18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Sparsity-Promoting Sensor Management for Estimation: An Energy Balance Point of View

Sijia Liu*, Feishe Chen†, Aditya Vempaty*, Makan Fardad*, Lixin Shen† and Pramod K. Varshney*
*Department of Electrical Engineering and Computer Science
†Department of Mathematics
Syracuse University, Syracuse, NY 13244 USA
Email: {sliu17, fchen05, avempaty, makan, lshen03, varshney}@syr.edu

***Abstract*—In the context of parameter estimation, we study the problem of sensor management under a sparsity-promoting framework, where a sensor being off at a certain time instant is represented by the corresponding column of the estimator coefficient matrix being identically zero. In order to achieve a balance between activating the most informative sensors and uniformly allocating sensor energy, we propose a novel sparsity-promoting approach by adding an ℓ_2-norm penalty function that discourages successive selections of the same group of sensors. We employ the alternating direction method of multipliers (ADMM) to solve the resulting ℓ_2-norm optimization problem, which can then be split into a sequence of analytically solvable subproblems. We finally provide numerical results and comparison with other sensor scheduling algorithms in the literature to demonstrate the effectiveness of our approach.**

I. Introduction

Sensor management is an important problem in applications of sensor networks such as parameter estimation, detection, and optimal control [1]–[3]. Due to the constraints on communication bandwidth and sensor battery life, it may not be desirable to have all sensors report their measurements at all time instants. Therefore, the problem of sensor management arises, which seeks to activate different subsets of sensors at different times in order to attain an optimal tradeoff between energy use and estimation or detection accuracy.

In this paper, we focus on the problem of sensor management for parameter estimation. Recently, this problem has been studied under a sparsity-promoting framework [4]–[7]. In [4], the problem of sensor selection was addressed by seeking the optimal sparse estimator gain matrix, where a one-to-one correspondence between active sensors and the nonzero columns of the estimator gain matrix was first introduced. As an extension of [4], a more general sparsity-promoting framework in sensor management was proposed in [5] for finding optimal periodic sensor schedules subject to measurement frequency constraints. In [6], the design of sensor selection scheme was transformed to the recovery of a sparse matrix. In [7], a sparsity-aware sensor selection problem was formulated where the number of selected sensors was minimized subject to a certain estimation quality.

In the current literature [4]–[7], the existing sparsity-promoting techniques may lead to scenarios in which a fixed set of sensors, which we hereafter refer to as the most *informative sensors*, are successively selected. The behavior of successive selections of the most informative sensors could be either due to their larger mutual information with the target [8] or due to stronger correlation with the field point of interest [9]. Successive selections would result in faster energy depletion of these sensors, which may render the network nonfunctional [10].

To address the imbalance of energy usage we introduce a new sparsity-promoting penalty function that discourages successive selections of the same group of sensors. This idea has been exploited in our earlier work [11], where with the help of reweighted ℓ_1 minimization method the proposed sparsity-promoting sensor management problem was solved via quadratic programming. Different from [11], our main contributions are threefold.

- Instead of using a weighted ℓ_1-norm, we use the standard unweighted ℓ_2-norm to characterize the column-sparsity of estimator coefficient matrices. Use of the ℓ_2-norm leads to a more efficient optimization framework.
- We show that the proposed ℓ_2-norm optimization problem is convex, and its solution can be efficiently found via alternating direction method of multipliers (ADMM).
- We present a comparison of both the performance and the computational complexity of our methodology with the work reported in [11]. We empirically show that our approach yields the estimation performance that is close to that of the method in [11], but has lower computational complexity.

II. Motivation: Example of Field Estimation

In this section, we begin with a motivating example of field estimation to review the conventional sparsity-promoting optimization framework for sensor management. We show that the conventional method may cause an imbalance in energy usage when the sparsity of sensor activations is promoted. To address this issue, we seek an approach that discourages the excessive use of the fixed group of sensors in a network.

Consider a generic system where multiple sensors are deployed to monitor a spatio-temporally correlated random field. Measurements from these multiple sensors at different time instants are used to estimate the field intensity at an unobserved location over multiple unobserved time instants. Let $f(\mathbf{s}, t)$ denote the field intensity at location $\mathbf{s}$ and time instant t, and with mean μ_f and variance σ_f^2. For simplicity of representation, we stack all the sensor measurements into a single vector, and the measurement of the mth sensor at the sampling time t_k, denoted by $y_{k+(m-1)K}$, is given by

$$y_{k+(m-1)K} = f(\mathbf{s}_m, t_k) + v_{k+(m-1)K}, \quad (1)$$

for $k = 1, 2, \ldots, K$ and $m = 1, 2, \ldots, M$, where K is the number of samples per sensor, M is the number of sensors, $\mathbf{s}_m$ denotes the location of the mth sensor, and $v_{k+(m-1)K}$ is a zero-mean Gaussian noise with variance σ_v^2. The vector of field intensities to be estimated is given by

$$\mathbf{x} = [f(\varsigma, \tau_1), f(\varsigma, \tau_2), \ldots, f(\varsigma, \tau_N)]^T, \quad (2)$$

where ς is the location where the field intensity is to be estimated at time instants $\{\tau_n\}_{n=1,2,\ldots,N}$. An example of field estimation is presented in Fig. 1.

To perform the estimation task, consider an unbiased linear estimator

$$\hat{\mathbf{x}} = \mathbf{W}\mathbf{y} + \mathbf{a} = \sum_{m=1}^{M}\sum_{k=1}^{K} y_{k+(m-1)K}\mathbf{W}_{k+(m-1)K} + \mathbf{a}, \quad (3)$$

where $\mathbf{W} \in \mathbb{R}^{N \times KM}$ is an unknown coefficient matrix determined by the minimum mean square error criterion, $\mathbf{W}_j \in \mathbb{R}^N$ denotes the jth column of $\mathbf{W}$, $\mathbf{a} = \mu_f(\mathbf{I} - \mathbf{W})\mathbf{1}$ is a vector to ensure unbiasedness, $\mathbf{I}$ and $\mathbf{1}$ denote the identity matrix and the all-ones vector, respectively, and $\mathbf{y} = [y_1, y_2, \ldots, y_{KM}]^T$.

The trace of estimation error covariance matrix for the estimate in (3) is given by [11]

$$J(\mathbf{W}) = \bar{\mathbf{w}}^T\mathbf{P}\bar{\mathbf{w}} - 2\mathbf{q}^T\bar{\mathbf{w}} + N\sigma_f^2, \quad (4)$$

where $\bar{\mathbf{w}} \in \mathbb{R}^{KMN}$ denotes the rowwise vector of $\mathbf{W}$, $\mathbf{P} = \mathbf{I}_{N\times N} \otimes [\mathrm{Cov}(\mathbf{z}, \mathbf{z}) + \sigma_v^2\mathbf{I}]$, $\otimes$ denotes the Kronecker product, $\mathbf{z} = [f(\mathbf{s}_1, t_1), f(\mathbf{s}_1, t_2), \ldots, f(\mathbf{s}_M, t_K)]^T \in \mathbb{R}^{KM}$, $\mathrm{Cov}(\cdot, \cdot)$ represents the covariance matrix of two random vectors, and $\mathbf{q} = [\mathrm{Cov}(x_1, \mathbf{z}), \mathrm{Cov}(x_2, \mathbf{z}), \ldots, \mathrm{Cov}(x_N, \mathbf{z})]^T$.

