CORRELATION-AWARE SPARSITY-ENFORCING SENSOR PLACEMENT FOR SPATIO-TEMPORAL FIELD ESTIMATION

Venkat Roy, Geert Leus

Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology, Delft, The Netherlands

ABSTRACT

In this work, we propose a generalized framework for designing optimal sensor constellations for spatio-temporally correlated field estimation using wireless sensor networks. The accuracy of the field intensity estimate in every point of a given service area strongly depends upon the number and the constellation of the sensors along with the spatio-temporal statistics of the field. We formulate and solve a sparsity-enforcing optimization problem to select the best sensor locations that achieve some desired estimation performance. The sparsity-enforcing iterative selection algorithm is aware of the non-separable space-time covariance structure of the field.

Index Terms— Wireless sensor network, field estimation, Bayesian framework, convex optimization, sparsity.

1. INTRODUCTION

Restricted use of the sensors in any sensor network over a specified service area due to life-time, bandwidth and other resource-related constraints is a wonted problem in environmental field estimation applications. To alleviate this problem, sparsity can be enforced in selecting only a subset of *informative* sensors which is well-investigated in [1] and references therein. Sparsity can also be introduced in sensor selection for the conventional kriging approach [2] by minimizing the kriging error variance while penalizing the kriging weights [3]. An elegant optimization framework has been proposed in [4] that handles the aforementioned performance-constrained and cardinality-aware optimization problems. A similar optimization framework can be utilized to handle a generalized non-linear measurement model [5], and can be successfully applied for different applications like anchor placement for localization [6], sensor selection for direction-of-arrival (DOA) estimation [7], etc. Also a Bayesian extension of the optimization framework is illustrated in [4], where a Bayesian error metric is minimized, constrained by the number of sensors to be selected. In [8], an information-theoretic approach is utilized for near-optimal sensor placements for Gaussian processes, where the mutual information of the selected and not-selected sensor locations is minimized exploiting the submodularity of the mutual information. Another significant contribution for selecting the optimal sensor locations is presented in [9], where the cost function is related to the frame potential property of the measurement matrix.

In this paper, we propose an *off-line sparsity-enforcing* sensor constellation design methodology for stationary field estimation, in a Bayesian framework. It is assumed that the first and second order statistics of the field intensity to be estimated are known a priori. Leveraging the space-time wide sense stationarity of the physical field, we select the best sensor locations constrained by an estimation performance metric which depends on the spatial and temporal lags. The developed approach can be generalized for any spatial resolution in different temporal sensing windows without any knowledge about the dynamics of the field. We use a space-time non-separable stationary covariance model which can readily be used for fitting spatio-temporal environmental and climatological fields [10]. The performance constraint is a cost function which is related to the mean squared error (MSE) covariance matrix. To be more specific, if an estimate of the field intensity in every point of the service area is given by the vector $\hat{\mathbf{u}}$, then what is the optimal way to deploy the sensors so that some desired $\mathrm{tr}(\mathbb{E}[(\mathbf{u}-\hat{\mathbf{u}})(\mathbf{u}-\hat{\mathbf{u}})^T])$ is achieved, where $\mathbf{u}$ is the true value of the field intensity vector. In terms of experimental design, the aforementioned problem can also be viewed as an *A-optimal* design that minimizes the trace of the MSE covariance matrix, i.e., the sum of the eigen values of $\mathbb{E}[(\mathbf{u}-\hat{\mathbf{u}})(\mathbf{u}-\hat{\mathbf{u}})^T]$ [11, Pg. 384].

The above problem is solved in two different space-time sensing/observation modalities which are,

- *Sliding measurement window*: The measurement window contains measurements from the past and the future as well as from the present snapshot, where the field is to be estimated. The next window shifts by one snapshot.
- *Non-overlapping measurement window*: The measurement window contains measurements at all snapshots where the field is to be estimated. The next window shifts by the total number of snapshots in the window.

The optimal sensor constellations for the above two scenarios are designed assuming all the field locations are candidate sensor locations. The effect of both the spatial as well as the temporal covariances are analyzed for both these scenarios. The developed sensor selection framework is based on the assumption that the physical field to be estimated is a Gaussian process. This is a realistic assumption, as the realizations of many physical stochastic processes like a Wiener process (used for modeling Brownian motion) or air pollution models (specially the ozone distribution) exhibit Gaussian behavior. Along with this, as mentioned in [12], stationarity is also a valid assumption for some environmental fields like rainfall, where off-line sensor (e.g., rain-gauge) deployment based on observations in different temporal regimes, is a standard practice.

