2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

SPARSITY-AWARE FIELD ESTIMATION VIA ORDINARY KRIGING

Sijia Liu[1], *Engin Masazade*[2], *Makan Fardad*[1], *Pramod K. Varshney*[1]

[1] Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA
[2] Department of Electrical and Electronics Engineering, Yeditepe University, Istanbul, 34755 Turkey
[1]{sliu17, makan, varshney}@syr.edu, [2]engin.masazade@yeditepe.edu.tr

***Abstract*—In this paper, we consider the problem of estimating a spatially varying field in a wireless sensor network, where resource constraints limit the number of sensors selected in the network that provide their measurements for field estimation. Based on a one-to-one correspondence between the selected sensors and the nonzero elements of Kriging weights, we propose a sparsity-promoting ordinary Kriging approach where we minimize the Kriging error variance while penalizing the number of nonzero Kriging weights. This yields a combinatorial optimization problem, which is intractable in general. To solve the proposed non-convex optimization problem, we employ the alternating direction method of multipliers (ADMM) and the reweighted ℓ_1 minimization method, respectively. Numerical results are provided to illustrate the effectiveness of our proposed approaches that provide a balance between the estimation accuracy and the number of selected sensors.**

Keywords*—Field estimation, alternating direction method of multipliers, sparsity, sensor networks, convex optimization.

I. INTRODUCTION

In this paper, we study the problem of field estimation, where in a given region of interest the spatially-correlated field is monitored by deployed sensors. The field intensity at a particular point of interest is estimated by using the sensor measurements. Over the last few decades, field estimation problems have been widely studied in the literature [1]–[4]. In [1], the authors explored the spatial correlation in a dynamic field and proposed a general model to reconstruct this correlation via the sensor data. In [2], [3], Kriging, a class of geostatistical techniques, is presented where the value of a random field at an unobserved location is interpolated from sensor measurements. In [4], the authors compared different types of Kriging estimators, where it was indicated that *simple* Kriging is mathematically the simplest but *ordinary* Kriging is the most commonly used approach.

In wireless sensor networks, due to limited communication and energy resources, it is desirable to select only a subset of sensors for field estimation. Therefore, the issue of sensor selection arises, which aims to strike a balance between estimation accuracy and sensor activations. From the energy point of view, the problem of sensor selection in the context of field estimation has drawn extensive attention recently [1], [5]–[8]. In [5], [6], the authors assume that the field evolves as a linear dynamical system, and thus the sensor selection problem can be formulated via a Kalman filter. In [1], [7], the sensors are scheduled in time and a staggered sensing strategy is proposed to improve the estimation performance based on ordinary Kriging. In [8], simple Kriging is applied for field estimation, and an iterative algorithm is proposed to obtain a suboptimal sensor selection scheme by minimizing the estimation error variance. However, this approach deals with the estimation and sensor selection procedures separately.

Compared to the existing work, in this paper we propose a sparsity-promoting ordinary Kriging procedure which carries out the tasks of sensor selection and field estimation jointly. This is based on the fact that a sensor being off is equivalent to the corresponding element of Kriging weights being zero. As our objective, we minimize the estimation error variance while including a cardinality function that counts and penalizes the number of active sensors in terms of the nonzero elements of Kriging weights. The proposed cardinality-aware optimization problem is combinatorial in nature, and we apply the alternating direction method of multipliers (ADMM) [9] to find its locally optimal solutions. The ADMM algorithm has been successfully applied to the problem of optimal sparse feedback control [10]–[12] and target tracking [13], here, we apply it to the field estimation problem in this paper.

Since the convergence of cardinality-aware ADMM is not guaranteed [9], [12], [13], we further study a relaxed convex problem by replacing the cardinality function with the reweighted ℓ_1 norm [6], [14]. Then, the resulting problem can be solved by either using quadratic programming or ADMM where we show that the reweighted ℓ_1-norm based ADMM yields lower computational complexity as compared to quadratic programming for large-scale sensor networks. The effectiveness of our approach is demonstrated in terms of estimation performance and computational complexity in numerical examples.