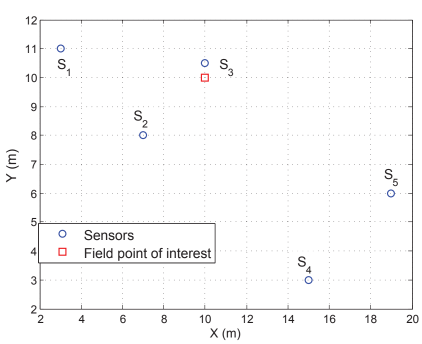

Fig. 1: A random field with $M = 5$ sensors and field intensity is estimated at the point of interest $\varsigma = (10, 10)$.

It is clear from (3) that the non-zero columns of $\mathbf{W}$ characterize the selected sensor measurements. For example, if only the ith sensor reports a measurement at time t_j. It follows from (3) that $\mathbf{W}\mathbf{y} = y_{j+(i-1)K}\mathbf{W}_{j+(i-1)K}$. To seek an optimal tradeoff between estimation accuracy and sensor activations, a sparsity-promoting optimization problem has been proposed in [7], [9], [12]

$$\underset{\mathbf{W}}{\text{minimize}} \quad J(\mathbf{W}) + \gamma h(\mathbf{W}). \quad (5)$$

In (5), $J(\mathbf{W})$ is the trace of estimation error covariance, $\gamma > 0$ is a sparsity-promoting parameter since sparser sensor schedules can be achieved by making γ larger, and $h(\mathbf{W})$ gives the number of nonzero columns of $\mathbf{W}$

$$h(\mathbf{W}) = \sum_{m=1}^{M}\sum_{k=1}^{K} \mathrm{card}(\|\mathbf{W}_{k+(m-1)K}\|_p), \quad (6)$$

where $\|\cdot\|_p$ denotes an arbitrary ℓ_p-norm ($p = 1$ or 2 typically), and the cardinality of a scalar $x \in \mathbb{R}$ is defined as

$$\mathrm{card}(x) = \begin{cases} 1 & x \neq 0 \\ 0 & x = 0. \end{cases}$$

We remark that different choices of ℓ_p-norm will lead to different convex relaxations of the cardinality function in (6). It has been suggested in [9] and [13] that ADMM and proximal algorithms are suitable tools to solve the nonconvex problem (5). As will be evident later, our framework is a generalization of (5).

Drawback of conventional formulation: The total number of selected sensors in the solution of (5) decreases when γ increases, thereby promoting sparsity in sensor activations. However, there exist scenarios in which a fixed subset of sensors (the most informative ones) will be successively selected and thus, result in unbalanced energy usage among sensors. We demonstrate this phenomenon in Fig. 3 of Sec. V. Motivated by the unbalanced energy usage, we will present a new sparsity-promoting framework for sensor management, which achieves a balance between activating the most informative sensors and uniformly allocating sensor energy over the entire network.

III. Problem Statement

In this section, we formally state the problem addressed in this paper. We introduce a quadratic function, with respect to the number of times that each sensor is selected over a time horizon, to penalize the overuse of the same group of sensors. By relaxing the cardinality function, we propose an ℓ_2-norm optimization problem to find sparse sensor schedules in an energy-balanced manner.

Let κ_m denote the number of times the mth sensor is selected over K time steps. We then use the quadratic function κ_m^2 to characterize the *cost* of each sensor being repeatedly selected due to its relatively fast growth as a function of κ_m. For instance given a system with $M = 2$ sensors, it is clear from (7) that the penalty value of using the first sensor 4 times and the second sensor 0 times ($4^2 + 0^2 = 16$ units) is greater than the penalty of using each sensor 2 times ($2^2 + 2^2 = 8$ units). Based on this motivation, we now formalize the new sparsity-promoting function that penalizes successive selections of each sensor

$$g(\mathbf{W}) = \sum_{m=1}^{M}\kappa_m^2, \;\; \kappa_m = \sum_{k=1}^{K}\mathrm{card}(\|\mathbf{W}_{k+(m-1)K}\|_p). \quad (7)$$

Using (7), we modify the conventional formulation (5) as

$$\underset{\mathbf{W}}{\text{minimize}} \quad J(\mathbf{W}) + \gamma h(\mathbf{W}) + \eta g(\mathbf{W}), \tag{8}$$

where $\gamma, \eta > 0$ are regularization parameters. It is worth mentioning that problem (8) is not restricted to the application of field estimation discussed in Sec. II, but suitable for a broad class of sensor scheduling problems where a *linear estimator* is used.

In problem (8), sparser sensor schedules can be achieved by making either γ or η larger. For large γ, the resulting sparse sensor schedule may contain less total number of active sensors, but some sensors will be selected more frequently. Conversely, the sparse sensor schedule resulting from large η may include more number of total active sensors, but the set of selected sensors is more diverse, which leads to the balanced usage of sensor energy in the network. When $\eta = 0$, problem (8) reduces to the conventional problem (5).

Note that the trace of estimation error covariance defined in (4) is a convex quadratic function with respect to the rowwise vector of $\mathbf{W}$. In contrast, the sparsity-promoting penalties h and g are functions of the columnwise vector of $\mathbf{W}$. Lemma 1 shows an association between the rowwise vector $\bar{\mathbf{w}}$ and columnwise vector $\mathbf{w}$ of $\mathbf{W}$, respectively.

Lemma 1: Consider a matrix $\mathbf{W} \in \mathbb{R}^{K \times MN}$, the rowwise vector $\bar{\mathbf{w}}$ and columnwise vector $\mathbf{w}$ of $\mathbf{W}$ satisfy

$$\mathbf{w} = \mathbf{D}\bar{\mathbf{w}}. \tag{9}$$

In (9),

$$\mathbf{D} = [\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \cdots, \mathbf{e}_{j_{KMN}}]^T,$$

where $\mathbf{e}_{j_l} \in \mathbb{R}^{KMN}$ denotes the basis vector with a 1 in the j_lth coordinate and 0s elsewhere, and the value j_l is given by $j_l = (n_l - 1)KM + (m_l - 1)K + k_l$, with $m_l = \left\lfloor \frac{l-1}{KN} \right\rfloor + 1$, $k_l = \left\lfloor \frac{l-(m_l-1)KN-1}{N} \right\rfloor + 1$, $n_l = l-(m_l-1)KN-(k_l-1)N$, and $\lfloor \cdot \rfloor$ maps a real number to the largest integer smaller than itself.