Notations: Matrices are in upper case bold while column vectors are in lower case bold. $[\mathbf{X}]_{ij}$ is the (i, j)-th entry of the matrix $\mathbf{X}$. $[\mathbf{x}]_i$ is the i-th entry of the vector $\mathbf{x}$. $\mathrm{tr}[\mathbf{X}]$ denotes the trace of $\mathbf{X}$, i.e., the sum of the diagonal elements of $\mathbf{X}$. $\mathbf{I}_N$ is the identity matrix of size $N \times N$. $(\cdot)^T$ is the transpose operator, $\hat{\mathbf{x}}$ is the estimate of $\mathbf{x}$, $\triangleq$ defines an entity, $\|\mathbf{x}\|_p = (\sum_{i=0}^{N-1} |[\mathbf{x}]_i|^p)^{1/p}$ is the ℓ_p norm

This work has been supported by the SHINE project, the flagship project of DIRECT (Delft Institute for Research on ICT at the Delft University of Technology)

of $\mathbf{x}$. $\mathbf{0}_N$ and $\mathbf{1}_N$ are the vectors of all zeros and ones of length N, respectively. $\mathbf{X} \in \mathbb{S}^N$ denotes that $\mathbf{X}$ is in the set of symmetric matrices of size $N \times N$.

2. SYSTEM MODEL

2.1. Data model

In this section, we assume a finite uniform discretization of the entire service area into N spatial points. It is assumed that the field intensity within any pixel is characterized by the intensity at the centre of the pixel. The field intensities in all the N pixels describe the spatial field distribution over the area at any time instance t, denoted by $\mathbf{u}(t) \in \mathbb{R}^N$. All of these N locations are potential sensor locations. A deployed sensor, in any of these pixels monitors the field intensity within that particular pixel, i.e, only a single dimension of the field intensity vector $\mathbf{u}(t)$. Based on this, a *compressive* measurement model (less measurements than unknowns) can be constructed as

$$\mathbf{y}(t) = \mathbf{C}_t \mathbf{u}(t) + \mathbf{e}(t), \quad t = 1, 2 \dots \tag{1}$$

where $\mathbf{y}(t) \in \mathbb{R}^M$ collects the set of M measurements out of N locations ($M \leq N$) at any snapshot t. The *compressive measurement matrix* $\mathbf{C}_t \in \{0,1\}^{M \times N}$ determines the presence/absence of a sensor in any of the N spatial points, at any t. This is designed by selecting M out of N rows of $\mathbf{I}_N$. The indices of the M selected rows are encoded by the support (set of indices of non-zero entries) of a selection vector defined as $\mathbf{w}_t = [w_{t1}, \dots, w_{tN}]^T \in \{0,1\}^N$, for any snapshot t. So, $[\mathbf{w}_t]_j = 1(0)$ denotes if the measurement at the j-th spatial point is available (unavailable), where $j = 1, \dots, N$. The M measurements are corrupted by additive spatio-temporally white Gaussian noise $\mathbf{e}(t) \sim \mathcal{N}(\mathbf{0}_M, \mathbf{\Sigma}_e)$, where $\mathbf{\Sigma}_e = \sigma_e^2 \mathbf{I}_M$. Further, $\mathbf{e}(t)$ is uncorrelated with $\mathbf{u}(t)$.

2.2. Statistical characterization of the field

In a Bayesian framework, to reconstruct $\mathbf{u}(t)$ from less observations than its dimension, prior knowledge about $\mathbf{u}(t)$ is utilized. We assume that, at any time instance t, the realization of the field at any location $\mathbf{x} = [x, y]^T$, i.e., $u(\mathbf{x}, t)$, is a Gaussian random variable. It is also assumed that $u(\mathbf{x}, t)$ is a zero-mean spatio-temporally (second-order) stationary isotropic process [2]. The elements of the spatio-temporal covariance matrix are derived from a generalized model of a non-separable covariance function, widely used for environmental prediction [10]. For any temporal lag τ, i.e., the difference between two snapshots of $\mathbf{u}(t)$, and any two spatial locations $\mathbf{x}_i, \mathbf{x}_j$, with $h_{ij} \triangleq \|\mathbf{x}_i - \mathbf{x}_j\|_2$, the elements of the spatio-temporal covariance matrix are given by,

$$[\mathbf{\Gamma}_\tau]_{ij} = f_c(h_{ij}, \tau) = \frac{\sigma_u^2}{(a|\tau|^{2\alpha} + 1)^\beta} \exp\left[-\frac{c\, h_{ij}^{2\phi}}{(a|\tau|^{2\alpha} + 1)^{\beta\phi}} \right]. \tag{2}$$

Here, $\alpha, \phi \in (0, 1]$ are the smoothing parameters, while a and c are the non-negative scaling parameters for time and space, respectively. The parameter $\beta \in [0, 1]$ is responsible for the space-time interaction of the covariance. The available lag-0 and lag-τ space-time covariance matrices can be given by $\mathbb{E}[\mathbf{u}(t)\mathbf{u}(t)^T] = \mathbf{\Gamma}_0$, and $\mathbb{E}[\mathbf{u}(t)\mathbf{u}(t-\tau)^T] = \mathbf{\Gamma}_\tau$. It is also clear from (2) that $\mathbf{\Gamma}_\tau = \mathbf{\Gamma}_{(-\tau)}^T$.