II. SYSTEM MODEL

In this paper, the task of the wireless sensor network is to estimate the field intensity at a particular location $\mathbf{s}_0$ given in a square region (denoted by $\mathcal{A}$) with length L. As shown in Fig. 1, we assume that $N = 33$ sensors are randomly distributed in $\mathcal{A}$ according to a Poisson distribution [5]. And the field intensity at the point of interest $\mathbf{s}_0 = (4, 5)$ will be estimated. We emphasize that the proposed framework in this article is also applicable for estimating field intensity at *multiple* locations.

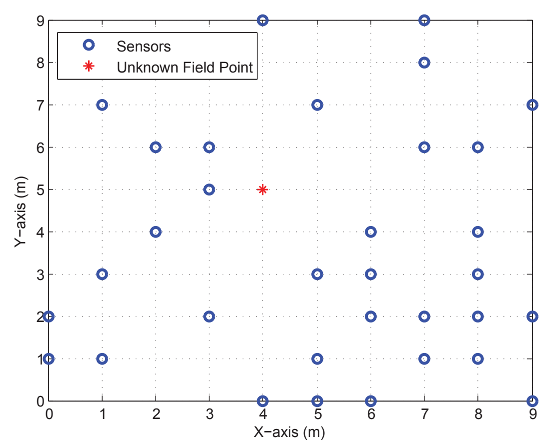


Fig. 1: A random field with $N = 33$ sensors and field intensity is estimated at $\mathbf{s}_0 = (4, 5)$.

Let $\{z(\mathbf{s}), \mathbf{s} \in \mathcal{A}\}$ denote a spatially-correlated random field, where $\mathbf{s} = (x, y)$ denotes the location of a field point. The covariance of two arbitrary field points $z(\mathbf{s}_a)$ and $z(\mathbf{s}_b)$ characterizes their spatial correlation and is described by an exponential model [1], [7]

$$\mathrm{Cov}(z(\mathbf{s}_a), z(\mathbf{s}_b)) = \tau e^{-\lambda \|\mathbf{s}_a - \mathbf{s}_b\|_2}, \quad (1)$$

where $\|\cdot\|_2$ is the 2-norm of a vector, τ is the variance of the field at an arbitrary location and λ governs the strength of spatial correlation which increases (or decreases) as λ decreases (or increases). Here we assume that the value of the correlation parameter is known in advance. Moreover, we assume that the mean value of the field is constant but unknown over the entire region, i.e., $\mathrm{E}(\mathbf{z}(\mathbf{s})) = \mu$. Therefore, we use ordinary Kriging for estimating the field intensity at a particular point of interest [3].

A. *Ordinary Kriging for Field Estimation*

Ordinary Kriging is a linear estimator to interpolate the value of a random field at an unobserved location from sensor measurements [2]. Let $\mathbf{s}_i$ denote the position of the ith sensor $1 \leq i \leq N$ and $z(\mathbf{s}_i)$ be the corresponding sensor measurement, where all sensor measurements are assumed to be noise-free [8]. The field intensity to be estimated is represented by $z(\mathbf{s}_0)$, where $\mathbf{s}_0$ denotes the Cartesian coordinates of an unobserved location. The ordinary Kriging estimator is defined as [2], [3],

$$\hat{z}(\mathbf{s}_0) = \sum_{i=1}^{N} w_i z(\mathbf{s}_i) \tag{2}$$

where $\hat{z}(\mathbf{s}_0)$ is the ordinary Kriging estimate, w_i is the Kriging weight assigned to the ith sensor. Ordinary Kriging finds the vector of Kriging weights,$\mathbf{w} = [w_1, \ldots, w_N]^T$, which minimize the Kriging error variance subject to the constraint $\sum_{i=1}^{N} w_i = 1$. The Kriging error variance (KEV) (also known as mean square error in estimation theory) is then defined as [7], [8],

$$\begin{aligned}\sigma^2(\mathbf{w}) &= \mathrm{Var}[\hat{z}(\mathbf{s}_0) - z(\mathbf{s}_0)] \\ &= \mathbf{w}^T \boldsymbol{\mathcal{K}} \mathbf{w} - 2\boldsymbol{\kappa}^T \mathbf{w} + \mathrm{Cov}[z(\mathbf{s}_0), z(\mathbf{s}_0)],\end{aligned} \tag{3}$$

where

$$\boldsymbol{\mathcal{K}} = \mathrm{Cov}\{[z(\mathbf{s}_1), \ldots, z(\mathbf{s}_N)]^T, [z(\mathbf{s}_1), \ldots, z(\mathbf{s}_N)]^T\} \tag{4}$$

$$\boldsymbol{\kappa} = \mathrm{Cov}\{[z(\mathbf{s}_1), \ldots, z(\mathbf{s}_N)]^T, z(\mathbf{s}_0)\} \tag{5}$$

and the covariance matrix is given by the $\mathrm{Cov}(\cdot)$ operator in (1) acting on its individual entries.

