

SPARSITY-PROMOTING ADAPTIVE SENSOR SELECTION FOR NON-LINEAR FILTERING

Sundeep Prabhakar Chepuri and Geert Leus

Faculty of Electrical Engineering, Mathematics, and Computer Science (EEMCS)
Delft University of Technology (TU Delft), The Netherlands
Email: {s.p.chepuri; g.j.t.leus}@tudelft.nl.

ABSTRACT

Sensor selection is an important design task in sensor networks. We consider the problem of adaptive sensor selection for applications in which the observations follow a non-linear model, e.g., target/bearing tracking. In adaptive sensor selection, based on the dynamical state model and the state estimate from the previous time step, the most informative sensors are selected to acquire the measurements for the next time step. This is done via the design of a sparse selection vector. Additionally, we model the evolution of the selection vector over time to ensure a smooth transition between the selected sensors of subsequent time steps. The original non-convex optimization problem is relaxed to a semi-definite programming problem that can be solved efficiently in polynomial time.

Index Terms— Sensor networks, adaptive sensor selection, sensor placement, non-linear measurement model, non-linear filtering, convex optimization, sparsity.

1. INTRODUCTION

Wireless sensors are widely used in a large variety of applications and services related to habitat monitoring, safety and security, logistics, and surveillance, to list a few. Each of the sensor nodes are capable of sensing, processing, and communicating to other nodes or a central unit (fusion center). These nodes are often spatially deployed to function as a network and perform a specific sensing task. We seek to extract relevant information by optimally processing the data acquired from the sensors by identifying the informative, redundant, and identical observations.

Optimal sensor placement forms an important and fundamental design task for such distributed sensor networks in which the spatial locations of the sensor nodes are optimized to guarantee a certain performance measure such as energy-efficiency, information measure, or estimation/detection accuracy. The sensor placement is even more important for location-related services such as localization, navigation, and tracking. In these problems the estimation error depends on the spatial constellation of the sensors as certain constellations deteriorate the estimation performance. The sensor placement problem can be interpreted as a sensor selection problem in which the best subset of the available sensor locations are selected subject to a specific performance constraint.

There exists a vast literature on sensor selection [1, and reference therein], and it has been applied to a wide variety of problems: dynamical systems [2–6], power grid monitoring [7], field estimation [8], and reliable sensor selection [9]. The sensor selection prob-

lem is typically formulated as an optimal experimental design problem [10] in which a scalar cost related to the mean squared error (MSE) covariance matrix is optimized to select the best subset of K sensors, where K is typically known [1, 7]. This is a non-convex Boolean optimization problem which incurs a combinatorial search over all the $\binom{K}{M}$ possible combinations, and hence, it is often simplified via convex relaxation techniques [1] or solved using greedy heuristics. Alternatively, the selection is performed based on the coherence among the columns of the system matrix as in [8, 11] which also optimizes the MSE. The optimal experimental design problems are well-studied for observations that follow a linear model for which the MSE covariance matrix has a known closed form. However, the error covariance matrix generally does not admit a known closed form for non-linear observation models, and hence the above methods cannot be used directly.

Recently, we proposed a (static) sensor selection framework for observations related to a non-linear model [12, 13]. In [13], we design a sparse selection vector that selects the most informative subset of the sensors such that a certain Cramér-Rao bound (CRB) optimality on the estimates is guaranteed. The current paper is an extension of the framework proposed in [13] for systems that obey a known dynamical model. More specifically, the proposed framework is aimed at applications related to target/bearing tracking, dynamic field estimation, and non-linear filtering in general. A similar problem has been considered in [14] for target tracking, where the sensors that yield the minimum *a posteriori* covariance are selected, with the *a posteriori* covariance matrix computed using the predicted estimate of an extended Kalman filter (EKF). However, the predicted estimates are typically far away from the true state (depending on the process/measurement noise), and hence, the sensor locations are not optimized for the true state. To alleviate this problem, different from [14], we constrain the posterior CRB (PCRB) to be satisfied for every point within a region characterized by the predicted estimates and the *a posteriori* covariance matrix, such that, the true state lies in this region with an overwhelming probability. The selection is performed via the design of a *sparse selection vector* which leads to an elegant semi-definite programming (SDP) problem. We further model the *evolution* of the selection vector in time using Gauss-Markov recursions to control the smoothness of the selection vector over time. This evolution model enables mobile sensing, and the path design of the mobile sensors.