2.3. The two observation scenarios and the related MSE

Before we detail the two observation scenarios as mentioned in the introduction, let us first review the MSE matrix in the Bayesian paradigm. If an unknown parameter $\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_\theta)$ is estimated from Gaussian observations $\mathbf{y}$, using a minimum mean square error (MMSE) estimator, then the error covariance/MSE matrix can be given by

$$\mathbb{E}_{\boldsymbol{\theta}|\mathbf{y}}[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T] = \mathbf{\Sigma}_\theta - \mathbf{\Sigma}_{\theta y} \mathbf{\Sigma}_{yy}^{-1} \mathbf{\Sigma}_{y\theta}. \tag{3}$$

Here, $\mathbf{\Sigma}_{\theta y}$ and $\mathbf{\Sigma}_{y\theta}$ are the cross covariance matrices between the unknown parameter and the observations [13].

In the first sensing scenario, the measurement window consists of measurements from the past, present, and the future. Let us assume that the measurements are given by $\tilde{\mathbf{y}} = [\mathbf{y}^T(t - N_s + 1), \dots, \mathbf{y}^T(t), \dots, \mathbf{y}^T(t + N_s - 1)]^T \in \mathbb{R}^{M(2N_s - 1)}$, where $2N_s - 1$ is the size of the measurement window. The parameter to be estimated is $\boldsymbol{\theta} = \mathbf{u}(t) \in \mathbb{R}^N$. Note that the measurement window can be slided over the entire swath of t and the measurements to estimate $\mathbf{u}(t)$ overlap those of $\mathbf{u}(t+1)$. Using (1) for all the measurements at $2N_s - 1$ snapshots we have $\tilde{\mathbf{y}} = \tilde{\mathbf{C}}_t \tilde{\mathbf{u}} + \tilde{\mathbf{e}}$, where $\tilde{\mathbf{C}}_t = \mathbf{I}_{2N_s-1} \otimes \mathbf{C}_t$, $\tilde{\mathbf{u}} = [\mathbf{u}^T(t - N_s + 1), \dots, \mathbf{u}^T(t), \dots, \mathbf{u}^T(t + N_s - 1)]^T \in \mathbb{R}^{N(2N_s - 1)}$ and $\tilde{\mathbf{e}} = [\mathbf{e}^T(t - N_s + 1), \dots, \mathbf{e}^T(t), \dots, \mathbf{e}^T(t + N_s - 1)]^T \in \mathbb{R}^{M(2N_s - 1)}$, respectively.

The individual terms for the expression in (3) can be derived as follows. First of all, we have $\mathbf{\Sigma}_\theta = \mathbf{\Gamma}_0$. The cross terms are $\mathbf{\Sigma}_{\theta y} = \mathbb{E}[\mathbf{u}(t)[\mathbf{u}^T(t - N_s + 1), \dots, \mathbf{u}^T(t), \dots, \mathbf{u}^T(t + N_s - 1)]\tilde{\mathbf{C}}_t^T] = \mathbf{\Gamma}_X \tilde{\mathbf{C}}_t^T$, where $\mathbf{\Gamma}_X = [\mathbf{\Gamma}_{N_s - 1}, \dots, \mathbf{\Gamma}_0, \dots, \mathbf{\Gamma}_{-(N_s - 1)}]$ of size $N \times N(2N_s - 1)$. For the measurements we have $\mathbb{E}[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^T] = \tilde{\mathbf{C}}_t \mathbf{\Gamma}_Y \tilde{\mathbf{C}}_t^T + \sigma^2 \mathbf{I}$ with $\mathbf{\Gamma}_Y = \mathbb{E}[\tilde{\mathbf{u}}\tilde{\mathbf{u}}^T]$, and $\sigma^2 \mathbf{I} = \mathbb{E}[\tilde{\mathbf{e}}\tilde{\mathbf{e}}^T]$. Here, $\mathbf{\Gamma}_Y$ is the symmetric positive-definite space-time covariance matrix of size $N(2N_s - 1) \times N(2N_s - 1)$, whose diagonal and off-diagonal blocks are $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_\tau$ with $\tau = -2N_s + 2, \dots, 2N_s - 2$, respectively. Substituting these terms in (3) we obtain the MSE matrix as a function of the selection vector $\mathbf{w}_t$ which is given as

$$\mathbf{\Sigma}_a(\mathbf{w}_t) = \mathbf{\Gamma}_0 - \mathbf{\Gamma}_X \tilde{\mathbf{C}}_t^T [\tilde{\mathbf{C}}_t \mathbf{\Gamma}_Y \tilde{\mathbf{C}}_t^T + \sigma^2 \mathbf{I}]^{-1} \tilde{\mathbf{C}}_t {\mathbf{\Gamma}_X}^T. \tag{4}$$