Remark 1: In (3), the Kriging error variance (KEV) yields a *quadratic* form with respect to $\mathbf{w}$ and depends on the spatial correlation of the field given by (1).

The ordinary Kriging weights $\mathbf{w}$ are obtained by minimizing the KEV, which yields the following *constrained* optimization problem

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{minimize}} & \sigma^2(\mathbf{w}) \\ \text{subject to} & \sum_{i=1}^{N} w_i = 1.\end{array} \tag{6}$$

In (6), the optimal $\mathbf{w}$ can be computed via the Lagrangian multiplier method.

III. SPARSITY-PROMOTING KRIGING

It is clear from the ordinary Kriging estimator in (2) that the number of nonzero elements in $\mathbf{w}$ characterizes the number of selected sensors for field estimation. From the perspective of energy, it is important to seek sparse Kriging weights to achieve the optimal tradeoff between the estimation performance and the number of selected sensors. Motivated by [10], [12], [13], we count and penalize the number of nonzero elements of $\mathbf{w}$ by incorporating the cardinality function in the objective function. Therefore, the problem of sparse ordinary Kriging can be cast as

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{minimize}} & \frac{1}{2}\sigma^2(\mathbf{w}) + \gamma \mathrm{card}(\mathbf{w}) \\ \text{subject to} & \sum_{i=1}^{N} w_i = 1,\end{array} \tag{7}$$

where the KEV $\sigma^2(\mathbf{w})$ is defined in (3), $\mathrm{card}(\cdot)$ denotes the cardinality function (i.e. the number of nonzero elements of a vector), and the positive scalar γ is a sparsity-promoting parameter which implies the relative importance of achieving good estimation performance versus activating a small number of sensors (i.e., the sparsity of $\mathbf{w}$).

Note that the presence of $\mathrm{card}(\cdot)$ in (7) results in a combinatorial problem that is computationally intractable in general [15]. However, it has been recently observed in [10]–[13] that the alternating direction method of multipliers (ADMM) is a powerful tool in solving optimization problems that include cardinality functions. Another widely-used technique is to replace the cardinality term with the weighted ℓ_1 norm [6], [12], [14]. In this paper, it will be shown that either of the above methods can be successfully applied to solving problem (7).

A. *Sparsity-Promoting Kriging via ADMM*

We begin by reformulating the optimization problem (7) in a way that lends itself to the application of ADMM. By introducing the indicator function [10] of the constraint set and the auxiliary variables $\mathbf{v}$, the problem (7) can be expressed as

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{minimize}} & \frac{1}{2}\sigma^2(\mathbf{w}) + \gamma \mathrm{card}(\mathbf{v}) + \mathcal{I}(\mathbf{w}) \\ \text{subject to} & \mathbf{w} = \mathbf{v},\end{array} \tag{8}$$

where $\mathcal{I}(\mathbf{w})$ is an indicator function corresponding to the constraint set $\sum_{i=1}^{N} w_i = 1$, and defined as follows

$$\mathcal{I}(\mathbf{w}) = \begin{cases} 0 & \text{if } \sum_{i=1}^{N} w_i = 1 \\ +\infty & \text{otherwise.} \end{cases}$$

The augmented Lagrangian [9] corresponding to optimization problem (8) is given by

$$\begin{aligned}\mathcal{L}(\mathbf{w}, \mathbf{v}, \boldsymbol{\xi}) =& \frac{1}{2}\sigma^2(\mathbf{w}) + \gamma \mathrm{card}(\mathbf{v}) + \mathcal{I}(\mathbf{w}) \\ &+ \boldsymbol{\xi}^T(\mathbf{w} - \mathbf{v}) + \frac{\rho}{2}\|\mathbf{w} - \mathbf{v}\|_2^2,\end{aligned} \tag{9}$$

where the vector $\boldsymbol{\xi}$ is the Lagrange multiplier (a.k.a the dual variable), and the scalar $\rho > 0$ is a penalty weight.

