbf{Z} \in \mathbb{R}^{M \times M}$  are respectively defined as

$$\mathbf{V} = \mathbf{P}^{-1} (\mathbf{P}^{-1} + \beta^{-1} \text{diag}(\mathbf{w}))^{-1} \mathbf{P}^{-1}, \quad (15)$$

$$\mathbf{Z} = \mathbf{K}_{\mathcal{E},\mathcal{C}} \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} (\mathbf{C} - \mathbf{V})^{-1} \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} \mathbf{K}_{\mathcal{C},\mathcal{E}}. \quad (16)$$

Thus, the original optimization problem is equivalent to minimizing  $\text{tr}(\mathbf{Z}) = J(\mathbf{w})$  for  $\mathbf{Z}$ ,  $\mathbf{V}$ , and  $\mathbf{w}$  with the constraints of (15) and (16). To derive an SDP, semidefinite relaxation is applied to the equality constraints (15) and (16) by replacing “=” with “ $\succeq$ ”. Moreover, the boolean constraint  $\mathbf{w} \in \{0, 1\}^N$  is relaxed to the real-valued constraint  $\mathbf{w} \in [0, 1]^N$ .

Finally, the SDP is derived as

$$\underset{\mathbf{w}, \mathbf{Z}, \mathbf{V}}{\text{minimize}} \quad \text{tr}(\mathbf{Z}), \quad (17)$$

$$\text{subject to} \quad 0 \leq w_n \leq 1 \quad (n = 1, \dots, N), \quad (18)$$

$$\mathbf{1}^T \mathbf{w} = K, \quad (19)$$

$$\begin{bmatrix} \mathbf{V} & \mathbf{P}^{-1} \\ \mathbf{P}^{-1} & \mathbf{P}^{-1} + \beta^{-1} \text{diag}(\mathbf{w}) \end{bmatrix} \succeq 0, \quad (20)$$

$$\begin{bmatrix} \mathbf{Z} & \mathbf{K}_{\mathcal{E},\mathcal{C}} \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} \\ \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} \mathbf{K}_{\mathcal{C},\mathcal{E}} & \mathbf{C} - \mathbf{V} \end{bmatrix} \succeq 0, \quad (21)$$

where  $\mathbf{1}$  is the vector of all ones and (18) indicates that each element of  $\mathbf{w}$  is within the range between 0 and 1. Here, it can be shown that (20) and (21) are equivalent to the relaxed constraint of (15) and (16) by using the Schur complement of block matrices. Note that  $\text{tr}(\mathbf{Z}) \geq J(\mathcal{S})$ ; therefore, this SDP is an optimization problem relaxed from (9). After solving the SDP, the sensor placement can be obtained by choosing the  $K$  largest elements of  $\mathbf{w}$  or by performing a randomized rounding algorithm [27].

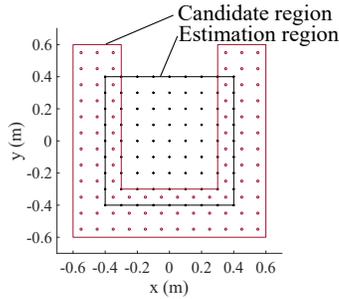


Fig. 2. Evaluation locations and candidate sensor locations for numerical experiment. The red and black dots represent  $\mathcal{C}$  and  $\mathcal{E}$ , respectively.

In a particular case in which the noise is uncorrelated and has the same covariance, i.e.,  $\Sigma = \lambda \mathbf{I}$ , the SDP can be significantly simplified. The cost function  $J(\mathbf{w})$  is rewritten as

$$J(\mathbf{w}) = \text{tr} \left( \mathbf{K}_{\mathcal{E},\mathcal{C}} \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} \left( \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} + \lambda^{-1} \text{diag}(\mathbf{w}) \right)^{-1} \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} \mathbf{K}_{\mathcal{C},\mathcal{E}} \right) \quad (22)$$

