Sensor Selection with Correlated Measurements for Target Tracking in Wireless Sensor Networks

*Sijia Liu**, *Engin Masazade*†, *Makan Fardad** and *Pramod K. Varshney**
*Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY, USA
Email: {sliu17, makan, varshney}@syr.edu
†Department of Electrical and Electronics Engineering, Yeditepe University, Istanbul, Turkey
Email: engin.masazade@yeditepe.edu.tr

***Abstract*—We study the problem of adaptive sensor management for target tracking, where at every instant we search for the best sensors to be activated at the next time step. In our problem formulation, the measurements may be corrupted by correlated noises, and the impact of correlated measurements on sensor selection is studied. Specifically, we adopt an alternative conditional posterior Cramér-Rao lower bound (C-PCRLB) as the optimization criterion for sensor selection, where the trace of the conditional Fisher information matrix is maximized subject to an energy constraint. We demonstrate that the proposed sensor selection problem can be transformed into the problem of maximizing a convex quadratic function over a bounded polyhedron. This optimization problem is NP-hard in nature, and thus we employ a linearization method and a bilinear programming approach to obtain locally optimal sensor schedules in a computationally efficient manner.**

***Index Terms*—Target tracking, sensor selection, convex optimization, wireless sensor networks.**

I. Introduction

In target tracking using a wireless sensor network, sensors observe an unknown state and report noisy measurements to a fusion center which determines an estimate of the target state. However, due to resource constraints such as communication bandwidth and sensor battery life, it is desirable that only a subset of sensors be activated over time and space [1]. Thus, the issue of sensor selection/scheduling arises.

Several variations of sensor selection/scheduling problems have been studied in the literature [2]–[8] according to the type of quantity to be estimated (parameter or random process), measurement models (linear or nonlinear) and cost functions (estimation performance or energy consumption). In the aforementioned literature, a posterior Cramér-Rao lower bound (PCRLB) is commonly used as the optimization criterion for sensor selection in target tracking [2], [3]. However, the PCRLB is not suitable for adaptive sensor management since it is an *offline* bound, which can be determined by the dynamic model, measurement model and prior knowledge of the system state at the initial time [9]. Therefore, we adopt an alternative *conditional* PCRLB (C-PCRLB) as the performance metric for sensor selection, where the considered C-PCRLB is an *online* bound which incorporates the effect of actual sensor measurements [10].

In [2]–[8], the sensor selection problem was studied by assuming the measurement noise to be uncorrelated, which implies that each measurement contributes to the Fisher information in an *additive* manner. By exploiting this structure, sensor selection problems with uncorrelated measurements can be efficiently solved via convex relaxations. Here, we consider a more general scenario in which the measurement noises are correlated, and for example, the correlated measurements occur in a spatially-correlated sensor field [11].

The sensor selection problem with correlated measurements has been recently studied in [12], [13] for parameter estimation and state tracking, respectively. In [12], the problem was formulated by minimizing the number of selected sensors subject to an estimation quality. In [13], the problem was solved by maximizing the trace of Fisher information under an energy constraint. The proposed sensor selection problems in [12], [13] are nonconvex due to the Boolean selection variables as well as the impact of correlated noises. Thus, locally optimal solutions were obtained by using semidefinite programs (SDPs). Compared with the existing literature [12], [13], our contributions are listed as follows.

- We propose an adaptive sensor selection framework, in which an alternative conditional PCRLB is used as the optimization criterion.
- We demonstrate that the problem of sensor selection with correlated measurements (in terms of maximizing the trace of information matrix under an energy constraint) can be transformed into the problem of *maximizing* a *convex* quadratic function over a bounded polyhedron.
- We show that locally optimal solutions of the proposed sensor selection problem can be efficiently found by solving linear programs rather than SDPs.

II. System Model

Consider a discrete-time dynamical system

$$\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \mathbf{u}_t \tag{1}$$

$$\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{v}_t, \tag{2}$$

where $\mathbf{x}_t \in \mathbb{R}^n$ is the target state at time t, $\mathbf{y}_t \in \mathbb{R}^m$ is the measurement vector whose ith entry corresponds to a scalar observation from the ith sensor, $\mathbf{F}$ is the state transition matrix, and $\mathbf{h}(\cdot)$ denotes a nonlinear measurement model. The inputs $\mathbf{u}_t$ and $\mathbf{v}_t$ are white, Gaussian, zero-mean random vectors with covariance matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively. We remark that the covariance matrix $\mathbf{R}$ may not be a diagonal matrix, since the noise experienced by different sensors could be *correlated.*