Proof: The proof is straightforward and the details are omitted for the sake of brevity. ■

Using Lemma 1, problem (8) can be rewritten as an optimization problem with respect to $\mathbf{w}$

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \; & J(\mathbf{w}) + \gamma \sum_{m=1}^{M} \sum_{k=1}^{K} \text{card}(\|\mathbf{w}_{k,m}\|_p) \\ & + \eta \sum_{m=1}^{M} \left(\sum_{k=1}^{K} \text{card}(\|\mathbf{w}_{k,m}\|_p) \right)^2, \end{aligned} \tag{10}$$

where with an abuse of notation, $J(\mathbf{w})$ is the mean square error as a function of the columnwise vector of $\mathbf{W}$, and $\mathbf{w}_{k,m}$ denotes the $(k + mK - K)$th subvector of $\mathbf{w}$ which corresponds to $\mathbf{w}_{k,m} = \mathbf{W}_{k+(m-1)K}$ for $k = 1, 2, \ldots, K$ and $m = 1, 2, \ldots, M$.

Problem (10) is not convex due to the presence of the cardinality function. For tractability, we replace the cardinality function with an ℓ_1-norm [14], [15],

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \; & J(\mathbf{w}) + \gamma \sum_{m=1}^{M} \sum_{k=1}^{K} \|\mathbf{w}_{k,m}\|_p \\ & + \eta \sum_{m=1}^{M} \left(\sum_{k=1}^{K} \|\mathbf{w}_{k,m}\|_p \right)^2. \end{aligned} \tag{11}$$

In problem (11), different choices of ℓ_p-norm will lead to different approximations of problem (10). When $p = 1$, we use an ℓ_1-norm to characterize the column-sparsity of $\mathbf{W}$. The problem (11) with $p = 1$ has been addressed in our earlier work [11], where we have shown that the ℓ_1-based (11) can be transformed into a convex quadratic program. When $p = 2$, an ℓ_2-norm is used to characterize the column-sparsity of $\mathbf{W}$ in (11). Recall that the vector $\mathbf{w}_{k,m}$ is a subvector of the columnwise vector of $\mathbf{W}$ and therefore, the column-sparsity of $\mathbf{W}$ is precisely the group-sparsity of $\mathbf{w}$. In this sense, the use of ℓ_2-norm in (11) is well motivated by the problem of group Lasso [16], which uses ℓ_2-norm to promote the group-sparsity of a vector.

As the major contribution of this paper, we will solve problem (11) with $p = 2$. Namely,

$$\underset{\mathbf{w}}{\text{minimize}} \; J(\mathbf{w}) + \gamma \sum_{m=1}^{M} \sum_{k=1}^{K} h_{k,m}(\mathbf{w}) + \eta \sum_{m=1}^{M} g_m(\mathbf{w}), \tag{12}$$

where

$$h_{k,m}(\mathbf{w}) := \|\mathbf{w}_{k,m}\|_2, \tag{13}$$

and

$$g_m(\mathbf{w}) := \left(\textstyle\sum_{k=1}^{K} \|\mathbf{w}_{k,m}\|_2 \right)^2. \tag{14}$$

IV. Optimal Solutions via ADMM

In this section, we first prove that the proposed new penalty function $g_m(\mathbf{w})$ is a convex function and thus problem (12) is a convex optimization problem.

Theorem 1: Problem (12) is a convex optimization problem.

Proof: See Appendix A. ■

We now develop an ADMM-based algorithm to find the optimal solution of problem (12). It has been shown in [5], [17], [18] that ADMM is well-suited for problems that involve (nonsmooth) sparsity-inducing regularizers such as ℓ_1 and ℓ_2 norms. The major advantage of ADMM is that it allows us to split the optimization problem (12) into a sequence of subproblems, each of which can be solved analytically. We will elaborate on ADMM in what follows.

We begin by reformulating the optimization problem (12) in a way that lends itself to the application of ADMM,

$$\begin{aligned} \underset{\mathbf{w},\mathbf{v},\mathbf{u}}{\text{minimize}} \quad & J(\mathbf{w}) + \gamma \sum_{m=1}^{M} \sum_{k=1}^{K} h_{k,m}(\mathbf{v}) + \eta \sum_{m=1}^{M} g_m(\mathbf{u}) \\ \text{subject to} \quad & \mathbf{w} = \mathbf{v}, \; \mathbf{w} = \mathbf{u}. \end{aligned} \tag{15}$$

The augmented Lagrangian of (15) is given by

$$\mathcal{L}(\mathbf{w},\mathbf{v},\mathbf{u},\boldsymbol{\mu},\boldsymbol{\xi}) = J(\mathbf{w}) + \gamma\sum_{m=1}^{M}\sum_{k=1}^{K} h_{k,m}(\mathbf{v}) + \eta\sum_{m=1}^{M} g_m(\mathbf{u}) + \boldsymbol{\mu}^T(\mathbf{w}-\mathbf{v}) + \boldsymbol{\eta}^T(\mathbf{w}-\mathbf{u}) + \frac{\rho}{2}\|\mathbf{w}-\mathbf{v}\|_2^2 + \frac{\rho}{2}\|\mathbf{w}-\mathbf{u}\|_2^2,$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\xi}$ are Lagrangian multipliers, and $\rho > 0$ is a penalty weight.

The ADMM algorithm iteratively executes the following three steps [17] for $l = 1, 2, \ldots$

$$\mathbf{w}^{l+1} = \arg\min_{\mathbf{w}} \left\{ J(\mathbf{w}) + \frac{\rho}{2}\|\mathbf{w}-\mathbf{a}_1^l\|_2^2 + \frac{\rho}{2}\|\mathbf{w}-\mathbf{a}_2^l\|_2^2 \right\}, \tag{16}$$

$$\{\mathbf{v}^{l+1},\mathbf{u}^{l+1}\} = \arg\min_{\mathbf{v},\mathbf{u}} \left\{ \gamma\sum_{m=1}^{M}\sum_{k=1}^{K} h_{k,m}(\mathbf{v}) + \frac{\rho}{2}\|\mathbf{v}-\mathbf{b}^l\|_2^2 + \eta\sum_{m=1}^{M} g_m(\mathbf{u}) + \frac{\rho}{2}\|\mathbf{u}-\mathbf{c}^l\|_2^2 \right\}, \tag{17}$$

$$\boldsymbol{\mu}^{l+1} = \boldsymbol{\mu}^l + \rho(\mathbf{w}^{l+1} - \mathbf{v}^{l+1}),\ \boldsymbol{\xi}^{l+1} = \boldsymbol{\xi}^l + \rho(\mathbf{w}^{l+1} - \mathbf{u}^{l+1}), \tag{18}$$

until $\|\mathbf{w}^{l+1}-\mathbf{v}^{l+1}\|_2^2 + \|\mathbf{w}^{l+1}-\mathbf{u}^{l+1}\|_2^2 \leq \epsilon$ and $\|\mathbf{v}^{l+1}-\mathbf{v}^l\|_2^2 + \|\mathbf{u}^{l+1}-\mathbf{u}^l\|_2^2 \leq \epsilon$, where $\mathbf{a}_1^l = \mathbf{v}^l - \frac{1}{\rho}\boldsymbol{\mu}^l$, $\mathbf{a}_2^l = \mathbf{v}^l - \frac{1}{\rho}\boldsymbol{\xi}^l$, $\mathbf{b}^l = \mathbf{w}^{l+1} + \frac{1}{\rho}\boldsymbol{\mu}^l$, $\mathbf{c}^l = \mathbf{w}^{l+1} + \frac{1}{\rho}\boldsymbol{\xi}^t$, and ϵ is a stopping tolerance.