The ADMM algorithm finds the minimum of (8) by solving the following optimization problems iteratively [9], [12] (the iteration step $k = 0, 1, \ldots$),

$$\mathbf{w}^{k+1} := \underset{\mathbf{w}}{\arg\min}\ \mathcal{L}(\mathbf{w}, \mathbf{v}^k, \boldsymbol{\xi}^k), \tag{10}$$

$$\mathbf{v}^{k+1} := \underset{\mathbf{v}}{\arg\min}\ \mathcal{L}(\mathbf{w}^{k+1}, \mathbf{v}, \boldsymbol{\xi}^k), \tag{11}$$

$$\boldsymbol{\xi}^{k+1} := \boldsymbol{\xi}^k + \rho(\mathbf{w}^{k+1} - \mathbf{v}^{k+1}), \tag{12}$$

until both of the conditions

$$\|\boldsymbol{w}^{k+1} - \boldsymbol{v}^{k+1}\|_2 \leq \epsilon, \quad \|\boldsymbol{v}^{k+1} - \boldsymbol{v}^k\|_2 \leq \epsilon$$

are satisfied.

The rationale behind using ADMM is that we could effectively separate the original non-differentiable problem into a "$\mathbf{w}$-minimization step" (10) and a "$\mathbf{v}$-minimization step" (11), of which the former can be addressed using the Lagrangian multiplier method and the latter can be solved analytically.

1) $\mathbf{w}$-minimization step: Completing the squares with respect to $\mathbf{w}$ in the augmented Lagrangian (9) and recalling the definition of $\mathcal{I}(\mathbf{w})$, the $\mathbf{w}$-minimization step in (10) can be expressed as [9], [12]

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{minimize}} & \phi(\mathbf{w}) = \frac{1}{2}\sigma^2(\mathbf{w}) + \frac{\rho}{2}\|\mathbf{w} - \mathbf{a}^k\|_2^2 \\ \text{subject to} & \mathbf{1}^T \mathbf{w} = 1,\end{array} \tag{13}$$

where $\mathbf{a}^k = \mathbf{v}^k - \frac{1}{\rho}\boldsymbol{\xi}^k$ and $\mathbf{1}$ denotes the vector of all ones.

Note that the optimal solution of problem (13) can be obtained by using the Lagrangian multiplier method. To be specific, by introducing the multiplier η, the Lagrangian of (13) can be written as

$$\mathcal{L}_\phi = \frac{1}{2}\sigma^2(\mathbf{w}) + \frac{\rho}{2}\|\mathbf{w} - \mathbf{a}^k\|_2^2 + \eta(\mathbf{1}^T\mathbf{w} - 1).$$

Taking the first derivatives of $\mathcal{L}_\phi$ with respect to $\mathbf{w}$ and η and setting them equal to zero, yields the optimal $\mathbf{w}$ that satisfies

$$\begin{bmatrix}\mathbf{w}\\ \eta\end{bmatrix} = \begin{bmatrix}\boldsymbol{\mathcal{K}} + \rho\mathbf{I} & \mathbf{1}\\ \mathbf{1}^T & 0\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{\kappa} + \rho\boldsymbol{a}^k\\ 1\end{bmatrix} \tag{14}$$

where $\mathbf{I}$ is an identity matrix, $\mathbf{1}$ is a vector consisting of all ones, $\boldsymbol{\mathcal{K}}$ and $\boldsymbol{\kappa}$ are defined in (4) and (5), respectively.