Without loss of generality, we assume that the target state is a 4×1 dimensional vector $\mathbf{x}_t = [x_{t,1}, x_{t,2}, x_{t,3}, x_{t,4}]^T$, where $(x_{t,1}, x_{t,2})$ and $(x_{t,3}, x_{t,4})$ denote the target location and velocity in the 2-D plane, respectively. The state equation (1) is assumed to follow a white noise acceleration model [6], where

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ \mathbf{Q} = q \begin{bmatrix} \frac{\Delta^3}{3} & 0 & \frac{\Delta^2}{2} & 0 \\ 0 & \frac{\Delta^3}{3} & 0 & \frac{\Delta^2}{2} \\ \frac{\Delta^2}{2} & 0 & \Delta & 0 \\ 0 & \frac{\Delta^2}{2} & 0 & \Delta \end{bmatrix}. \tag{3}$$

In (3), Δ and q denote the sampling interval and the process noise parameter, respectively.

The measurement equation (2) is given by a power attenuation model [14], where for the mth sensor at time t, the nonlinear measurement function follows

$$h_i(\mathbf{x}_t) = \sqrt{\frac{P_0}{1 + d_{t,i}^2}}, \quad i = 1, 2, \ldots, m. \tag{4}$$

In (4), P_0 is the signal power of the target, $d_{t,i} = \sqrt{(x_{t,1} - \alpha_i)^2 + (x_{t,2} - \beta_i)^2}$, $(x_{t,1}, x_{t,2})$ is the target location at time t, (α_i, β_i) is the position of the ith sensor, and $\mathbf{h}(\mathbf{x}_t) = [h_1(\mathbf{x}_t), h_2(\mathbf{x}_t), \ldots, h_m(\mathbf{x}_t)]^T$.

In this paper, we employ a sampling importance resampling filter [15, Algorithm 1] to track the target for a nonlinear dynamical system (1)-(2). The SIR filter is a particle filter, which refers to a sequential Monte Carlo method of using point mass (also referred to as particle) representations to approximate the probability densities in estimation. The particle filter has been widely used in problems of nonlinear filtering; see more applications in [16]. Considering the issue of sensor selection, at every time step we require the search for the best sensors to be activated at the next time step. The sensor selection problem will be elaborated on in the following sections.

III. Problem Formulation

Recently, a conditional posterior Cramér-Rao lower bound (C-PCRLB) has been proposed in [9] by incorporating the history of actual sensor measurements. Compared with the conventional PCRLB, the C-PCRLB provides a tighter error bound on mean squared error (MSE). Nevertheless, obtaining C-PCRLB is computationally intensive due to the presence of the auxiliary Fisher information matrix; see [9, Thm.1]. To reduce the complexity, reference [10] introduced an *alternative* C-PCRLB, which yields an estimate of the *exact* C-PCRLB. In this paper, we adopt the alternative C-PCRLB as the performance criterion for sensor selection.

A. Alternative conditional posterior Cramér-Rao lower bound

Let $\mathbf{x}_{0:t}$ and $\mathbf{y}_{1:t}$ denote the state vector and measurements up to time t. Then the MSE of the estimate $\hat{\mathbf{x}}_{0:t}$ (conditioned on the measurements $\mathbf{y}_{1:t-1}$) is lower bounded by the inverse of the *conditional* Fisher information matrix [9]

$$E\{[\hat{\mathbf{x}}_{0:t}-\mathbf{x}_{0:t}][\hat{\mathbf{x}}_{0:t}-\mathbf{x}_{0:t}]^T|\mathbf{y}_{1:t-1}\}\geq \mathbf{J}^{-1}(\mathbf{x}_{0:t}|\mathbf{y}_{1:t-1}).$$

Let $\mathbf{J}(\mathbf{x}_t|\mathbf{x}_{1:t-1})$ denote the matrix whose inverse equals the lower-right corner submatrix of $\mathbf{J}^{-1}(\mathbf{x}_{0:t}|\mathbf{y}_{1:t-1})$. Then the inversion of $\mathbf{J}(\mathbf{x}_t|\mathbf{y}_{1:t-1})$ provides a lower bound on the MSE of estimating $\mathbf{x}_t$.