We note that the ADMM step (17) can be further decomposed into two subproblems,

$$\mathbf{v}^{l+1} = \arg\min_{\mathbf{v}} \left\{ \gamma\sum_{m=1}^{M}\sum_{k=1}^{K} h_{k,m}(\mathbf{v}) + \frac{\rho}{2}\|\mathbf{v}-\mathbf{b}^l\|_2^2 \right\}, \tag{19}$$

and

$$\mathbf{u}^{l+1} = \arg\min_{\mathbf{v}} \left\{ \eta\sum_{m=1}^{M} g_m(\mathbf{u}) + \frac{\rho}{2}\|\mathbf{u}-\mathbf{c}^l\|_2^2 \right\}. \tag{20}$$

In the subsections that follow, we will derive the analytical solutions of '$\mathbf{w}$-minimization' problem (16), '$\mathbf{v}$-minimization' problem (19), and '$\mathbf{u}$-minimization' problem (20).

A. $\mathbf{w}$-*minimization problem*

Using (4) and (9), the $\mathbf{w}$-minimization problem (16) can be rewritten as

$$\underset{\mathbf{w}}{\text{minimize}} \quad \tfrac{1}{2}\mathbf{w}^T\mathbf{H}\mathbf{w} - \left[2(\mathbf{D}^{-1})^T\mathbf{q} + \rho\mathbf{a}_1^l + \rho\mathbf{a}_2^l\right]^T\mathbf{w},$$

where $\mathbf{H} = 2(\mathbf{D}^{-1})^T\mathbf{P}\mathbf{D}^{-1} + 2\rho\mathbf{I}$, and the matrix $\mathbf{D}$ is invertible as shown by its definition in (9). According to (4), we also obtain $\mathbf{P} \succ 0$ and thus $\mathbf{H} \succ 0$, where $\mathbf{X} \succ 0$ indicates that $\mathbf{X}$ is a positive definite matrix. The minimizer of problem (16) is then given by

$$\mathbf{w}^{l+1} = \mathbf{H}^{-1}\left[2(\mathbf{D}^{-1})^T\mathbf{q} + \rho\mathbf{a}_1^l + \rho\mathbf{a}_2^l\right]. \tag{21}$$

B. $\mathbf{v}$-*minimization problem*

Using (13), the $\mathbf{v}$-minimization problem (19) becomes

$$\underset{\mathbf{v}}{\text{minimize}} \quad \sum_{m=1}^{M}\sum_{k=1}^{K}\|\mathbf{v}_{k,m}\|_2 + \frac{1}{2\hat{\gamma}}\|\mathbf{v}-\mathbf{b}^l\|_2^2,$$

where $\mathbf{v}_{k,m}$ is the $(k+mK-K)$th subvector of $\mathbf{v}$, and $\hat{\gamma} = \gamma/\rho$. The solution of problem (19) is then given by a block soft thresholding operator [13]

$$\mathbf{v}_{k,m}^{l+1} = \begin{cases} \left(1 + \frac{\hat{\gamma}}{\|\mathbf{b}_{k,m}^l\|_2}\right)\mathbf{b}_{k,m}^l & \|\mathbf{b}_{k,m}^l\|_2 \geq \hat{\gamma} \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{22}$$

for $k = 1, 2, \ldots, K$ and $m = 1, 2, \ldots, M$, where $\mathbf{b}_{k,m}^l$ is the $(k+mK-K)$th subvector of $\mathbf{b}^l$.

C. $\mathbf{u}$-*minimization problem*

Using (14), the $\mathbf{u}$-minimization problem (20) is

$$\underset{\mathbf{u}}{\text{minimize}} \quad \eta\sum_{m=1}^{M}\left(\sum_{k=1}^{K}\|\mathbf{u}_{k,m}\|_2\right)^2 + \frac{\rho}{2}\|\mathbf{u}-\mathbf{c}^l\|_2^2, \tag{23}$$

which can be decomposed into M subproblems

$$\underset{\{\mathbf{u}_{k,m}\}_{k=1,2,\ldots,K}}{\text{minimize}} \quad \left(\sum_{k=1}^{K}\|\mathbf{u}_{k,m}\|_2\right)^2 + \frac{1}{2\hat{\eta}}\sum_{k=1}^{K}\|\mathbf{u}_{k,m}-\mathbf{c}_{k,m}\|_2^2 \tag{24}$$

for $m = 1, 2, \ldots, M$, where $\hat{\eta} = \eta/\rho$, and we use $\mathbf{c}_{k,m}$ instead of $\mathbf{c}_{k,m}^l$ for ease of notation.

We remark that the quadratic function with respect to the sum of ℓ_2 norms makes finding the minimizer of (24) more challenging, since it is a nonlinear composition of nonsmooth functions. However, Proposition 1 will demonstrate that the solution of (24) is achievable via quadratic programming (QP).

Proposition 1: The minimizer of problem (24) is given by

$$\mathbf{u}_{k,m}^* = r_k^*\frac{\mathbf{c}_{k,m}^T}{\|\mathbf{c}_{k,m}^T\|_2}, \quad k = 1, 2, \ldots, K. \tag{25}$$

In (25), $\mathbf{r}^* := [r_1^*, r_2^*, \ldots, r_K^*]^T$ is the minimizer of QP

$$\begin{array}{ll} \underset{\mathbf{r}}{\text{minimize}} & \mathbf{r}^T\mathbf{1}\mathbf{1}^T\mathbf{r} + \frac{1}{2\hat{\eta}}\|\mathbf{r}-\mathbf{f}\|_2^2, \\ \text{subject to} & \mathbf{r} \geq \mathbf{0}, \end{array} \tag{26}$$

where $\mathbf{f} = [\|\mathbf{c}_{1,m}^T\|_2, \|\mathbf{c}_{2,m}^T\|_2, \ldots, \|\mathbf{c}_{K,m}^T\|_2]^T$.

Proof: See Appendix B. ■

It is clear from (23) and Proposition 1 that the solution of the $\mathbf{u}$-minimization problem can be found through M QPs, each of which has complexity $O(K^{3.5})$ [19]. However, instead of using QP, we will show that the analytical solution of (26) is also tractable. We first present an important feature of problem (26) in Lemma 2.

Lemma 2: If the entries of $\mathbf{f}$ satisfy $f_1 \leq f_2 \leq \ldots \leq f_K$, then the solution of (26) yields $r_1^* \leq r_2^* \leq \ldots \leq r_K^*$.

Proof: The proof is straightforward, proceeding by contradiction. Details are omitted for the sake of brevity. ■

The rationale behind exploiting the result in Lemma 2 is that the analytical solution of (26) is tractable by sorting $\mathbf{f}$ in an ascending order. Let $\mathcal{I}$ denote the index set that describes the rearrangement of $\{f_i\}_{i=1,2,\ldots,K}$ in an ascending order. And let $\mathbf{r}^*_{\mathcal{I}}$ be the solution of (26) rearranged by $\mathcal{I}$. For instance, if $\mathbf{f} = [5,4]^T$, then $\mathcal{I} = \{2,1\}$ and $\mathbf{r}^*_{\mathcal{I}} = [r^*_2, r^*_1]^T$. The closed-form of $\mathbf{r}^*_{\mathcal{I}}$ is given in Proposition 2. Together with $\mathcal{I}$, we then obtain the solution of (26), namely, $\mathbf{r}^*$.