2) $\mathbf{v}$-minimization step: Completing the squares with respect to $\mathbf{v}$ in the augmented Lagrangian (9), the $\mathbf{v}$-minimization step can be written as [9], [12]

$$\underset{\boldsymbol{v}}{\text{minimize}}\, \psi(\boldsymbol{v}) = \gamma \text{card}(\boldsymbol{v}) + \frac{\rho}{2}\|\boldsymbol{v} - \boldsymbol{b}^k\|_2^2, \tag{15}$$

where $\boldsymbol{b}^k = \frac{1}{\rho}\boldsymbol{\xi}^k + \boldsymbol{w}^{k+1}$. Using the separability properties of both the cardinality function and the 2 norm of a vector, the problem (15) can be decomposed

$$\underset{\boldsymbol{v}_i}{\text{minimize}}\, \psi_i(\boldsymbol{v}_i) = \gamma \text{card}(v_i) + \frac{\rho}{2}(v_i - b_i^k)^2, \tag{16}$$

where v_i and b_i^k are the ith element of the vector $\mathbf{v}$ and $\mathbf{b}^k$, respectively. The optimal solution of (16) has been proposed by [12, Eq.15], i.e.,

$$v_i = \begin{cases} 0 & |b_i^k| \le \sqrt{2\gamma/\rho} \\ b_i^k & |b_i^k| > \sqrt{2\gamma/\rho} \end{cases} \tag{17}$$

The ADMM algorithm iteratively proceeds as (10)-(12) until its convergence is achieved. For a nonconvex problem, such as the one considered here, the convergence of ADMM is not guaranteed [9]. However, our simulations and those in other works such as [9], [12], [13], [16] demonstrate that ADMM converges well when the value of ρ is chosen to be appropriately large. We refer interested readers to [9, Sec.3] for the discussion on the choice of ρ.

B. The Reweighted ℓ_1 Minimization Method

The cardinality-aware optimization problem (7) is nonconvex and intractable in general. A common alternative is to replace the cardinality with the ℓ_1 norm, which results in a relaxed convex problem for (7). However, the use of ℓ_1 norm leads to the undesired dependence on the magnitude of elements in a vector compared to its cardinality [14]. Therefore, the authors in [14] proposed a method of reweighted ℓ_1 minimization, where the cardinality is substituted with a weighted ℓ_1 norm and the corresponding weights are redefined iteratively.

To be specific, we replace $\text{card}(\mathbf{w})$ in (7) with weighted ℓ_1 norm $\sum_{i=1}^N c_i|w_i|$, where c_i is the positive weight corresponding to the optimization variable w_i. Then the algorithm for reweighted ℓ_1 minimization can be stated as follows [14],

1) Set the iteration count $k = 0$ and set $c_i^0 = 1$ for $i = 1, \ldots, N$.
2) Solve the weighted ℓ_1 minimization problem

$$\begin{aligned} &\underset{\mathbf{w}^k}{\text{minimize}} && \tfrac{1}{2}\sigma^2(\mathbf{w}^k) + \gamma\textstyle\sum_{i=1}^N c_i^k|w_i^k| \\ &\text{subject to} && \textstyle\sum_{i=1}^N w_i^k = 1, \end{aligned} \tag{18}$$

whose solution is denoted by $\mathbf{w}^k = [w_1^k, \ldots, w_N^k]$.
3) Update the weights $c_i^{k+1} = \frac{1}{w_i^k + \nu}$, where the parameter $\nu > 0$ is used to ensure the validity of inversion of the zero-value component in $\mathbf{w}^k$.
4) Terminate if either k reaches a specified maximum number of iterations or the solution $\mathbf{w}^k$ has converged. Otherwise, increase k and go to Step 2.

In Step 2, the ℓ_1-norm based optimization problem (18) can be further cast as a quadratic program (QP) [15],

$$\begin{aligned} &\underset{\mathbf{w}^k, \mathbf{t}}{\text{minimize}} && \tfrac{1}{2}\sigma^2(\mathbf{w}^k) + \gamma\mathbf{1}^T\mathbf{t} \\ &\text{subject to} && \textstyle\sum_{i=1}^N w_i^k = 1 \\ &&& -t_i \le c_i^k w_i^k \le t_i,\ i = 1, 2, \ldots, N \end{aligned} \tag{19}$$

where $\sigma^2(\mathbf{w})$ is defined in (3), and $\mathbf{t} = [t_1, \ldots, t_N]^T$ is a new vector variable.