For the positive definite matrices $\mathbf{\Gamma}_Y$ and $\sigma^2 \mathbf{I}$, using the matrix inversion lemma (MIL), we have the matrix identity $\mathbf{\Gamma}_Y - \mathbf{\Gamma}_Y \tilde{\mathbf{C}}_t^T (\tilde{\mathbf{C}}_t \mathbf{\Gamma}_Y \tilde{\mathbf{C}}_t^T + \sigma^2 \mathbf{I})^{-1} \tilde{\mathbf{C}}_t \mathbf{\Gamma}_Y = (\mathbf{\Gamma}_Y^{-1} + \frac{1}{\sigma^2} \tilde{\mathbf{C}}_t^T \tilde{\mathbf{C}}_t)^{-1}$ [14]. Using the above identity we have

$$\begin{aligned} \tilde{\mathbf{C}}_t^T (\tilde{\mathbf{C}}_t \mathbf{\Gamma}_Y \tilde{\mathbf{C}}_t^T + \sigma^2 \mathbf{I})^{-1} \tilde{\mathbf{C}}_t &= \\ \mathbf{\Gamma}_Y^{-1} [\mathbf{\Gamma}_Y - (\mathbf{\Gamma}_Y^{-1} + \frac{1}{\sigma^2} \tilde{\mathbf{C}}_t^T \tilde{\mathbf{C}}_t)^{-1}] \mathbf{\Gamma}_Y^{-1}&. \end{aligned} \tag{5}$$

Substituting (5) in (4) we obtain the following expression

$$\mathbf{\Sigma}_a(\mathbf{w}_t) = \mathbf{\Gamma}_0 - \mathbf{\Gamma}_X \mathbf{\Gamma}_Y^{-1} [\mathbf{\Gamma}_Y - (\mathbf{\Gamma}_Y^{-1} + \frac{1}{\sigma^2} \tilde{\mathbf{C}}_t^T \tilde{\mathbf{C}}_t)^{-1}] \mathbf{\Gamma}_Y^{-1} \mathbf{\Gamma}_X^T \tag{6}$$

Comment: *In some cases, due to the parameters of the covariance function, number of grid points, size of the mesh etc. [15],* $\mathbf{\Gamma}_Y$ *can be ill-conditioned. In such cases, the aforementioned formulation can be generalized by replacing* $\mathbf{\Gamma}_Y$ *by* $\mathbf{Z} - \beta \mathbf{I}$*, where* $\mathbf{Z} = (\mathbf{\Gamma}_Y + \beta \mathbf{I})$ *is the well-conditioned matrix, with* $\beta > 0$. *Applying the fact that* $\tilde{\mathbf{C}}_t \tilde{\mathbf{C}}_t^T = \mathbf{I}$*, and using the MIL as before, an alternative form of (6) can be given as,*

$$\mathbf{\Sigma}_a(\mathbf{w}_t) = \mathbf{\Gamma}_0 - \mathbf{\Gamma}_X \mathbf{Z}^{-1} [\mathbf{Z} - (\mathbf{Z}^{-1} + \frac{1}{\sigma^2 - \beta} \tilde{\mathbf{C}}_t^T \tilde{\mathbf{C}}_t)^{-1}] \mathbf{Z}^{-1} \mathbf{\Gamma}_X^T.$$

However, for the time being, assuming $\mathbf{\Gamma}_Y$ is well-conditioned, and using the fact that $\mathbf{C}_t^T \mathbf{C}_t = \mathrm{diag}(\mathbf{w}_t)$, (6) can also be written as,

$$\begin{aligned} \mathbf{\Sigma}_a(\mathbf{w}_t) = \mathbf{\Gamma}_0 - \mathbf{\Gamma}_X \mathbf{\Gamma}_Y^{-1} [\mathbf{\Gamma}_Y - (\mathbf{\Gamma}_Y^{-1} +& \\ \frac{1}{\sigma^2} (\mathbf{I}_{2N_s - 1} \otimes \mathrm{diag}(\mathbf{w}_t)))^{-1}] \mathbf{\Gamma}_Y^{-1} \mathbf{\Gamma}_X^T&. \end{aligned} \tag{7}$$

In the second scenario, the available measurements are the same as before i.e., $\tilde{\mathbf{y}} \in \mathbb{R}^{M(2N_s-1)}$ but the parameter to be estimated now is $\boldsymbol{\theta} = \tilde{\mathbf{u}} \in \mathbb{R}^{N(2N_s-1)}$. The set of time-varying sensor constellations, i.e., the observation matrices are $\mathbf{C}_{t-N_s+1}, \dots, \mathbf{C}_t, \dots, \mathbf{C}_{t+N_s-1}$. In this case, we have $\boldsymbol{\Sigma}_\theta = \mathbb{E}[\tilde{\mathbf{u}}\tilde{\mathbf{u}}^T] = \boldsymbol{\Gamma}_Y$.