Proposition 2: The minimizer of problem (26), with an index rearrangement $\mathcal{I}$, is given by

$$r^*_{\mathcal{I},i} = \begin{cases} 0 & 1 \le i \le \iota - 1,\ i \in \mathbb{N} \\ f_i - \dfrac{2\hat{\eta}\sum_{k=\iota}^{K} f_k}{1 + 2\hat{\eta}(K - \iota + 1)} & \iota \le i \le K,\ i \in \mathbb{N}, \end{cases} \tag{27}$$

for $i = 1, 2, \ldots, K$, where $\mathbb{N}$ denotes the set of natural numbers, $\mathcal{I}$ is the index set that describes the rearrangement of $\{f_i\}_{i=1,2,\ldots,K}$ in an ascending order, f_i is the ith entry of $\mathbf{f}$ given in (26), ι is the index of the first positive element in the set of numbers $\left\{ f_i - \frac{2\hat{\eta}\sum_{k=i}^{K} f_k}{1+2\hat{\eta}(K-i+1)} \right\}_{i=1,2,\ldots,K}$ and $\iota = K+1$ if no positive element exists.

Proof: See Appendix C. ■

According to Proposition 2, the complexity of solving problem (26) is given by $O(K \log K)$ owing to the sorting of the entries of $\mathbf{f}$ in an ascending order. This is much lower than the complexity of QP with $O(K^{3.5})$.

D. Summary of ADMM and complexity analysis

The ADMM algorithm for solving (12) is summarized in the following pseudocode format.

Algorithm 1 ADMM for solving (12)

Require: Choose ρ and ϵ. Initialize ADMM with $\mathbf{w}^0 = \mathbf{v}^0 = \mathbf{u}^0 = \mathbf{1}$, $\boldsymbol{\mu} = \boldsymbol{\xi} = \mathbf{0}$.

1: **for** $l = 0, 1, \ldots$ **do**
2: Obtain $\mathbf{w}^{l+1}$ from (21).
3: Obtain $\mathbf{v}^{l+1}$ from (22).
4: Obtain $\mathbf{u}^{l+1}$ from Propositions 1 and 2.
5: if $\|\mathbf{w}^{l+1} - \mathbf{v}^{l+1}\|_2^2 + \|\mathbf{w}^{l+1} - \mathbf{u}^{l+1}\|_2^2 \le \epsilon$ and $\|\mathbf{v}^{l+1} - \mathbf{v}^{l}\|_2^2 + \|\mathbf{u}^{l+1} - \mathbf{u}^{l}\|_2^2 \le \epsilon$, **quit**.
6: **end for**

Typically, ADMM takes a few tens of iterations to converge with modest accuracy [17], [18]. In our implementation, it typically takes $O(L^{0.5})$ iterations when $\epsilon = 10^{-3}$, where $L = KMN$. At each iteration of ADMM, the major computation cost lies in Step 2 of Algorithm 1 due to the matrix inversion, with complexity $O(L^{2.373})$ [20]. Thus, the total computational complexity of Algorithm 1 is approximated by $O\left(L^{2.873}\right)$.

For additional perspective, we compare the computational complexity of Algorithm 1 to the existing method in [11], where a reweighted ℓ_1 based QP was used to obtain the energy-balanced sensor schedules. The computational complexity of the method in [11] can be approximated by $O(L^4)$, where the reweighted ℓ_1 method takes $O(L^{0.5})$ iterations for convergence, and the complexity of QP is $O(L^{3.5})$. It can be seen that Algorithm 1 reduces the computational complexity by a factor of $O(L^{1.127})$.

V. Numerical Results

In this section, we consider a sensor network with $M \in \{5, 10\}$ sensors on a 20×20 grid, where each sensor takes $K = 10$ measurements at the sampling time $0.2k$ for $k = 1, 2, \ldots, K$. The task of the sensor network is to reconstruct the field intensities at the coordinate $\boldsymbol{\varsigma} = (10, 10)$ over time slots $\{\tau_n\}_{n=1,2,\ldots,N}$, where $\tau_n = 0.2n + 0.1$ and $N = 9$. When $M = 5$, the deployment of sensors and the field point to be estimated is shown in Fig. 1. We assume that the correlation model of the random field is given by $\mathrm{Cov}(f(\mathbf{s},t), f(\mathbf{s}',t')) = \sigma_f^2 \exp\{-c_s\|\mathbf{s} - \mathbf{s}'\|_2 - c_t(t-t')^2\}$, where $\sigma_f^2 = 1$, $c_s = 0.1$, and $c_t = 0.1$. In Algorithm 1, we select the ADMM parameters as $\rho = 100$ and $\epsilon = 10^{-3}$.

In what follows, we define a factor to measure the energy imbalance in the usage of each sensor due to successive selections. Let $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots, \omega_M] \in \mathbb{N}^M$ denote the sensor activation scheme of M sensors over K time steps, where $\omega_i \in \{0, 1, \ldots, K\}$ represents the number of times the ith sensor is selected over K time steps, and $\mathbf{1}^T\boldsymbol{\omega}$ gives the total number of sensor activations. For simplicity, we consider a specific sensor schedule that satisfies $\omega_1 \ge \omega_2 \ge \ldots \ge \omega_M$. From the perspective of energy balance, we expect to select sensors as uniformly as possible over K time steps. The most balanced schedule is given by $\tilde{\boldsymbol{\omega}} := [\omega_0 + \tilde{a}_1, \omega_0 + \tilde{a}_2, \ldots, \omega_0 + \tilde{a}_M]^T$, where $\omega_0 = \left\lfloor \frac{\mathbf{1}^T\boldsymbol{\omega}}{M} \right\rfloor$ yields the maximum number of times each sensor is uniformly scheduled, and $\tilde{a}_i$ is equal to 1 for $i = 1, 2, \ldots, (\mathbf{1}^T\boldsymbol{\omega} - M\omega_0)$, and 0 otherwise. Given the actual sensor schedule $\boldsymbol{\omega}$ and the balanced schedule $\tilde{\boldsymbol{\omega}}$, we adopt the distance $d_{\mathrm{im}} = \|\boldsymbol{\omega} - \tilde{\boldsymbol{\omega}}\|_2$ to measure the energy imbalance between $\boldsymbol{\omega}$ and $\tilde{\boldsymbol{\omega}}$.

In our numerical examples, a normalized value of d_{im} is used for multiple simulation trials. Specifically, let $d_{\mathrm{im}}^{(n)}$ denote the distance between $\boldsymbol{\omega}^{(n)}$ and $\tilde{\boldsymbol{\omega}}^{(n)}$ associated with the sensor schedule for the nth simulation, where $n \in \{1, 2, \ldots, N_{\mathrm{sim}}\}$. We can then define an *energy imbalance measure (EIM)* as follows,

$$\rho_{\mathrm{im}}^{(n)} = \frac{d_{\mathrm{im}}^{(n)} - \underline{d}}{\bar{d} - \underline{d}} \in [0, 1], \tag{28}$$

where $\underline{d}$ and $\bar{d}$ denote the minimum and maximum value of $\{d_{\mathrm{im}}^{(n)}\}_{n=1,2,\ldots,N_{\mathrm{sim}}}$, respectively. Clearly if $\rho_{\mathrm{im}}^{(n)} < \rho_{\mathrm{im}}^{(m)}$, the sensor schedule $\boldsymbol{\omega}^{(n)}$ yields a more balanced energy usage.