For the linear Gaussian model (1), $\mathbf{J}(\mathbf{x}_t|\mathbf{y}_{1:t-1})$ can be computed as follows [10, Corollary 1],

$$\mathbf{J}_t=\left(\mathbf{Q}+\mathbf{F}\mathbf{J}_{t-1}^{-1}\mathbf{F}^T\right)^{-1}+\mathbf{\Phi}_t \tag{5}$$

$$\mathbf{\Phi}_t:=E_{p_t^c}\left\{-\Delta_{\mathbf{x}_t}^{\mathbf{x}_t}\ln p(\mathbf{y}_t|\mathbf{x}_t)\right\}, \tag{6}$$

where for simplicity we have used, and henceforth will continue to use $\mathbf{J}_t$ instead of $\mathbf{J}(\mathbf{x}_t|\mathbf{y}_{1:t-1})$, $p_t^c:=p(\mathbf{x}_t,\mathbf{y}_t|\mathbf{y}_{1:t-1})$, and $\Delta_{\mathbf{x}}^{\mathbf{x}}$ denotes the second-order partial derivative with respect to $\mathbf{x}$. Due to the assumption of Gaussian noise, the observation likelihood $p(\mathbf{y}_t|\mathbf{x}_t)$ in (6) is given by

$$p(\mathbf{y}_t|\mathbf{x}_t)=\frac{(2\pi)^{-2}}{\det(\mathbf{R})^{1/2}}e^{-\frac{1}{2}[\mathbf{y}_t-\mathbf{h}(\mathbf{x}_t)]^T\mathbf{R}^{-1}[\mathbf{y}_t-\mathbf{h}(\mathbf{x}_t)]}, \tag{7}$$

Using the fact that $\mathbf{y}_t$, $\mathbf{x}_t$ and $\mathbf{y}_{1:t-1}$ form a Markov chain, we have $p_t^c=p(\mathbf{x}_t|\mathbf{y}_{1:t-1})p(\mathbf{y}_t|\mathbf{x}_t)$. Substituting (7) into (6), we obtain

$$\mathbf{\Phi}_t=E_{\hat{p}_t}\left[(\nabla_{\mathbf{x}_t}\mathbf{h}^T)\mathbf{R}^{-1}(\nabla_{\mathbf{x}_t}\mathbf{h}^T)^T\right],\ \ \hat{p}_t:=p(\mathbf{x}_t|\mathbf{y}_{1:t-1}), \tag{8}$$

where we replace $\mathbf{h}(\mathbf{x}_t)$ with $\mathbf{h}$ for simplicity.

It is clear from (8) that the alternative conditional PCRLB incorporates the history of sensor measurements $\mathbf{y}_{1:t-1}$ and thus is an *online* bound. This is in contrast with the conventional PCRLB (see [16, Eqs. 34-36]), which involves a similar form as (8) except for the incorporation of $\mathbf{y}_{1:t-1}$.

B. Sensor selection problem

We introduce a vector of binary variables

$$\mathbf{w}=[w_1,w_2,\ldots,w_m]^T,\ w_i\in\{0,1\}$$

to characterize the sensor schedule during one snapshot. Suppose, for example, that at time t only the jth sensor reports a measurement. In this case, it follows from (4) that $w_ih_{t,i}=0$ for all $i\neq j$.

By incorporating the selection variables $\mathbf{w}$, Fisher information (5) can be rewritten as [12]

$$\mathbf{J}_t=\left(\mathbf{Q}+\mathbf{F}\mathbf{J}_{t-1}^{-1}\mathbf{F}^T\right)^{-1}+\mathbf{\Psi}_t \tag{9}$$

$$\begin{aligned}\mathbf{\Psi}_t&:=E_{\hat{p}_t}[(\nabla_{\mathbf{x}_t}\mathbf{h}^T)\mathbf{D}_w\mathbf{R}^{-1}\mathbf{D}_w(\nabla_{\mathbf{x}_t}\mathbf{h}^T)^T]\\&=E_{\hat{p}_t}[(\nabla_{\mathbf{x}_t}\mathbf{h}^T)(\mathbf{w}\mathbf{w}^T\circ\mathbf{R}^{-1})(\nabla_{\mathbf{x}_t}\mathbf{h}^T)^T],\end{aligned} \tag{10}$$

where $\mathbf{D}_w$ denotes the diagonal matrix $\mathrm{diag}(\mathbf{w})$, and $\circ$ represents the Hadamard (elementwise) product.

Remark 1: *If $\mathbf{R}$ is a diagonal matrix, then $\mathbf{\Psi}_t$ reduces to a linear matrix function with respect to the selection variables, i.e.,*

$$\mathbf{\Psi}_t=\sum_{i=1}^m\frac{w_i}{R_{ii}}E_{\hat{p}_t}\left[(\nabla_{\mathbf{x}_t}h_i)(\nabla_{\mathbf{x}_t}h_i)^T\right], \tag{11}$$

where we use the fact that $w_i^2=w_i$ for $i=1,2,\ldots,m$. This implies that sensor measurements contribute in an additive manner for the information matrix. Such a structure has been explored in [4], [7] when the measurement noises are uncorrelated.