Let us define the joint space-time observation matrix as $\breve{\mathbf{C}} = \text{blkdiag}(\mathbf{C}_{t-N_s+1}, \dots, \mathbf{C}_t, \dots, \mathbf{C}_{t+N_s-1})$ of size $M(2N_s - 1) \times N(2N_s - 1)$, where $\breve{\mathbf{C}}$ is a block diagonal matrix with $\mathbf{C}_{t-N_s+1}, \dots, \mathbf{C}_t, \dots, \mathbf{C}_{t+N_s-1}$ as diagonal blocks. The spatio-temporal noise components are the same as before, i.e., $\tilde{\mathbf{e}} \in \mathbb{R}^{M(2N_s-1)}$.

In this case, the error covariance matrix can be directly derived from (3) as a multivariate function of the selection vectors $\mathbf{w}_{t-N_s+1}, \dots, \mathbf{w}_t, \dots, \mathbf{w}_{t+N_s-1}$. This can be given as $\boldsymbol{\Sigma}_b(\mathbf{w}_{t-N_s+1}, \dots, \mathbf{w}_{t+N_s-1}) = \boldsymbol{\Gamma}_Y - \boldsymbol{\Gamma}_Y\breve{\mathbf{C}}^T(\breve{\mathbf{C}}\boldsymbol{\Gamma}_Y\breve{\mathbf{C}}^T + \sigma^2\mathbf{I})^{-1}\breve{\mathbf{C}}\boldsymbol{\Gamma}_Y = (\boldsymbol{\Gamma}_Y^{-1} + \frac{1}{\sigma^2}\breve{\mathbf{C}}^T\breve{\mathbf{C}})^{-1}$. Using the same trick as before, i.e., $\mathbf{C}_t^T\mathbf{C}_t = \text{diag}(\mathbf{w}_t)$, the MSE matrix as a function of the selection vectors can be represented as,

$$\boldsymbol{\Sigma}_b(\mathbf{w}_{t-N_s+1}, \dots, \mathbf{w}_{t+N_s-1}) = [\boldsymbol{\Gamma}_Y^{-1} + \frac{1}{\sigma^2}\text{blkdiag}[\text{diag}(\mathbf{w}_{t-N_s+1}), \dots, \text{diag}(\mathbf{w}_{t+N_s-1})]]^{-1}. \quad (8)$$

3. PROBLEM FORMULATION

3.1. Main problem

We formulate an optimal *space-time sensor location selection* problem assuming that all the mid points of the pixels are candidate sensor locations (see Section 2.1). This is accomplished by designing all the *selection vectors* $\mathbf{w}_t$, for both scenarios, constrained by the desired accuracy requirement (MSE) expressed as a function of $\mathbf{w}_t$. Let us assume that, the desired upper bounds on the MSE to estimate $\mathbf{u}(t)$ and $\tilde{\mathbf{u}}$ are given by γ_a and γ_b, respectively. The optimization problem for the first case can be expressed as

$$\underset{\mathbf{w}_t \in \{0,1\}^N}{\arg\min} \quad \|\mathbf{w}_t\|_0 \quad (9a)$$

$$\text{s.t.} \quad \text{tr}[\boldsymbol{\Sigma}_a(\mathbf{w}_t)] \le \gamma_a. \quad (9b)$$

For the next case, i.e., to design the selection vectors for all $2N_s - 1$ snapshots, the optimization problem is given by

$$\underset{\mathbf{w}_{t-N_s+1},\dots,\mathbf{w}_{t+N_s-1} \in \{0,1\}^N}{\arg\min} \quad \sum_{i=-N_s+1}^{N_s-1} \|\mathbf{w}_{t+i}\|_0 \quad (10a)$$

$$\text{s.t.} \quad \text{tr}[\boldsymbol{\Sigma}_b(\mathbf{w}_{t-N_s+1}, \dots, \mathbf{w}_{t+i}, \dots, \mathbf{w}_{t+N_s-1})] \le \gamma_b. \quad (10b)$$

Note that the performance constraints in (9b) and (10b) are related to the functions of the spatio-temporal covariance matrix as derived in (7) and (8), respectively.