In Fig. 2, we present the results for the conventional sensor management problem (5) for a small network with $M = 5$ sensors. Fig. 2-(a) presents EIM and the total number of sensor activations as functions of regularization parameter γ. Here the total number of sensor activations is normalized over $N_{\mathrm{sim}} = 10$ numerical trials, each of which yields the solution of (5) for a given value of γ. We observe that when γ increases, EIM increases although less sensors are activated. This is not surprising since sensors with strong correlation to the field point of interest are successively selected. The specific sensor activation schemes that correspond to the marked values of γ are shown in Fig. 2-(b). For an extreme case of $\gamma = 8$, only the

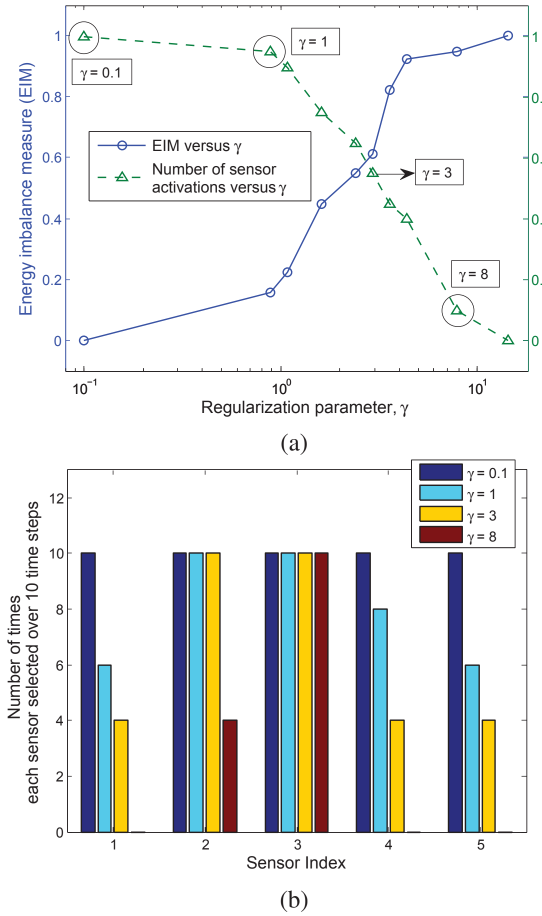

Fig. 2: Conventional sparsity-promoting approach in sensor management, varying γ with $\eta = 0$: (a) EIM and total number of sensor activations, (b) sensor schedules.

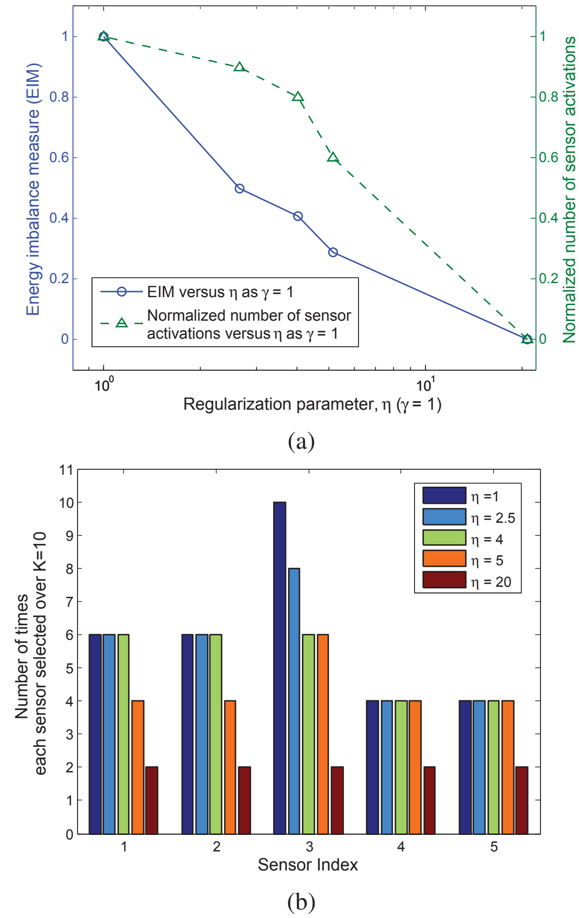

Fig. 3: Promoting sparsity in sensor management from an energy balanced point of view, varying η with $\gamma = 1$: (a) EIM and total number of sensor activations, (b) sensor schedules.

2nd and 3rd sensors are used, but the 3rd sensor is activated at every time step.

In Fig. 3, we present the results for our proposed sensor scheduling problem (10) by varying η and setting $\gamma = 1$. In Fig. 3-(a), we observe that both EIM and the number of sensor activations decrease when η increases. This is in contrast to the result shown in Fig. 2, where EIM increased when the sparsity of sensor schedules was promoted. The energy-balanced sensor schedule is obtained since the new sparsity-promoting penalty $\mathbf{g}$ in (10) enforces sensor energy to be consumed in a balanced way. In Fig. 3-(b), we present the specific sensor schedules for different values of η. As we can see, when η becomes larger than γ, sensors are selected as uniformly as possible. This is because a larger η places more emphasis on minimizing the new sparsity-promoting function, which discourages the successive selection of the most informative sensors.

In Fig. 4, we present the mean squared error (MSE) of field estimation as a function of the number of sensor activations. In this example, we consider a larger network with $M = 10$ sensors. For sensor scheduling, we employ three different methods: our proposed approach for solving problem (11) with $p = 2$ and $\gamma = 0$, the method of [11] for solving problem (11) with $p = 1$ and $\gamma = 0$, and the conventional method for solving problem (11) with $p = 2$ and $\eta = 0$. First, we observe that when the new penalty on energy imbalance is considered, both our approach and the method in [11] yield a higher MSE compared to the conventional sparsity-promoting method. This is because the most informative sensors can be successively selected when the energy-balance concern is ignored. Second, we observe that our approach leads to a slightly higher MSE than that of the method in [11]. This is due to the fact that in [11], a better proxy of the cardinality function (i.e., the reweighted ℓ_1 method [21]) is used to solve the cardinality involved optimization problem (8), while our approach employs an unweighted ℓ_2 formulation. Our numerical results show that the method in [11] fails to grant energy-balanced sensor schedules if an unweighted ℓ_1-norm is used in optimization. In contrast, our approach performs well to promote sparsity in an energy-balanced fashion, and has

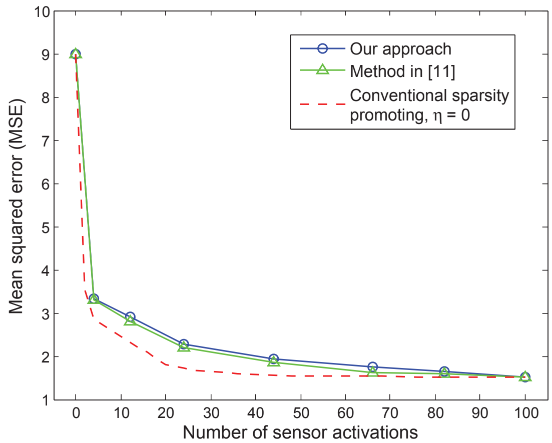

Fig. 4: Estimation performance comparison for different sensor scheduling approaches.

lower computational complexity than the method in [11].