By using one additional variable  $\mathbf{Z} \in \mathbb{R}^{M \times M}$ , we can formulate the SDP as

$$\underset{\mathbf{w}, \mathbf{Z}}{\text{minimize}} \quad \text{tr}(\mathbf{Z}), \quad (23)$$

$$\text{subject to} \quad 0 \leq w_n \leq 1 \quad (n = 1, \dots, N), \quad (24)$$

$$\mathbf{1}^T \mathbf{w} = K, \quad (25)$$

$$\begin{bmatrix} \mathbf{Z} & \mathbf{K}_{\mathcal{E},\mathcal{C}} \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} \\ \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} \mathbf{K}_{\mathcal{C},\mathcal{E}} & \mathbf{K}_{\mathcal{C},\mathcal{C}}^{-1} + \lambda^{-1} \text{diag}(\mathbf{w}) \end{bmatrix} \succeq 0. \quad (26)$$

## V. EXPERIMENTS

We evaluate the proposed sensor placement method by numerical experiments. To investigate the effects of enabling arbitrarily setting estimation and candidate positions, the proposed algorithms are compared with the algorithms that employ the candidate positions as the estimation positions, in which both the candidate and estimation positions are fixed to the candidate positions of the proposed algorithms. The proposed algorithms and the algorithms for comparison are denoted as  $\mathcal{E} \neq \mathcal{C}$  and  $\mathcal{E} = \mathcal{C}$ , respectively. Note that the  $\mathcal{E} = \mathcal{C}$  case only considers the MSE at the candidate positions and is closely related to those proposed in [7]. In the two-dimensional field estimation ( $D = \mathbb{R}^2$ ), as shown in Fig. 2, the sensor positions were determined within the candidate region using each algorithm, and the estimation performance was evaluated in the estimation region. In the proposed algorithms, the sets of estimation positions  $\mathcal{E}$  and candidate positions  $\mathcal{C}$  are obtained by discretizing the estimation and candidate regions. Their intervals were 0.10 m for  $\mathcal{E}$  and 0.08 m for  $\mathcal{C}$ , and the numbers of positions were  $N = |\mathcal{E}| = 81$  and  $M = |\mathcal{C}| = 90$ . In the algorithms compared, both the estimation and candidate positions were set to the positions of  $\mathcal{C}$  used in the proposed algorithms.

The kernel function  $\kappa(\cdot, \cdot)$  was set as the Gaussian kernel in (2) with  $\gamma = 0.3$  m, and the noise covariance matrix was set as  $\Sigma_{\mathcal{C},\mathcal{C}} = \lambda \mathbf{I}$  with  $\lambda = 10^{-2}$ . Both the greedy

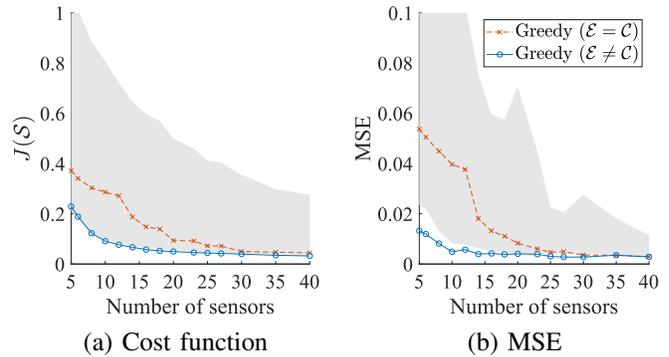


Fig. 3. Value of cost function and MSE with respect to number of sensors for greedy algorithms. Best and worst performances of random sampling are indicated by the gray area.

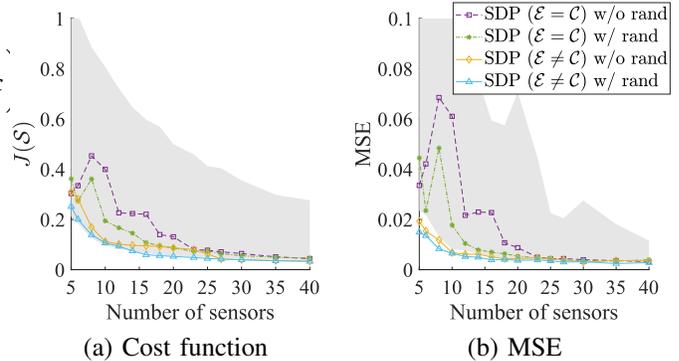


Fig. 4. Value of cost function and MSE with respect to number of sensors for SDP-based algorithms. Results of SDP with and without randomized rounding algorithm are shown.

algorithm in Algorithm 1 and the convex relaxation in (23)–(26) were evaluated. The SDP was solved using CVX [28]. After solving the SDP, we also performed the randomized rounding algorithm with  $10^3$  iterations for refinement.