To find the optimal sensor selection scheme $\mathbf{w}$, various scalar functions of the Fisher information matrix can be used as performance metrics; examples include $\mathrm{tr}(\mathbf{J}_t^{-1})$ [3], $\det(\mathbf{J}_t)$ [4] and $\mathrm{tr}(\mathbf{J}_t)$ [13], [17]. Compared to the sensor selection problem with uncorrelated measurements [3], [4], the derivation of $\det(\mathbf{J}_t)$ and $\mathrm{tr}(\mathbf{J}_t^{-1})$ with *correlated* noise becomes much more involved due to the presence of off-diagonal entries of $\mathbf{R}$. For simplicity of computation, we consider the maximization of $\mathrm{tr}(\mathbf{J}_t)$, which can be further transformed into a convex quadratic function with respect to $\mathbf{w}$.

Now, we pose the sensor selection problem by maximizing $\mathrm{tr}(\mathbf{J}_t)$

$$\underset{\mathbf{w}}{\text{maximize}}\quad \mathrm{tr}(\mathbf{\Psi}) \tag{12a}$$

$$\text{subject to}\quad \mathbf{1}^T\mathbf{w}\leq s, \tag{12b}$$

$$w_i\in\{0,1\},\ i=1,2,\ldots,m.$$

where we replace $\mathbf{\Psi}_t$ with $\mathbf{\Psi}$ for simplicity, the term $\mathrm{tr}[(\mathbf{Q}+\mathbf{F}\mathbf{J}_{t-1}^{-1}\mathbf{F}^T)^{-1}]$ in $\mathrm{tr}(\mathbf{J}_t)$ is omitted since it is given *a priori* at time t, and s denotes the maximum number of sensors to be activated. In what follows, we will refer to inequality (12b) as the energy constraint.

IV. Sensor Selection with Correlated Measurements

In this section, we show that the proposed sensor selection problem (12) can be transformed into the maximization of a convex quadratic function over a bounded polyhedron.

From (10), we can rewrite (12a) as

$$\mathrm{tr}(\mathbf{\Psi})=\mathbf{w}^T\mathbf{\Omega}\mathbf{w}, \tag{13}$$

where the (i,j)th entry of $\mathbf{\Omega}$ is given by

$$\Omega_{ij}=E_{\hat{p}_t}[(\nabla_{\mathbf{x}_t}h_i)^T(\mathbf{R})_{ij}(\nabla_{\mathbf{x}_t}h_j)], \tag{14}$$

and $(\mathbf{R})_{ij}$ is the (i,j)th entry of $\mathbf{R}$. In (13), we can further show that the matrix $\mathbf{\Omega}$ is *positive semidefinite* (noted by $\mathbf{\Omega}\succeq 0$) since $\mathbf{\Omega}$ can be rewritten as $\mathbf{\Omega}=E_{\hat{p}_t}[\mathbf{A}(\mathbf{x}_t)(\mathbf{R}^{-1}\otimes\mathbf{I})\mathbf{A}(\mathbf{x}_t)^T]$, where $\mathbf{A}(\mathbf{x}_t)=\mathrm{blkdiag}[(\frac{\partial h_1(\mathbf{x}_t)}{\partial\mathbf{x}_t})^T,(\frac{\partial h_2(\mathbf{x}_t)}{\partial\mathbf{x}_t})^T,\ldots,(\frac{\partial h_m(\mathbf{x}_t)}{\partial\mathbf{x}_t})^T]\in\mathbb{R}^{m\times mn}$. [1] There-

[1]The notation $\mathrm{blkdiag}[\mathbf{A}_1,\ldots,\mathbf{A}_m]$ signifies a block-diagonal matrix whose diagonal entries are the matrices $\mathbf{A}_1,\ldots,\mathbf{A}_m$.

fore, the objective function in (12) is a *convex quadratic* function. We omit details of the derivations for the sake of brevity.

Problem (12) can therefore be written as

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\mathbf{\Omega}\mathbf{w} \\ \text{subject to} & \mathbf{1}^T\mathbf{w} \leq s,\ w_i \in \{0,1\},\ i=1,2,\ldots,m, \end{array} \tag{15}$$

where we elaborate on the computation of $\mathbf{\Omega} \succeq 0$, using particle representations, in Appendix A.