3.2. Sparsity-enforcing correlation-aware sensor location selection

The optimization problems of (9) and (10) are intractable due to the following reasons: the ℓ_0 norm in the cost function makes the problem NP-hard and non-convex and the Boolean constraint $\mathbf{w}_t \in \{0,1\}^N$, is non-convex and combinatorially complex. We use standard convex relaxations and formulate the problems of (9) and (10) with the MSE cost functions derived in (7) and (8) for the two aforementioned sensing modalities. The overall optimization problems in semi-definite form with the appropriate convex relaxations can be formulated as the following optimization problems. The relaxed version of (9) in semi-definite form is given as

$$\underset{\mathbf{w}_t \in \mathbb{R}^N, \mathbf{U} \in \mathbb{S}^{N(2N_s-1)}}{\arg\min} \quad \|\mathbf{w}_t\|_1 \quad (11a)$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{U} & \mathbf{I} \\ \mathbf{I} & [\boldsymbol{\Gamma}_Y^{-1} + \frac{1}{\sigma_a^2}(\mathbf{I}_{(2N_s-1)} \otimes \text{diag}(\mathbf{w}_t))] \end{bmatrix} \succeq 0, \quad (11b)$$

$$\text{tr}\,[\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_X\boldsymbol{\Gamma}_Y^{-1}[\boldsymbol{\Gamma}_Y - \mathbf{U}]\boldsymbol{\Gamma}_Y^{-1}\boldsymbol{\Gamma}_X^T] \le \gamma_a, \quad (11c)$$

$$0 \le [\mathbf{w}_t]_j \le 1, \quad j = 1, \dots, N. \quad (11d)$$

In (11a), the ℓ_1-norm is a standard convex relaxation for $\|\mathbf{w}_t\|_0$, that enforces sparsity in selection. The threshold on the MSE performance in (11c) is given by γ_a. The $\mathbf{w}_t \in \{0,1\}^N$ constraint is relaxed to $\mathbf{w}_t \in [0,1]^N$ in (11d). The solution of the optimization problem (11) results in a constant $\mathbf{w}_t = \mathbf{w}$, which can be used to estimate $\mathbf{u}(t)$ from all the measurements from $(2N_s - 1)$ snapshots.

Next, we design the set of sparse selection vectors, to organize the sensor deployment in every snapshot in order to estimate the field intensity in all of these snapshots jointly. Assuming the performance threshold is γ_b, the optimization problem to design $\mathbf{w}_{t-N_s+1}, \dots, \mathbf{w}_{t+N_s-1}$ can be given as

$$\underset{\substack{\mathbf{w}_{t-N_s+1}, \dots, \mathbf{w}_{t+N_s-1} \in \mathbb{R}^N \\ \mathbf{V} \in \mathbb{S}^{N(2N_s-1)}}}{\arg\min} \quad \sum_{i=-N_s+1}^{N_s-1} \|\mathbf{w}_{t+i}\|_1 \quad (12a)$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{V} & \mathbf{I} \\ \mathbf{I} & \boldsymbol{\Sigma}_b^{-1}(\mathbf{w}_{t-N_s+1}, \dots, \mathbf{w}_{t+i}, \dots, \mathbf{w}_{t+N_s-1}) \end{bmatrix} \succeq 0, \quad (12b)$$

$$\text{tr}\,[\mathbf{V}] \le \gamma_b, \quad (12c)$$

$$0 \le [\mathbf{w}_{t+i}]_j \le 1, \quad i = -N_s + 1, \dots, N_s - 1;\ j = 1, \dots, N, \quad (12d)$$

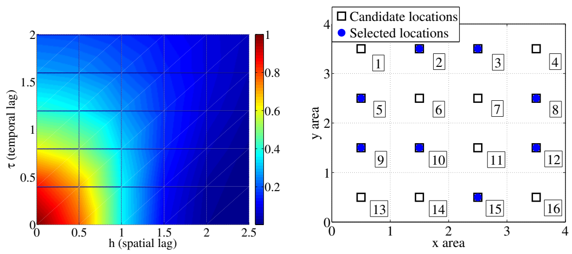
$$\|\mathbf{w}_{t+i}\|_1 \ge p, \quad i = -N_s + 1, \dots, N_s - 1. \quad (12e)$$

Note that, the constraint in (12e) implies that at least p sensor locations will be selected in every snapshot. This *design constraint* is applied to utilize the measurements on individual snapshots more efficiently, i.e., to distribute the selected sensing locations over all snapshots. Using this constraint, we avoid solutions where all locations are selected in a single snapshot, which may satisfy the performance constraint but not in an energy-efficient design. The value of p generally depends upon the application and the available resources.

3.3. Discussion

The solution of (11) and (12) gives respectively the selection vector $\mathbf{w}$ and the set of selection vectors $\mathbf{w}_{t-N_s+1}, \dots, \mathbf{w}_{t+N_s-1}$ achieving the desired performance γ_a and γ_b. It is clear that lowering the values of γ_a and γ_b, i.e., putting a tighter threshold on the performance, more sensor locations are needed to be selected. One way to calculate the performance threshold to estimate $\mathbf{u}(t)$, i.e., γ_a, is by scaling the best MSE which occurs when all sensors are present, i.e., $\gamma_a = \lambda\text{tr}[\boldsymbol{\Sigma}_a(\mathbf{1}_N)]$, where $\lambda \ge 1$.