VI. Conclusion

In this paper, we proposed an ℓ_2-norm based sparsity-promoting penalty function to discourage temporally successive selections of the same group of sensors. With the aid of this penalty function, we developed a novel sparsity-promoting framework for sensor management from an energy balance point of view. We demonstrated that the resulting ℓ_2 optimization problem is convex and it can be efficiently solved via the ADMM algorithm. In numerical experiments, we introduced an energy imbalance measure to illustrate the effectiveness of our approach, and provided comparisons with other sensor scheduling algorithms in the literature.

In future work, a combination of reweighted ℓ_1-norm and ℓ_2-norm could be employed to make a better proxy for the column-cardinality of estimator coefficient matrices. Also, a study on the choice of sparsity-promoting parameters will be considered for a given network sparsity level.

Acknowledgment

The work of S. Liu, A. Vempaty, and P. K. Varshney was supported by the U.S. Air Force Office of Scientific Research (AFOSR) under the Grant FA9550-10-1-0458. The work of M. Fardad was supported by the National Science Foundation (NSF) under awards CNS-1329885 and CMMI-0927509. The work of F. Chen and L. Shen was supported by the National Science Foundation (NSF) under the award DMS-1115523.

Appendix A
Proof of Theorem 1

To prove the convexity of problem (12), it is sufficient to study the convexity of the function $g_m(\mathbf{w}) = p^2(\mathbf{w})$, where $p(\mathbf{w}) := \sum_{k=1}^{K} \|\mathbf{w}_{k,m}\|_2$, which is a convex function.

Since $p(\mathbf{w})$ is convex, for any two points $\mathbf{v}$, $\mathbf{u}$ and $\theta \in [0,1]$, we obtain

$$p^2(\theta\mathbf{v} + (1-\theta)\mathbf{u}) \leq (\theta p(\mathbf{v}) + (1-\theta)p(\mathbf{u}))^2. \tag{29}$$

By comparing the right hand side of (29) and the function $\theta g_m(\mathbf{v}) + (1-\theta)g_m(\mathbf{u})$, we have

$$\begin{aligned}&(\theta p(\mathbf{v}) + (1-\theta)p(\mathbf{u}))^2 - \theta g_m(\mathbf{u}) - (1-\theta)g_m(\mathbf{v})\\ =&\theta^2 p^2(\mathbf{v}) + (1-\theta)^2 p^2(\mathbf{u}) + 2\theta(1-\theta)p(\mathbf{v})p(\mathbf{u})\\ &-\theta p^2(\mathbf{v}) - (1-\theta)p^2(\mathbf{u})\\ =&-\theta(1-\theta)(p(\mathbf{v}) - p(\mathbf{u}))^2 \leq 0.\end{aligned} \tag{30}$$

From (29) and (30), we obtain

$$g_m(\theta\mathbf{v} + (1-\theta)\mathbf{u}) \leq \theta g_m(\mathbf{v}) + (1-\theta)g_m(\mathbf{u}),$$

which proves that the function $g_m(\mathbf{w})$ is convex. This proves that the optimization problem (12) is convex. ■

Appendix B
Proof of Proposition 1

Problem (24) can be written as

$$\begin{array}{ll}\underset{\{\mathbf{u}_{k,m}\},\mathbf{r}}{\text{minimize}} & \mathbf{r}^T\mathbf{1}\mathbf{1}^T\mathbf{r} + \dfrac{1}{2\hat{\eta}}\displaystyle\sum_{k=1}^{K}\|\mathbf{u}_{k,m} - \mathbf{c}_{k,m}\|_2^2\\ \text{subject to} & r_k = \|\mathbf{u}_{k,m}\|_2,\ k = 1,2,\ldots,K,\end{array} \tag{31}$$

where we use $\{\mathbf{u}_{k,m}\}$ instead of $\{\mathbf{u}_{k,m}\}_{k=1,2,\ldots,K}$ for simplicity.

Problem (31) can be further transformed to

$$\underset{\mathbf{r}\geq\mathbf{0}}{\text{minimize}}\left\{\underset{\{\|\mathbf{u}_{k,m}\|_2=r_k\}}{\text{minimize}}\ \mathbf{r}^T\mathbf{1}\mathbf{1}^T\mathbf{r} + \frac{1}{2\hat{\eta}}\sum_{k=1}^{K}\|\mathbf{u}_{k,m} - \mathbf{c}_{k,m}\|_2^2\right\}, \tag{32}$$

where the inner minimization problem is with respect to $\{\mathbf{u}_{k,m}\}$, while the outer is with respect to $\mathbf{r}$. Note that the equivalence between (31) and (32) can be verified, proceeding by contradiction and using the fact that the solution of problem (32) is unique.

We first consider the inner minimization problem of (32)

$$\begin{array}{ll}\underset{\{\mathbf{u}_{k,m}\}}{\text{minimize}} & \dfrac{1}{2\hat{\eta}}\displaystyle\sum_{k=1}^{K}\|\mathbf{u}_{k,m} - \mathbf{c}_{k,m}\|_2^2\\ \text{subject to} & \|\mathbf{u}_{k,m}\|_2 = r_k,\ k = 1,2,\ldots,K,\end{array} \tag{33}$$

which can be decomposed into K subproblems

$$\begin{array}{ll}\underset{\mathbf{u}_{k,m}}{\text{minimize}} & \frac{1}{2\hat{\eta}}\|\mathbf{u}_{k,m} - \mathbf{c}_{k,m}\|_2^2\\ \text{subject to} & \|\mathbf{u}_{k,m}\|_2 = r_k\end{array} \tag{34}$$

for $k = 1,2,\ldots,K$.

From a geometrical point of view, the minimizer of (34) can be interpreted as a point lying at the surface of Euclidean ball $\|\mathbf{u}_{k,m}\|_2 \leq r_k$ such that its distance from a give point $\mathbf{c}_{k,m}$ is minimized. Therefore, the solution of (34) is a vector with *length* r_k and *direction* $\frac{\mathbf{c}_{k,m}}{\|\mathbf{c}_{k,m}\|_2}$. Furthermore, the solution of (33) can be given by $\mathbf{u}_{k,m}^* = r_k\frac{\mathbf{c}_{k,m}^T}{\|\mathbf{c}_{k,m}^T\|_2}$ for $k = 1,2,\ldots,K$. Substituting $\{\mathbf{u}_{k,m}^*\}$ into (32), the outer minimization problem becomes QP (26). The proof is now complete. ■

APPENDIX C
PROOF OF PROPOSITION 2

Without loss of generality, we assume that $f_1 \leq f_2 \leq \ldots \leq f_K$. Then according to Lemma 2, we have $r_1^* \leq r_2^* \leq \ldots \leq r_K^*$ for the solution of problem (26). Our goal is to find this value of $\mathbf{r}^*$ in a closed-form.