Remark 2: *For the special case in which measurement noises are uncorrelated (i.e., $\mathbf{R}$ is a diagonal matrix) [4], we have $\Omega_{ij} = 0$ for $i \neq j$ from* (14)*. Then, problem* (15) *becomes*

$$\begin{array}{ll} \underset{\mathbf{w}}{\textit{maximize}} & \sum_{i=1}^m w_i\Omega_{ii} \\ \textit{subject to} & \mathbf{1}^T\mathbf{w} \leq s,\ w_i \in \{0,1\},\ i=1,2,\ldots,m, \end{array} \tag{16}$$

where we use the fact that $w_i^2 = w_i$ for $i = 1, 2, \ldots, m$. In (16)*, the optimal sensor selection scheme is given by choosing the s sensors which yield the first s largest values of $\{\Omega_{ii}\}_{i=1,2,\ldots,m}$. This result is consistent with [4], [13, Theorem 3].*

A standard method of solving (15) is to replace the binary variable $w_{t,i}$ with a continuous variable $w_i \in [0,1]$. This leads to a relaxed problem

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\mathbf{\Omega}\mathbf{w} \\ \text{subject to} & \mathbf{1}^T\mathbf{w} \leq s,\ \mathbf{0} \leq \mathbf{w} \leq \mathbf{1}, \end{array} \tag{17}$$

which generally provides an upper bound on the optimal value of (15). However, in the scenario of $\mathbf{\Omega} \succeq 0$, Proposition 1 demonstrates that problem (17) is equivalent to problem (15).

Proposition 1: Problems (15) and (17) are equivalent.

Proof: The idea of the proof lies in the fact that an optimal solution to problem (17) occurs at a vertex of the associated polyhedron [18]. For the sake of brevity, details of the proof are omitted here. ■

Note that problem (17) is not a convex optimization problem, and it has been shown in [18] that finding a globally optimal solution to the problem of maximizing a convex quadratic function over a bounded polyhedron is NP-hard. Therefore, it is a nontrivial exercise to solve problem (17).

V. Optimization Methods

Typically, there exist two types of optimization approaches for solving problem (17): Global optimization methods (e.g., cutting plane [19] and branch and bound [20]) and local optimization methods (e.g., linearization [21] and bilinear programming [18]). The former yield an ϵ-globally optimal solution but are often slow, while the latter are fast but result in a locally optimal solution. To achieve real-time target tracking, it is desirable to solve problem (17) in a time-efficient manner. Thus, we employ local optimization methods to obtain the locally optimal solutions.

For convenience, we rewrite problem (17) as the minimization of a concave quadratic function

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & -\mathbf{w}^T\mathbf{\Omega}\mathbf{w} \\ \text{subject to} & \mathbf{1}^T\mathbf{w} \leq s,\ \mathbf{0} \leq \mathbf{w} \leq \mathbf{1}. \end{array} \tag{18}$$

A. Linearization method

An iterative linearization method is introduced in [22], [23] to solve the nonconvex quadratic program (QP) by linearizing the nonconvex parts of quadratic functions. Thus in (18), we replace the concave objective function with its first-order Taylor expansion around a feasible point $\bar{\mathbf{w}}$, namely, $-2\mathbf{w}^T\mathbf{\Omega}\bar{\mathbf{w}} + \bar{\mathbf{w}}^T\mathbf{\Omega}\bar{\mathbf{w}}^T$. This linearized function provides an affine upper bound of the concave objective function in (18). A locally optimal solution of problem (18) is then obtained by solving the linear program (LP)

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & -2\mathbf{w}^T\mathbf{\Omega}\bar{\mathbf{w}} + \bar{\mathbf{w}}^T\mathbf{\Omega}\bar{\mathbf{w}} \\ \text{subject to} & \mathbf{1}^T\mathbf{w} \leq s,\ \mathbf{0} \leq \mathbf{w} \leq \mathbf{1}. \end{array} \tag{19}$$

The linearizing method (also called convex-concave procedure [23]) is summarized in Algorithm 1.

Algorithm 1 Linearization method for solving (18)

Require: Given an initial feasible point $\mathbf{w}_0$ and ϵ.
1: **for** k = 1, 2, ..., **do**
2: set $\bar{\mathbf{w}} = \mathbf{w}_{k-1}$, and solve (19) for the solution $\mathbf{w}_k$.
3: terminate if $|\mathbf{w}_k^T\mathbf{\Omega}\mathbf{w}_k - \mathbf{w}_{k-1}^T\mathbf{\Omega}\mathbf{w}_{k-1}| \leq \epsilon$.
4: **end for**

We remark that the solution obtained from Algorithm 1 may not be a good approximation for the optimal solution of (18) due to the choice of the initial linearizing point $\mathbf{w}_0$. Reference [24, Sec. 2] has suggested an approach for selecting a *good* initial point such that the error bound between the objective value of (18) and that of (19) is minimized. However, this procedure requires solving $2m$ linear programs to obtain the lower and upper bound on $\mathbf{u}_i^T\mathbf{w}$ for $i = 1, 2, \ldots, m$, where $\{\mathbf{u}_i\}_{i=1,2,\ldots,m}$ denote the eigenvectors of $-\mathbf{\Omega}$. For ease of computation, here we set the initial value of $\mathbf{w}_0$ as a vector of containing s random numbers (drawn from the standard uniform distribution) and 0s elsewhere.