Remark 2: Note that the problem (18) can also be solved via ADMM. If we replace $\text{card}(\mathbf{v})$ with $\sum_{i=1}^N c_i|v_i|$ in (8), the steps of ADMM (10)-(12) will remain the same except the $\mathbf{v}$-minimization step (11) (i.e. (15)) which is given by

$$\underset{\boldsymbol{v}}{\text{minimize}}\, \psi(\boldsymbol{v}) = \gamma\sum_{i=1}^N c_i|v_i| + \frac{\rho}{2}\|\boldsymbol{v} - \boldsymbol{b}^k\|_2^2, \tag{20}$$

where $\mathbf{v}$, $\mathbf{b}^k$ have been defined in (15). Further, the analytical solution of (20) has been proposed in [12, Eq.14], i.e.,

$$v_i = \begin{cases} (1 - \frac{\gamma c_i}{\rho|b_i^k|})b_i^k & |b_i^k| > \frac{\gamma}{\rho}c_i \\ 0 & |b_i^k| \le \frac{\gamma}{\rho}c_i \end{cases} \tag{21}$$

where v_i and b_i^k are the ith elements of the vector $\mathbf{v}$ and $\mathbf{b}^k$, respectively.

As we relax the cardinality function to a weighted ℓ_1 norm, the resulting ADMM algorithm introduced in Remark 2 and the QP proposed in (19) are both guaranteed to converge due to the convexity of (18), but the converged solution is not the minimizer of the original problem (7). By contrast, the convergence of ADMM presented in Sec.III-A depends on the parameter ρ and the initial values of $\mathbf{w}$ and $\mathbf{v}$ [9]. Therefore, the reweighted ℓ_1-norm based methods are better as they do not require the initial guess of the optimization variables and algorithm parameters.

The ADMM-based algorithms typically take a few tens of iterations in many practical situations [9], [13]. At each iteration, the computational complexity of solving the $\mathbf{w}$-minimization problem is given by $O(N^{2.373})$, where N is the number of sensors, due to the calculation of the matrix inversion [17] in (14). For the $\mathbf{v}$-minimization problem, the analytical solution can be directly obtained from (17) or (21). Thus, the computational complexity of the ADMM-based approach is dominated by the cost of solving the $\mathbf{w}$-minimization problem, which is asymptotically towards $O(N^{2.373})$ for large N. For comparison, the QP-based algorithm roughly requires $O(N^{3.5})$ arithmetic operations by using a primal-dual interior point algorithm [18]. Therefore, the application of ADMM yields a lower computational complexity in large sensor networks.

We finally remark that estimating the value of random field at *multiple* unobserved locations, the Kriging error variance in the optimization problem (7) can be replaced with the trace of the Kriging error covariance, and our proposed approaches are still applicable. This extension is straightforward and omitted here for brevity.

IV. NUMERICAL EXAMPLES

In this section, we illustrate the utility of our proposed sparsity-promoting Kriging approach by considering the example shown in

Fig. 1. To specify the spatial correlation (1), we assume $\tau = 1$ and $\lambda = 0.1$. For ADMM, the algorithm parameters are selected as $\rho = 30$ and $\epsilon = 10^{-3}$, and the initial points are given by $\mathbf{v}^0 = \mathbf{1}/N$ and $\boldsymbol{\xi}^0 = \mathbf{0}$ in (10).

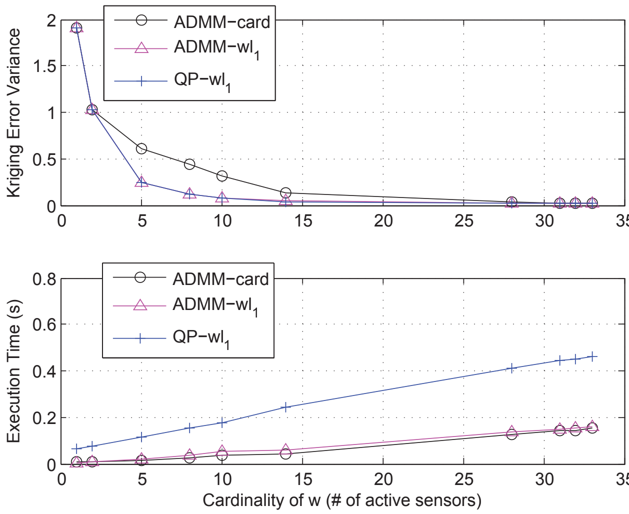


Fig. 2: Estimation performance and execution times versus the number of active sensors