2. STATE-SPACE MODEL AND PRELIMINARIES

We assume a non-linear measurement model for observing an unknown dynamic parameter $\theta_t \in \mathbb{R}^{N \times 1}$ corrupted by additive noise:

$$\mathbf{y}_t = \mathbf{h}_t(\theta_t) + \mathbf{n}_t, \quad (1)$$

This work was supported in part by STW under FASTCOM project (10551) and in part by NWO-STW under the VICI program (10382).

Algorithm 1 Extended Kalman filter (EKF) [16]

1. **Initialize** $t = 0$, $\hat{\boldsymbol{\theta}}_{0|0}$, and $\mathbf{P}_{0|0}$.
 2. **for** $t = 0$ to t_{\max}
 3. **given** $\mathbf{y}_t, \mathbf{h}_t(\cdot), \mathbf{A}_t, \boldsymbol{\Sigma}_t$, and \mathbf{Q}
 4. **prediction**
 - P1. $\hat{\boldsymbol{\theta}}_{t|t-1} = \mathbf{A}_t \hat{\boldsymbol{\theta}}_{t-1|t-1}$
 - P2. $\mathbf{P}_{t|t-1} = \mathbf{A}_t \mathbf{P}_{t-1|t-1} \mathbf{A}_t^T + \mathbf{Q}$
 5. **update**
 - U1. $\tilde{\mathbf{y}}_t = \mathbf{y}_t - \mathbf{h}_t(\hat{\boldsymbol{\theta}}_{t|t-1})$
 - U2. $\mathbf{S}_t = \tilde{\mathbf{H}}_t \mathbf{P}_{t|t-1} \tilde{\mathbf{H}}_t^T + \boldsymbol{\Sigma}_t$,
 where $\tilde{\mathbf{H}}_t = \left. \frac{\partial \mathbf{h}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{t|t-1}}$
 - U3. $\mathbf{L}_t = \mathbf{P}_{t|t-1} \tilde{\mathbf{H}}_t^T \mathbf{S}_t^{-1}$
 - U4. $\hat{\boldsymbol{\theta}}_{t|t} = \hat{\boldsymbol{\theta}}_{t|t-1} + \mathbf{L}_t \tilde{\mathbf{y}}_t$
 - U5. $\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{L}_t \tilde{\mathbf{H}}_t) \mathbf{P}_{t|t-1}$
 10. **end**
-

where the spatial measurements at time instance t are stacked in the measurement vector $\mathbf{y}_t = [y_{t,1}, y_{t,2}, \dots, y_{t,M}]^T \in \mathbb{R}^{M \times 1}$, and each of these measurements is a non-linear functional given by $h_{t,m}(\cdot) : \mathbb{R}^N \rightarrow \mathbb{R}$ for $m = 1, 2, \dots, M$, with the regressor $\mathbf{h}_t(\boldsymbol{\theta}_t) = [h_{t,1}(\boldsymbol{\theta}_t), \dots, h_{t,M}(\boldsymbol{\theta}_t)]^T$. The additive noise vector $\mathbf{n}_t = [n_{t,1}, n_{t,2}, \dots, n_{t,M}]^T \in \mathbb{R}^{M \times 1}$ is assumed to be zero-mean with a covariance matrix $\boldsymbol{\Sigma}_t = \text{diag}(\sigma_{t,1}^2, \sigma_{t,2}^2, \dots, \sigma_{t,M}^2) \in \mathbb{R}^{M \times M}$. The unknown parameter is assumed to obey the following dynamical model:

$$\boldsymbol{\theta}_{t+1} = \mathbf{A}_t \boldsymbol{\theta}_t + \mathbf{u}_t, \quad (2)$$

where $\mathbf{A}_t \in \mathbb{R}^{N \times N}$ is the state transition matrix, and $\mathbf{u}_t \in \mathbb{R}^{N \times 1}$ is the process noise that accounts for unmodeled dynamics. Here, we model $\mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$, where $\mathbf{Q} \in \mathbb{R}^{N \times N}$ represents the covariance matrix.