B. Bilinear programming approach

According to [18, Theorem 2.2], we obtain that problem (18) is equivalent to an associated bilinear program[2]

$$\begin{array}{ll} \underset{\mathbf{w},\mathbf{z}}{\text{minimize}} & -\mathbf{w}^T\mathbf{\Omega}\mathbf{z} \\ \text{subject to} & \mathbf{w} \in \mathcal{P},\ \mathbf{z} \in \mathcal{P}, \end{array} \tag{20}$$

where $\mathcal{P} = \{\mathbf{x} | \mathbf{1}^T\mathbf{x} \leq s, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$, and if $(\mathbf{w}^*, \mathbf{z}^*)$ is optimal for (20), then both $\mathbf{w}^*$ and $\mathbf{z}^*$ are optimal for (18).

To solve problem (20), we employ an alternative optimization method (also known as the mountain climbing algorithm [18], [25]) to seek a locally optimal solution of (20) as demonstrated in Algorithm 1. To initialize Algorithm 1, we use the same strategy of generating the starting point in Sec. V-A. At each iteration, we solve two LPs in Steps 2-3 of Algorithm 1.

Algorithm 2 Alternative optimization method for solving (20)

Require: given a feasible vector $\mathbf{z}_0 \in \mathcal{P}$ and ϵ.
1: **for** k = 1, 2, ..., **do**
2: solve LP: $\mathbf{w}_k = \arg\min_{\mathbf{w}\in\mathcal{P}}\{\psi(\mathbf{w}, \mathbf{z}_{k-1})\}$.
3: solve LP: $\mathbf{z}_k = \arg\min_{\mathbf{z}\in\mathcal{P}}\{\psi(\mathbf{w}_k, \mathbf{z})\}$.
4: terminate if $|\mathbf{w}_k^T\mathbf{\Omega}\mathbf{z}_k - \mathbf{w}_k^T\mathbf{\Omega}\mathbf{z}_{k-1}| \leq \epsilon$.
5: **end for**

VI. Numerical Results

In our simulations, we assume that $m = 30$ sensors are deployed in a 10×10 m^2 area. In state model (3), we set $\Delta = 1$ and $q = 0.1$. The initial state distribution of the target is assumed to be Gaussian with mean $\mu_0 = [0, 0, 0.8, 0.8]$ and covariance $\Sigma_0 = \text{diag}[1, 1, 0.01, 0.01]$. In measurement

[2] A bilinear program refers to the problem which involves a bilinear objective function $\psi(\mathbf{w}, \mathbf{z}) = \mathbf{a}^T\mathbf{w} + \mathbf{b}^T\mathbf{z} + \mathbf{w}^T\mathbf{P}\mathbf{z}$ and linear constraints.

model (2), we set $P_0 = 10^4$, and the noise covariance matrix $\mathbf{R}$ is given by an exponential model [11] $\mathrm{cov}(v_{t,i}, v_{t,j}) = \sigma_v^2 \exp\{-\rho\sqrt{(\alpha_i - \alpha_j)^2 + (\beta_i - \beta_j)^2}\}$, where $\sigma_v^2 = 1$, $\rho = 0.025$ corresponds to a strong correlation, and (α_i, β_i) denotes the position of the ith sensor. In Algorithm 1, we choose $\epsilon = 10^{-3}$. To evaluate the tracking performance, we perform 100 simulation trials, during each of which we use $K = 5000$ particles in the particle filter.

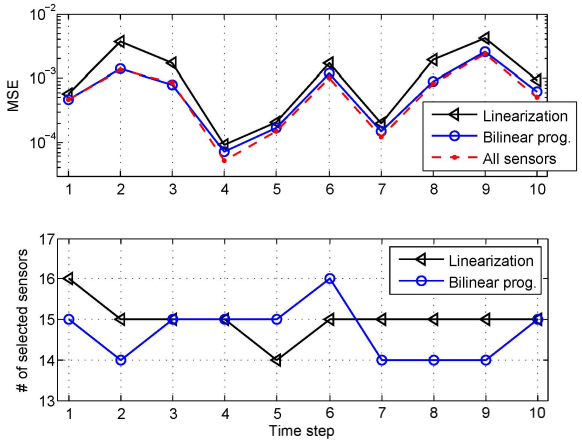

Fig. 1: MSE and number of selected sensors versus time steps.