In Fig. 2, we vary the sparsity-promoting parameter γ and obtain the trade-off curves between the estimation performance (in terms of KEV) and the number of selected sensors (in terms of the cardinality of $\mathbf{w}$). We also present the execution time[1] of our approaches versus the number of sensor activations. For simplicity of notation, we use “ADMM-card”, “QP-wl$_1$”, “ADMM-wl$_1$” to denote cardinality-aware ADMM, QP and ADMM with reweighted ℓ_1 norm, respectively. We observe that KEV ceases to improve significantly beyond the activation of 14 sensors, which means that only a subset of sensors can yield satisfactory estimation accuracy. Moreover, the use of reweighted ℓ_1 norm yields better estimation performance than ADMM-card for a fixed cardinality of $\mathbf{w}$. A possible reason is that ADMM-card yields a local optima whose accuracy relies on the selection of algorithm parameters and the initial point. Both of the ADMM-based methods require less computation time compared to quadratic progamming.

In Fig. 3, we present the KEV and the cardinality of $\mathbf{w}$ as a function of the magnitude of spatial correlation parameter λ for three different values of sparsity-promoting parameter $\gamma \in \{10^{-4}, 10^{-3}, 0.01\}$. It can be seen that the estimation error increases as λ increases. This is not surprising, since a larger value of λ indicates weaker spatial correlation which results in worse estimation performance [7]. Moreover, as the field becomes more spatially-correlated (λ decreases), less number of sensors are required to be activated for a given value of the sparsity-promoting parameter γ.

In Fig. 4, we demonstrate the optimal sparse sensor schedule (indicated by $\mathbf{w}$) obtained from the reweighted ℓ_1-norm based ADMM algorithm for different values of sparsity-promoting parameters $\gamma \in \{10^{-3}, 10^{-2}\}$. As we can see, a larger value of γ promotes a sparser sensor schedule. And the selected sensors tend to be spatially scattered over the field. The optimality of the obtained sensor schedules is also verified by comparing its estimation performance with that of a random selection strategy and that of exhaustive search which enumerates all possible activated sensors. The corresponding results

[1]The execution time is computed by MATLAB functions tic and toc on a 3 GHz computer

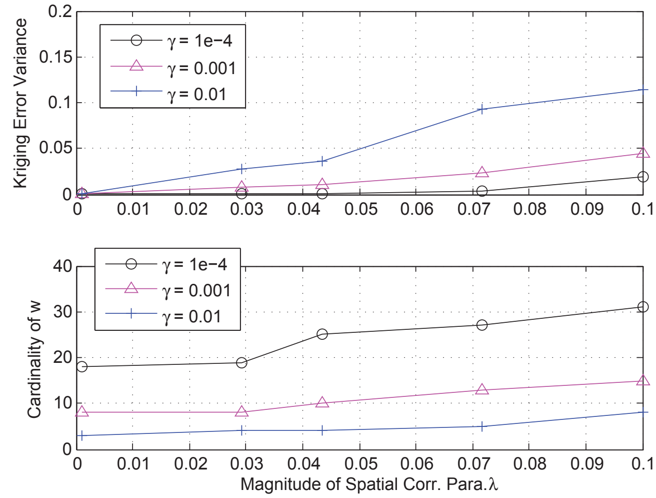
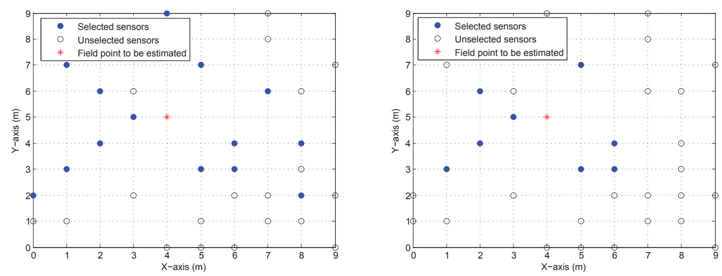


Fig. 3: Sparsity-promoting Kriging for different spatial correlation



(a) (b)

Fig. 4: Sensor selection via sparsity-promoting Kriging: (a) $\gamma = 10^{-3}$, (b) $\gamma = 10^{-2}$.

are omitted here for brevity. From another perspective, we may regard the positions (denoted by circles) as the potential locations where sensors can be deployed. Thus, our proposed sparsity-promoting Kriging also infers optimal sensor placement.