In Figs. 3 and 4, the values of the proposed cost function  $J(S)$  and the MSEs in the estimation region for 50 patterns of the field are plotted with respect to the number of sensors  $K$  for greedy algorithms and SDP-based algorithms, respectively. In both methods,  $J(S)$  was calculated using  $\mathcal{E}$  and  $\mathcal{C}$  adopted in the proposed algorithms to evaluate the expected squared error in the estimation region. The field to be estimated was drawn by mixing two Gaussian distributions. The peak positions and values were randomly chosen from the rectangular region of  $1.2 \text{ m} \times 1.2 \text{ m}$  and the range of  $[-1, 1]$ , respectively. The MSE was evaluated at the positions obtained by resampling the estimation region at intervals of 0.01 m. The results of the greedy algorithms and the SDP-based algorithms are shown in Figs. 3 and 4, respectively. The best and worst cases obtained by random sensor selection  $10^7$  times are also plotted with the gray area indicating the approximate performance bound. In the greedy algorithms (Fig. 3), it can be observed that the proposed method for  $\mathcal{E} \neq \mathcal{C}$  outperformed the method for  $\mathcal{E} = \mathcal{C}$ . Moreover, the proposed method attained almost the best performance in the random selection for the value of the cost function, and even a lower MSE at a small number of selected sensors  $K$ . The performance of the SDP-

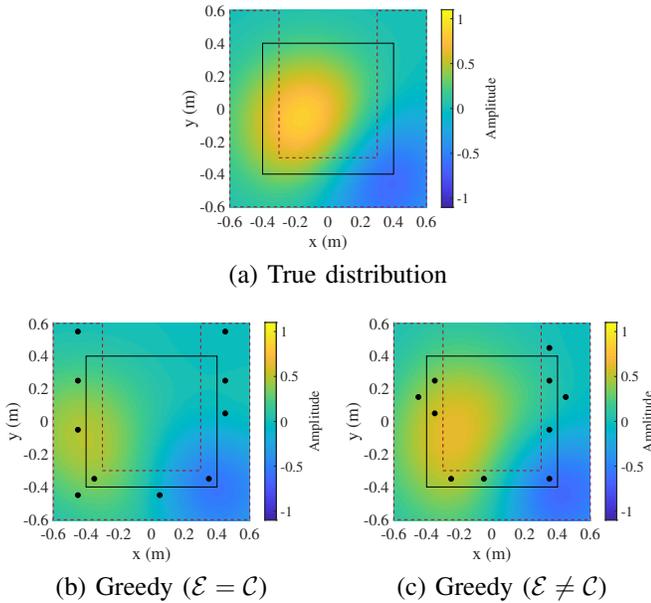


Fig. 5. Distributions of true and estimated fields for  $K = 10$  of greedy algorithm. Black dots represent selected sensors.

based algorithms depended on whether the random rounding algorithm was used or not; however, in most cases for  $K$ , the proposed method for  $\mathcal{E} \neq \mathcal{C}$  achieved a smaller cost function and a smaller MSE than the method for  $\mathcal{E} = \mathcal{C}$ . Fig. 5 shows the distributions of the true and estimated fields for  $K = 10$  of the greedy algorithms when two Gaussian distributions were placed at  $(-0.1 \text{ m}, -0.1 \text{ m})$  and  $(0.3 \text{ m}, -0.4 \text{ m})$  with their peak values of 1.0 and  $-0.8$ , respectively. It can be seen that accurate estimation was achieved using the proposed method (Fig. 5(c)) compared with the method for  $\mathcal{E} = \mathcal{C}$ , particularly inside the estimation region.

## VI. CONCLUSION

We proposed a sensor placement method that enables us to arbitrarily and independently set the estimation and candidate regions for the field estimation. The cost function is formulated as the expected squared error inside the estimation region on the basis of Gaussian process regression. The greedy algorithm and convex relaxation method are proposed to efficiently solve the optimization problem. In numerical simulations, accurate field estimation is achieved by the proposed method compared with the current method of evaluating optimization criteria only at the candidate positions.

## VII. ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number JP19H01116 and JST PRESTO Grant Number JPMJPR18J4.

## REFERENCES

- [1] M. R. Holdaway, "Spatial modeling and interpolation of monthly temperature using kriging," *Clim. Res.*, vol. 6, no. 3, pp. 215–225, 1996.
- [2] R. Fablet and F. Rousseau, "Joint interpolation of multisensor sea surface temperature fields using nonlocal and statistical priors," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2665–2675, 2016.