In Fig. 1, we present the MSE and the number of selected sensors as functions of time steps, where the energy budget is set by $s = m = 20$. For comparison, we also show the MSE when all of sensors are activated in target tracking. We observe that the bilinear programming approach outperforms the linearization method, since the optimization performance of the latter might be (negatively) affected by the choice of the linearizing point. Although the number of selected sensors obtained from the bilinear programming approach is less than s, the corresponding MSE is close to the estimation performance of using all $s = 20$ sensors. This is because the presence of correlation among sensors leads to information redundancy, which causes less diversity in received measurements. As a result, the estimation performance ceases to significantly improve when a very large number of sensors are activated.

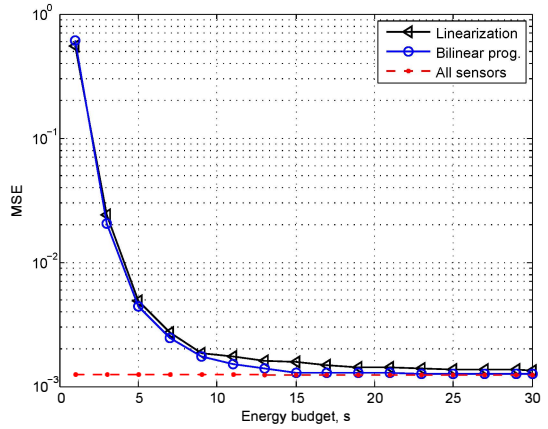

Fig. 2: MSE as a function of m (i.e., the bound on energy constraint).

In Fig. 2, we present the MSE by varying the energy budget of s, where the MSE is averaged over 10 time steps. As we can see, the estimation performance ceases to significantly improve when $m >$ 15. This result is consistent with our analysis in Fig. 1. For $m = 11$, we show the resulting sensor schedules for two time steps in Fig. 3. Note that the sensors closest to the target are selected due to their high received signal power. And the selected sensors tends to be spatially distributed rather than aggregating in a small neighborhood around the target, since observations from neighboring sensors are strongly correlated and lead to information redundancy.

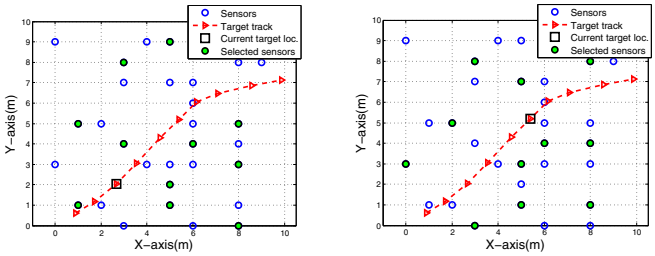

Fig. 3: Sensor schedules for two time steps with $s = 11$: (a) $t = 3$, (b) $t = 6$.

VII. Conclusion

In this paper, the sensor selection problem for target tracking with correlated measurements was studied. Using the trace of the conditional Fisher information matrix as an objective to be maximized subject to energy constraints, it was shown that the sensor selection problem can be interpreted as the maximization of a convex quadratic function over a bounded polyhedron. This optimization problem is nonconvex in nature, and we applied a linearization method and bilinear programming approach to find locally optimal solutions. Our proposed solution algorithms are based on linear programming and therefore scale gracefully with problem size. Numerical results show that correlated measurements bring in information redundancy, which has a large impact on sensor selection. In this paper, the problem of sensor selection was only considered for scheduling sensors at the next time step. Our future work will study the problem of non-myopic scheduling, where sensor schedules are determined for multiple future time steps. Also, we will study other scalar measures of Fisher information for sensor selection with correlated measurements.

Appendix A
Particle Representation of $\mathbf{\Omega}$

In particle filters, the posterior probability density $\hat{p}_t$ is given by $\hat{p}_t = p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) \approx \sum_{k=1}^{K} \frac{1}{K}\delta(\mathbf{x}_t - \mathbf{x}_t^{(k)})$, where K is the number of particles, $\delta(\cdot)$ is the δ-function, and $\mathbf{x}_t^{(k)}$ denotes the kth particle at time t. Substituting (4) and the particle representation of $\hat{p}_t$ into (14), we obtain

$$\Omega_{ij} = \frac{\bar{R}_{ij} P_0}{K} \sum_{p=1}^{K} \frac{(\alpha_i - x_{t,1}^{(k)})(\alpha_j - x_{t,1}^{(k)}) + (\beta_i - x_{t,2}^{(k)})(\beta_j - x_{t,2}^{(k)})}{\left\{\left[1 + (d_{i,t}^{(k)})^2\right]\left[1 + (d_{j,t}^{(k)})^2\right]\right\}^{\frac{3}{2}}}, \tag{21}$$