V. CONCLUSION

In this paper, the sensor selection problem for an ordinary Kriging estimator in field estimation was studied. We established a correspondence between active sensors and nonzero elements of Kriging weights and proposed a sparsity-promoting Kriging approach which minimizes the Kriging error variance while penalizing for the cardinality of Kriging weights. First, a cardinality-aware ADMM algorithm was proposed to obtain the locally optimal solutions of the original nonconvex problem. Next, we replaced the cardinality function with the weighted ℓ_1 norm and solved the relaxed problem with QP and ADMM, respectively. Numerical results show that the reweighted ℓ_1-norm based ADMM works well in terms of both optimization performance and computation time. While in this paper we assumed the field to be only correlated in space, our future work will study the problem of sensor selection for both spatially and temporally correlated fields.

ACKNOWLEDGMENT

This work was supported by U.S. Air Force Office of Scientific Research (AFOSR) under Grants FA9550-10-1-0263 and FA9550-10-1-0458, the National Science Foundation under award CMMI-0927509, and the Scientic and Technological Research Council of Turkey (TUBITAK) under Grant 113E220.

REFERENCES

[1] J. Wang, Y. Liu, and S. K. Das, "Energy-efficient data gathering in wireless sensor networks with asynchronous sampling," *ACM Transactions on Sensor Networks*, vol. 6, no. 22, pp. 1–37, June 2010.

[2] N. Cressie, *Statistics for Spatial Data*, John Wiley Sons, Inc., NY, 1991.

[3] N. Cressie, "Spatial prediction and ordinary Kriging," *Mathematical Geology*, vol. 20, no. 4, pp. 1405–421, May 1988.

[4] A. Boufassa and M. Armstrong, "Comparison between different Kriging estimators," *Mathematical Geology*, vol. 21, no. 3, pp. 331–345, 1989.

[5] H. Zhang, J. Moura, and B. Krogh, "Dynamic field estimation using wireless sensor networks: Tradeoffs between estimation error and communication cost," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2383–2395, June 2009.

[6] Y. Mo, R. Ambrosino, and B. Sinopoli, "Sensor selection strategies for state estimation in energy constrained wireless sensor networks," *Automatica*, vol. 47, no. 7, pp. 1330–1338, 2011.

[7] S. Liu, E. Masazade, and P. K. Varshney, "Temporally staggered sensing for field estimation with quantized data in wireless sensor networks," in *IEEE Statistical Signal Processing Workshop (SSP)*, Aug. 2012, pp. 512–515.

[8] Y. Weng, L. Xie, and W. Xiao, "Random field estimation with quantized measurements in sensor networks," in *Proceedings of the 29th Chinese Control Conference*, July 2010, pp. 6203–6208.

[9] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[10] Fu Lin, M. Fardad, and M. R. Jovanović, "Algorithms for leader selection in large dynamical networks: Noise-corrupted leaders," in *Proceedings of the 50th IEEE Conference on Decision and Control*, Dec. 2011, pp. 2932–2937.

[11] F. Lin, M. Fardad, and M. R. Jovanović, "Augmented Lagrangian approach to design of structured optimal state feedback gains," *IEEE Trans. Autom. Control*, vol. 56, no. 12, pp. 2923–2929, Dec. 2011.

[12] F. Lin, M. Fardad, and M. R. Jovanović, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Trans. Autom. Control*, vol. 58, pp. 2426–2431, 2013.

[13] E. Masazade, M. Fardad, and P. K. Varshney, "Sparsity-promoting extended Kalman filtering for target tracking in wireless sensor networks," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 845–848, Dec. 2012.

[14] E. Candès, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[16] S. Liu, M. Fardad, E. Masazade, and P. K. Varshney, "Optimal periodic sensor scheduling in large-scale dynamical networks," *http://arxiv.org/abs/1305.5601*, 2013.

[17] V. V. Williams, "Multiplying matrices faster than Coppersmith-Winograd," in *Proceedings of the 44th symposium on Theory of Computing*, 2012, pp. 887–898.

[18] A. Nemirovski, "Interior point polynomial time methods in convex programming," 2012 [Online], Available: *http://www2.isye.gatech.edu/ nemirovs/Lect_IPM.pdf*.