The *a posteriori* estimate $\hat{\boldsymbol{\theta}}_{t|t}$ satisfies the well-known posterior Cramér-Rao bound (PCRB) inequality given by

$$\mathbb{E}\{(\hat{\boldsymbol{\theta}}_{t|t} - \boldsymbol{\theta}_t)(\hat{\boldsymbol{\theta}}_{t|t} - \boldsymbol{\theta}_t)^T\} \geq \mathbf{F}_t^{-1}(\boldsymbol{\theta}_t), \quad (3)$$

where the posterior Fisher information matrix (FIM) $\mathbf{F}_t(\boldsymbol{\theta}_t)$ is computed using the following recursion [15]:

$$\mathbf{F}_t(\boldsymbol{\theta}_t) = (\mathbf{Q} + \mathbf{A}_t \mathbf{F}_{t-1}^{-1}(\boldsymbol{\theta}_{t-1}) \mathbf{A}_t^T)^{-1} + \tilde{\mathbf{H}}_t^T \boldsymbol{\Sigma}_t^{-1} \tilde{\mathbf{H}}_t, \quad (4)$$

with the Jacobian matrix $\tilde{\mathbf{H}}_t = \left. \frac{\partial \mathbf{h}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_t} \in \mathbb{R}^{M \times N}$. The second term in (4) is related to the log-likelihood of the measurements $\ln p(\mathbf{y}_t; \boldsymbol{\theta}_t)$. If the observations are independent, then the information measure from each observation is *additive*, which is intuitive as each independent measurement contributes some information [13, 16]. Using this property, we can further simplify (4) to arrive at

$$\mathbf{F}_t(\boldsymbol{\theta}_t) = (\mathbf{Q} + \mathbf{A}_t \mathbf{F}_{t-1}^{-1}(\boldsymbol{\theta}_{t-1}) \mathbf{A}_t^T)^{-1} + \sum_{m=1}^M \mathbf{F}_{t,m}(\boldsymbol{\theta}_t), \quad (5)$$

where $\mathbf{F}_{t,m} = \frac{1}{\sigma_{t,m}^2} \left(\frac{\partial h_{t,m}(\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}^T} \right) \left(\frac{\partial h_{t,m}(\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}^T} \right)^T$.

The *a posteriori* estimate $\hat{\boldsymbol{\theta}}_{t|t}$ can be obtained using any of the non-linear filters (e.g., extended Kalman filter (EKF), unscented Kalman filter (UKF), or particle filters). For the sake of completeness, the EKF is briefly summarized in Algorithm 1.

3. OPTIMIZATION PROBLEM

The *adaptive sensor selection* problem can be interpreted as a problem to select the best subset out of the M available sensors to acquire measurements for time step t such that a certain accuracy on the estimate $\hat{\boldsymbol{\theta}}_{t|t}$ is guaranteed. In order to perform selection, we introduce a *selection vector* $\mathbf{w}_t = [w_{t,1}, w_{t,2}, \dots, w_{t,M}]^T \in \{0, 1\}^M$, and modify the model in (1) to

$$y_{t,m} = w_{t,m} h_{t,m}(\boldsymbol{\theta}_t) + n_{t,m}, \quad m = 1, 2, \dots, M,$$

where $w_{t,m}$ is a *virtual Boolean selection parameter*, i.e., $w_{t,m} \in \{0, 1\}$, $m = 1, 2, \dots, M$, where $w_{t,m} = (0)1$ indicates that the sensor is (not) selected.

Remark 1 (Active sensing). *In active sensing, the sensors transmit probing signals (e.g., ranging signal). The selection parameter $w_{t,m}$ for active sensing is a soft parameter used for joint selection and resource allocation [12], i.e., $w_{t,m} \in [0, 1]$ is a resource (e.g., ranging energy) normalized to the maximum prescribed value, and hence, it is dimensionless.*

The FIM for the modified measurement model will be

$$\mathbf{F}_t(\mathbf{w}_t, \boldsymbol{\theta}_t) = \overbrace{(\mathbf{Q} + \mathbf{A}_t \mathbf{F}_{t-1}^{-1}(\boldsymbol{\theta}_{t-1}) \mathbf{A}_t^T)^{-1}}^{\mathbf{J}_{t-1}(\boldsymbol{\theta}_{t-1})} + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}(\boldsymbol{\theta}_t), \quad (6)$$

where $\mathbf{J}_{t-1}(\boldsymbol{\theta}_{t-1})$ is basically the prior FIM.