where $(x_{t,1}^{(k)}, x_{t,2}^{(k)})$ denotes the target position given by the kth particle, $d_{i,t}^{(k)} = \sqrt{(\alpha_i - x_{t,1}^{(k)})^2 + (\beta_i - x_{t,2}^{(k)})^2}$. The complexity of computing $\mathbf{\Omega}$ is approximately $O(m^2 K)$, since for each entry of $\mathbf{\Omega} \in \mathbb{R}^{m \times m}$, there exist K additive terms in (21).

Acknowledgment

This work was supported by U.S. Air Force Office of Scientific Research (AFOSR) under Grant FA9550-10-1-0458, National Science Foundation (NSF) under awards CNS-1329885 and CMMI-0927509, and the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant 113E220.

REFERENCES

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

[2] M. L. Hernandez, T. Kirubarajan, and Y. Bar-Shalom, "Multisensor resource deployment using posterior Cramér-Rao bounds," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 2, pp. 399–416, April 2004.

[3] R. Tharmarasa, T. Kirubarajan, M. Hernandez, and A Sinha, "PCRLB-based multisensor array management for multitarget tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 539–555, April 2007.

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

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

[6] 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.

[7] S. P. Chepuri and G. Leus, "Sparsity-promoting adaptive sensor selection for non-linear filtering," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 5080–5084.

[8] 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, Dec 2014.

[9] L. Zuo, R. Niu, and P. K. Varshney, "Conditional posterior Cramér-Rao lower bounds for nonlinear sequential Bayesian estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 1–14, Jan. 2011.

[10] Y. Zheng, O. Ozdemir, R. Niu, and P. K. Varshney, "New conditional posterior Cramér-Rao lower bounds for nonlinear sequential Bayesian estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5549–5556, Oct. 2012.

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

[12] H. Jamali-Rad, A. Simonetto, G. Leus, and X. Ma, "Sparsity-aware sensor selection for correlated noise," in *Proceedings of International Conference on Information Fusion (FUSION)*, July 2014, pp. 1–7.

[13] X. Shen and P. K. Varshney, "Sensor selection based on generalized information gain for target tracking in large sensor networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 363–375, Jan 2014.

[14] R. Niu and P. K. Varshney, "Target location estimation in sensor networks with quantized data," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4519–4528, Dec. 2006.

[15] E. Masazade, R. Niu, and P. K. Varshney, "Dynamic bit allocation for object tracking in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5048–5063, Oct. 2012.

[16] P. Tichavsky, C.H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Transactions on Signal Processin*, vol. 46, no. 5, pp. 1386–1396, may 1998.

[17] S. Liu, E. Masazade, X. Shen, and P. K. Varshney, "Adaptive non-myopic quantizer design for target tracking in wireless sensor networks," in *Proceedings of Asilomar Conference on Signals, Systems and Computers*, Nov 2013, pp. 1085–1089.

[18] H. Konno, "Maximization of a convex quadratic function under linear constraints," *Mathematical Programmin*, vol. 11, no. 1, pp. 117–127, 1976.

[19] H. Konno, "A cutting plane algorithm for solving bilinear programs," *Mathematical Programming*, vol. 11, no. 1, pp. 14–27, 1976.

[20] F. Al-Khayyal and J. E. Falk, "Jointly constrained biconvex programming," *Mathematics of Operations Research*, vol. 8, no. 2, pp. 273–286, 1983.

[21] P. M. Pardalos, "Global optimization algorithms for linearly constrained indefinite quadratic problems," *Computers & Mathematics with Applications*, vol. 21, no. 67, pp. 87 – 97, 1991.

[22] A. d'Aspremont and S. Boyd, "Relaxations and randomized methods for nonconvex qcqps," Stanford, CA: Stanford Univ., Autumn 2003. [Online]. Available: *http://web.stanford.edu/class/ee392o/relaxations.pdf*.

[23] T. Lipp and S. Boyd, "Variations and extensions of the convex-concave procedure," http://web.stanford.edu/~boyd/papers/pdf/cvx_ccv.pdf, 2014.

[24] J. B. Rosen and P. M. Pardalos, "Global minimization of large-scale constrained concave quadratic problems by separable programming," *Mathematical Programming*, vol. 34, no. 2, pp. 163–174, 1986.

[25] A. G. Nahapetyan, "Bilinear programming," in *Encyclopedia of Optimization*, pp. 279–282. Springer US, 2009.