- [3] N. Ueno, S. Koyama, and H. Saruwatari, "Sound field recording using distributed microphones based on harmonic analysis of infinite order," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 135–139, Jan. 2018.
- [4] A. Laborie, R. Bruno, and S. Montoya, "A new comprehensive approach of surround sound recording," in *Proc. 114th AES Int. Conf.*, Oct. 2003.
- [5] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, Cambridge, 2006.
- [6] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, 2009.
- [7] S. Liu, S. P. Chepuri, M. Fardad, E. Maşazade, G. Leus, and P. K. Varshney, "Sensor selection for estimation with correlated measurement noise," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3509–3522, 2016.
- [8] J. Ranieri, A. Chebira, and M. Vetterli, "Near-optimal sensor placement for linear inverse problems," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1135–1146, Mar. 2014.
- [9] 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, 2019.
- [10] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Proc. IEEE Conf. Decis. Control (CDC)*, Dec. 2010, pp. 2572–2577.
- [11] M. Naeem, S. Xue, and D. C. Lee, "Cross-entropy optimization for sensor selection problems," in *Proc. Int. Symp. Commun. Inf. Technol. (ISCIT)*, Sep. 2009, pp. 396–401.
- [12] S. Koyama, G. Chardon, and L. Daudet, "Joint source and sensor placement for sound field control based on empirical interpolation method," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Apr. 2018, pp. 501–505.
- [13] D. J. C. MacKay, "Information-based objective functions for active data selection," *Neural Comput.*, vol. 4, no. 4, pp. 590–604, 1992.
- [14] Y. Yang and R. S. Blum, "Sensor placement in gaussian random field via discrete simulation optimization," *IEEE Signal Process. Lett.*, vol. 15, pp. 729–732, 2008.
- [15] C. Ko, J. Lee, and M. Queyranne, "An exact algorithm for maximum entropy sampling," *Oper. Res.*, vol. 43, no. 4, pp. 684–691, 1995.
- [16] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *J. Mach. Learn. Res.*, vol. 9, pp. 235–284, 2008.
- [17] M. Coutino, S. P. Chepuri, and G. Leus, "Subset selection for kernel-based signal reconstruction," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Apr. 2018, pp. 4014–4018.
- [18] N. K. Nguyen and A. J. Miller, "A review of some exchange algorithms for constructing discrete d-optimal designs," *Comput. Stat. Data Anal.*, vol. 14, no. 4, pp. 489–498, 1992.
- [19] D. M. Steinberg and W. G. Hunter, "Experimental design: review and comment," *Technometrics*, vol. 26, no. 2, pp. 71–97, 1984.
- [20] S. Liu, E. Masazade, M. Fardad, and P. K. Varshney, "Sparsity-aware field estimation via ordinary kriging," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2014, pp. 3948–3952.
- [21] S. Liu, A. Vempaty, M. Fardad, E. Masazade, and P. K. Varshney, "Energy-aware sensor selection in field reconstruction," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1476–1480, 2014.
- [22] S. Liu, S. P. Chepuri, G. Leus, and A. O. Hero, "Distributed sensor selection for field estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2017, pp. 4257–4261.
- [23] G. R. G. Lanckriet, N. Cristianini, P. Bartlett, L. E. Ghaoui, and M. I. Jordan, "Learning the kernel matrix with semidefinite programming," *J. Mach. Learn. Res.*, vol. 5, pp. 27–72, 2004.
- [24] A. G. Wilson and R. P. Adams, "Gaussian process kernels for pattern discovery and extrapolation," in *Proc. Int. Conf. Mach. Learn. (ICML)*, June 2013, pp. 1067–1075.
- [25] N. Ueno, S. Koyama, and H. Saruwatari, "Kernel ridge regression with constraint of helmholtz equation for sound field interpolation," in *Proc. Int. Workshop Acoust. Signal Enhancement (IWAENC)*, Sep. 2018, pp. 1–440.
- [26] K. P. Murphy, *Machine Learning: a Probabilistic Perspective*, MIT Press, Cambridge, 2012.
- [27] S. P. Chepuri and G. Leus, "Sparse sensing for distributed detection," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1446–1460, Mar. 2016.
- [28] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Mar. 2014.