Note that the solution of problem (26) is unique since it is strongly convex [22]. Also, the optimal primal-dual feasible pair $(\mathbf{r}, \boldsymbol{\nu})$ is given by KKT conditions of (26):

- primal and dual feasibility: $\mathbf{r} \geq \mathbf{0}$ and $\boldsymbol{\nu} \geq \mathbf{0}$, (35a)
- complementary slackness: $r_i \nu_i = 0$ for $i = 1, 2, \ldots, K$, (35b)
- stationary condition: $\mathbf{r} = (\mathbf{I} + 2\hat{\eta}\mathbf{1}\mathbf{1}^T)^{-1}(\mathbf{f} + \hat{\eta}\boldsymbol{\nu})$. (35c)

According to (35c), we obtain

$$\mathbf{r} = \left(\mathbf{I} - \frac{2\hat{\eta}}{1 + 2\hat{\eta}K}\mathbf{1}\mathbf{1}^T\right)(\mathbf{f} + \hat{\eta}\boldsymbol{\nu}), \tag{36}$$

where we use the fact that $(\mathbf{I} + 2\hat{\eta}\mathbf{1}\mathbf{1}^T)(\mathbf{I} - \frac{2\hat{\eta}}{1+2K\hat{\eta}}\mathbf{1}\mathbf{1}^T) = \mathbf{I}$.

Now, from (35b), consider different cases for the values of optimal dual variable $\boldsymbol{\nu}$. If $\boldsymbol{\nu} = \mathbf{0}$, we have

$$r_i = f_i - \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1 + 2\hat{\eta}K}, \; i = 1, 2, \ldots, K. \tag{37}$$

According to (35a) and (35b), we can conclude that $\mathbf{r}^* = \mathbf{r}$ only if $f_1 > \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$. If $f_1 \leq \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$, then the solution of (26) is not given by (37), which yields that the vector of optimal dual variables $\boldsymbol{\nu}$ is not zero. Suppose that $\nu_j \neq 0$ for some $j \in \{1, 2, \ldots, K\}$, then from (35b), we have $r_j^* = 0$. Recall that $0 \leq r_1^* \leq r_2^* \leq \ldots \leq r_K^*$, implying that $r_1^* = 0$. Therefore, we obtain if $f_1 \leq \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$, $r_1^* = 0$. In summary, we have the following result.

If $f_1 > \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$, then $r_i^* = f_i - \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$, for $i = 1, 2, \ldots, K$. If $f_1 \leq \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$, then $r_1^* = 0$.

In the aforementioned conclusion, note that the values of r_i^* for $i = 2, 3 \ldots, K$ have not been determined when $f_1 \leq \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$. However, they can be found by letting $r_1 = 0$ and solving problem (26) with $K - 1$ variables. Similar to the previous analysis, by exploring KKT conditions we obtain the following results.

If $f_1 \leq \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$ and $f_2 > \frac{2\hat{\eta}\sum_{k=2}^K f_k}{1+2\hat{\eta}(K-1)}$, then $r_i^* = f_i - \frac{2\hat{\eta}\sum_{k=2}^K f_k}{1+2\hat{\eta}(K-1)}$, for $i = 2, 3, \ldots, K$. If $f_1 \leq \frac{2\hat{\eta}\sum_{k=1}^K f_k}{1+2\hat{\eta}K}$ and $f_2 \leq \frac{2\hat{\eta}\sum_{k=2}^K f_k}{1+2\hat{\eta}(K-1)}$, then $r_2^* = 0$.

Continuing further, the solution of (26) can be compactly written as

$$r_i^* = \begin{cases} 0 & 1 \leq i \leq \iota - 1, \, i \in \mathbb{N} \\ f_i - \dfrac{2\hat{\eta}\sum_{k=\iota}^K f_k}{1 + 2\hat{\eta}(K - \iota + 1)} & \iota \leq i \leq K, \, i \in \mathbb{N}, \end{cases} \tag{38}$$

for $i = 1, 2, \ldots, K$, where ι is the index of the first positive element in the numbers $\left\{f_i - \frac{2\hat{\eta}\sum_{k=i}^K f_k}{1+2\hat{\eta}(K-i+1)}\right\}_{i=1,2,\ldots,K}$ and $\iota = K + 1$ if no positive element exists. The proof is now complete. ■

REFERENCES

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

[2] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4938–4953, Oct. 2011.

[3] M. P. Vitus, W. Zhang, A. Abate, J. Hu, and C. J. Tomlin, "On efficient sensor scheduling for linear dynamical systems," *Automatica*, vol. 48, no. 10, pp. 2482–2493, Oct. 2012.

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

[5] S. Liu, M. Fardad, E. Masazade, and P. K. Varshney, "Optimal periodic sensor scheduling in networks of dynamical systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3055–3068, Jun. 2014.

[6] I. D. Schizas, "Distributed informative-sensor identification via sparsity-aware matrix decomposition," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4610–4624, Sept. 2013.

[7] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 684–698, Feb. 2015.

[8] E. Ertin, J. W. Fisher, and L. C. Potter, "Maximum mutual information principle for dynamic sensor query problems," in *Proc. Int. Conf. Inf. Processing Sensor Netw.*, 2003, pp. 405–416.

[9] S. Liu, M. Fardad, E. Masazade, and P. K. Varshney, "Sparsity-aware field estimation via ordinary Kriging," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP 2014)*, May 2014, pp. 3976–3980.

[10] Y. Chen, C. Chuah, and Q. Zhao, "Sensor placement for maximizing lifetime per unit cost in wireless sensor networks," in *Proc. 2005 IEEE Military Commun. Conf. (MILCOM)*, Oct. 2005, pp. 1097–1102.

[11] 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, Dec. 2014.

[12] S. Liu, M. Fardad, E. Masazade, and P. K. Varshney, "On optimal periodic sensor scheduling for field estimation in wireless sensor networks," in *Proc. IEEE Global Conf. Signal Inf. Processing (GlobalSIP)*, Dec. 2013, pp. 137–140.

[13] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 123–231, 2013.

[14] S. Boyd, "ℓ_1-norm methods for convex-cardinality problems," *http://web.stanford.edu/class/ee364b/lectures/l1_slides.pdf*, Lecture notes for EE364b, Stanford University.

[15] L. Yuan, J. Liu, and J. Ye, "Efficient methods for overlapping group lasso," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2104–2116, Sep. 2013.

[16] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. R. Stat. Soc. Ser. B. Methodol.*, vol. 68, no. 1, pp. 49–67, Feb. 2006.

[17] 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.

[18] S. Liu, S. Kar, M. Fardad, and P. K. Varshney, "Sparsity-aware sensor collaboration for linear coherent estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2582–2596, May 2015.

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

[20] V. V. Williams, "Multiplying matrices faster than Coppersmith-Winograd," in *Proc. 44th Symp. Theory Comp.*, 2012, pp. 887–898.

[21] 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, Dec. 2008.

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