As mentioned in [16], the unwanted dependence on the magnitude of the elements can be avoided by using an iterative re-weighted



(a) (b)

Fig. 1: (a) Space-time variation of $f_c(h, \tau)$. (b) Selected sensor locations with $N_s = 1$.

ℓ_1-norm minimization technique. Hence, we will solve (11) and (12) using the same sparsity enhancing iterative algorithm as in [16] . To design the selection vectors exactly to be in the set $\{0, 1\}$ we use the randomization method described in [5]. It is repeated briefly below. If a solution of the iterative version of (11) is $\hat{\mathbf{w}}_t = [\hat{w}_{t1}, \ldots, \hat{w}_{tN}]^T$, then many random realizations of $\mathbf{w}_t$ are generated, where the probability of $w_{ti} = 1$ is specified by $\hat{w}_{ti}$, with $i = 1, \ldots, N$. Here we denote the locations with high probability of being $w_{ti} = 1$, (i.e., $|\hat{w}_{ti}|$ is very close to 1) as *informative* sensor locations. Others are considered to be *non-informative* sensor locations. Then, those realizations of $\mathbf{w}_t$ are selected which satisfy the performance constraint, and the minimum norm $\mathbf{w}_t$ is selected. The resulting $\mathbf{w}_t \in \{0, 1\}^N$ is the close-to-optimal solution of (11) whose support denotes the sparsely distributed sensor locations. The same approach is followed to estimate $\mathbf{w}_{t-N_s+1}, \ldots, \mathbf{w}_{t+N_s-1}$ from the optimization problem (12).

4. SIMULATION RESULTS

We assume a given service area is uniformly discritized in $N = 16$ pixels, where we would like to measure the field intensities. The candidate sensor locations are the midpoints of the pixels (see Fig.1(b)). From all these coordinates, the distance matrix (matrix of all possible pair-wise Euclidean distances) is calculated and the spatial covariance matrix for any lag τ is calculated.

In Fig.1(a), we plot the non-separable space-time covariance function, with different space and time lags. The selected parameters are $\sigma_u = 1, \alpha, \ \phi = 1, a, \ c = 1$, and $\beta = 1$, i.e., we consider the maximum space-time interaction. We further take $\sigma^2 = 1$. We use the MATLAB implementation of CVX [17] for solving the semidefinite programming (SDP) problems of (11) and (12).

The performance threshold to estimate $\mathbf{u}(t)$, i.e., γ_a is calculated using $N_s = 1$, $\lambda = 1.5$, and the aforementioned smoothness and scaling parameters for the covariance function. The optimization problem of (10) is first solved with $N_s = 1$, i.e., only the present measurements are used. We use 20 iterations of the sparsity enhancing iterative algorithm with $\epsilon = 10^{-8}$ [16]. Then, the randomization method is applied to 5000 random realizations of $\hat{\mathbf{w}}_t$ to calculate $\mathbf{w}_t \in \{0, 1\}$. In Fig.1(b), the selected sensor locations are shown. The indices of the sensors are given in numbers in Fig.1(b). To explore the effect of the space-time covariance in sensor selection, the following numerical experiments are performed. We use $N_s = 2$ with the same γ_a. Keeping a, c, and β fixed, the smoothness parameters, i.e., α and ϕ are varied. The resulting sparse sensor locations are exhibited in Table 1. First of all, we note that the optimal sensor locations change when the past and future measurements are also used. The selection pattern in Fig.1(b) is for $N_s = 1$, whereas Table 1 shows the selection patterns for $N_s = 2$. With $N_s = 1$, no temporal correlation is involved, i.e., $\mathbf{\Gamma}_Y = \mathbf{\Gamma}_X = \mathbf{\Gamma}_0$. With $N_s = 2$, i.e., taking also the past and future measurements into account, the required number of sensor locations reduces.

Table 1: Numerical results: optimization problem (11)

No.	Parameters	Selected senor locations
I	$\alpha = 1; \ \phi = 1$	$\{2, 3, 5, 8, 12, 14, 15\}$
II	$\alpha = 0.5; \ \phi = 1$	$\{2, 3, 5, 8, 9, 12, 14\}$
III	$\alpha = 1; \ \phi = 0.5$	$\{2, 5, 8, 9, 12, 14, 15\}$
IV	$\alpha = 0.5; \ \phi = 0.5$	$\{2, 3, 5, 8, 9, 14, 15\}$

In the next case, we use $\gamma_b = (2N_s - 1)\gamma_a$. We solve the optimization problem of (12) for $N_s = 2$, i.e., we solve for $\mathbf{w}_{t-1}, \mathbf{w}_t, \mathbf{w}_{t+1}$ iteratively for 20 iterations with the same $\epsilon = 10^{-8}$. We take $p = 1$, i.e., at least one sensing location will be selected on every snapshot. We denote the sets of the indices of the non-zero entries of $\mathbf{w}_{t-1}, \mathbf{w}_t, \mathbf{w}_{t+1}$ as $k_{-1}, \ k_0$, and k_1. The obtained $k_{-1}, \ k_0$, and k_1 for different smoothing parameters with the fixed γ_b are given in Table 2.

Table 2: Numerical results: optimization problem (12)

No.	Parameters	Selected senor locations
I	$\alpha = 1; \ \phi = 1$	$k_{-1} = \{2, 5, 9, 12\}$ $k_0 = \{1, 3, 6, 7, 8, 11, 13, 14, 16\}$ $k_1 = \{2, 3, 5, 8, 12\}$
II	$\alpha = 0.5; \ \phi = 1$	$k_{-1} = \{2, 5, 8, 12, 14\}$ $k_0 = \{1, 3, 6, 10, 11, 12, 13\}$ $k_1 = \{2, 3, 5, 9, 14, 15\}$
III	$\alpha = 1; \ \phi = 0.5$	$k_{-1} = \{6, 7, 10\}$ $k_0 = \{1, 2, 3, 4, 5, 8, 9, 12, 13, 14, 15, 16\}$ $k_1 = \{6, 7, 11\}$
IV	$\alpha = 0.5; \ \phi = 0.5$	$k_{-1} = \{7, 10, 11\}$ $k_0 = \{1, 2, 3, 4, 5, 8, 9, 12, 13, 14, 15, 16\}$ $k_1 = \{6, 10, 11\}$

From Table 1 it is seen that, altering the spatial/temporal smoothness, the selected sensor locations slightly change. The set of the most informative sensor locations, i.e., for instance $\{2, 5, 8, 14\}$ are the same in most of the cases. In Table 2, the average number of locations over 3 snapshots is 6, for all four cases but their constellations change with different smoothing parameters. Here, the number of required locations per snapshot is less than when only a single snapshot of $\mathbf{u}(t)$ is estimated separately (Table 1).

From the resulting selection patterns in Tables 1 and 2, it is seen that the selected locations are more or less uniformly distributed over the entire area rather than forming clusters in a specific area or snapshot. Further simulations show that, increasing the values of γ_a, γ_b, less locations are selected but the aforementioned uniformity is generally present. However, selection patterns will vary with a different class of the covariance function f_c.

5. CONCLUSION AND FUTURE WORK

It is seen that, MSE-optimal space-time sensor deployment can be designed to be efficient (less required sensing locations) in different space-time sensing modalities, availing the prior correlation information of the physical field. As many practical environmental fields exhibit space-time sparsity in a proper representation basis, further research is envisioned to develop an optimal sensor constellation that is jointly MSE-optimal as well as satisfies the criteria for uniqueness in a sparse recovery framework.

6. REFERENCES

[1] S. Liu, A. Vempaty, M. Fardad, E. Masazade, and P.K. Varshney, "Energy-aware sensor selection in field reconstruction," *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1476–1480, December 2014.

[2] N. Cressie and K. Wikle, *Statistics for spatio-temporal data*, John Wiley & Sons, 2011.

[3] S. Liu, E. Masazade, M. Fardad, and P.K. Varshney, "Sparsity-aware field estimation via ordinary Kriging," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 3976–3980.

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

[5] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *arXiv preprint arXiv:1310.5251*, 2013.

[6] S. P. Chepuri and G. Leus, "Sparsity-exploiting anchor placement for localization in sensor networks," in *Proc. 21st European Signal Processing Conference (EUSIPCO)*, Marrakech, Morocco, September 2013, pp. 1–5.

[7] V. Roy, S. P. Chepuri, and G. Leus, "Sparsity-enforcing sensor selection for DOA estimation," in *Proc. 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Saint Martin, December 2013, pp. 340–343.

[8] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *The Journal of Machine Learning Research*, vol. 9, pp. 235–284, 2008.

[9] J. Ranieri, A. Chebira, and M. Vetterli, "Near-optimal sensor placement for linear inverse problems," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1135–1146, March 2014.

[10] T. Gneiting, "Nonseparable, stationary covariance functions for space–time data," *Journal of the American Statistical Association*, vol. 97, no. 458, pp. 590–600, June 2002.

[11] S. Boyd and S. Vandenberghe, *Convex optimization*, Cambridge university press, 2009.

[12] I. Rodriguez-Iturbe, M. Marani, P. D'Odorico, and A. Rinaldo, "On space-time scaling of cumulated rainfall fields," *Water resources research*, vol. 34, no. 12, pp. 3461–3469, December 1998.

[13] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*, PTR Prentice-Hall, Englewood Cliffs, NJ, 1993.

[14] Max Welling, "The Kalman Filter," *Lecture Note*, 2010.

[15] R. Ababou, A. C. Bagtzoglou, and E. F. Wood, "On the condition number of covariance matrices in kriging, estimation, and simulation of random fields," *Mathematical Geology*, vol. 26, no. 1, pp. 99–133, 1994.

[16] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[17] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," 2008.