As a performance measure, we constrain the estimation error $\boldsymbol{\epsilon}_t = \hat{\boldsymbol{\theta}}_{t|t} - \boldsymbol{\theta}_t$ to be within an origin centered circle of radius R_e with a probability higher than P_e , i.e., $\Pr(\|\boldsymbol{\epsilon}_t\|_2 \leq R_e) \geq P_e$. The true states $\boldsymbol{\theta}_t$ and $\boldsymbol{\theta}_{t-1}$ are not known in practice, unless the process noise is zero (deterministic trajectory). In order to select the measurements for the next time step t , the prior Fisher \mathbf{J}_{t-1} is evaluated at $\hat{\boldsymbol{\theta}}_{t-1|t-1}$ and the *predicted estimate* $\hat{\boldsymbol{\theta}}_{t|t-1}$ (cf. Algorithm 1) is assumed as the true state in which case the FIM in (4) is nothing but the inverse of the *a posteriori* covariance matrix $\mathbf{P}_{t|t}^{-1}$. In other words, the FIM in (6) is approximated to

$$\mathbf{F}_t(\mathbf{w}_t, \boldsymbol{\theta}_t) \approx \mathbf{J}_{t-1}(\hat{\boldsymbol{\theta}}_{t-1|t-1}) + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}(\hat{\boldsymbol{\theta}}_{t|t-1}). \quad (7)$$

A sufficient condition to satisfy this accuracy constraint is given by the inequality $\lambda_{\min}(\mathbf{F}_t(\mathbf{w}_t, \boldsymbol{\theta}_t)) \geq \lambda$, which can be alternatively expressed as the following linear matrix inequality (LMI)

$$\mathbf{J}_{t-1}(\hat{\boldsymbol{\theta}}_{t-1|t-1}) + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}(\hat{\boldsymbol{\theta}}_{t|t-1}) \succeq \lambda \mathbf{I}_N, \quad (8)$$

where $\lambda = \frac{N}{R_e^2} \log\left(\frac{1}{1-P_e}\right)$ [13]. Here, $\lambda_{\min}(\mathbf{A})$ denotes the minimum eigenvalue of a symmetric matrix \mathbf{A} . The solution set of \mathbf{w}_t satisfying this LMI is convex [10].

For a non-zero measurement noise the predicted state will always be away from the true state. Hence, we constrain (8) for every point within a certain domain \mathcal{T}_t determined by $\hat{\boldsymbol{\theta}}_{t|t-1}$ and $\mathbf{P}_{t|t-1}$. More specifically, \mathcal{T}_t is a circle having a radius $5\sqrt{\text{tr}\{\mathbf{P}_{t|t-1}\}}$ centered around $\hat{\boldsymbol{\theta}}_{t|t-1}$. Since $\hat{\boldsymbol{\theta}}_{t|t-1} \sim \mathcal{N}(\boldsymbol{\theta}_t, \mathbf{P}_{t|t-1})$, the *true state* lies within a circle of radius 5σ with an overwhelming probability, where $\sigma = \sqrt{\text{tr}\{\mathbf{P}_{t|t-1}\}}$ is the mean standard deviation. Similarly,

the past estimate lies within the domain \mathcal{T}_{t-1} which is a circle centered around $\hat{\boldsymbol{\theta}}_{t-1|t-1}$ having a radius of $5\sqrt{\text{tr}\{\mathbf{P}_{t-1|t-1}\}}$. The adaptive sensor selection problem can be briefly stated as follows.

Problem statement (Adaptive sensor selection). *Given the measurement process $h_{t,m}(\boldsymbol{\theta}_t)$, $m = 1, \dots, M$, find a vector $\mathbf{w}_t \in \{0, 1\}^M$ for each time step t that selects the minimum number of most informative sensors that satisfies the accuracy constraint $\mathbf{J}_{t-1}(\boldsymbol{\theta}_{t-1}) + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}(\boldsymbol{\theta}_t) \succeq \lambda \mathbf{I}_N$, $\forall \boldsymbol{\theta}_{t-1} \in \mathcal{T}_{t-1}$, and $\forall \boldsymbol{\theta}_t \in \mathcal{T}_t$.*

The adaptive sensor selection can be formulated as the design of a selection vector which can be expressed as the following optimization problem

$$\arg \min_{\mathbf{w}_t} \|\mathbf{w}_t\|_0 \quad (9a)$$

$$\text{s.t. } \mathbf{J}_{t-1}(\boldsymbol{\theta}_{t-1}) + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}(\boldsymbol{\theta}_t) \succeq \lambda \mathbf{I}_N, \quad (9b)$$

$$\forall \boldsymbol{\theta}_{t-1} \in \mathcal{T}_{t-1}, \forall \boldsymbol{\theta}_t \in \mathcal{T}_t,$$

$$\mathbf{w}_t \in \{0, 1\}^M, \quad (9c)$$

where the ℓ_0 (-quasi) norm refers to the number of non-zero entries of \mathbf{w}_t . This is non-convex Boolean optimization problem. The prior FIM is independent of the optimization variable, hence, it is sufficient to use a worst-case $\boldsymbol{\theta}_{t-1} \in \mathcal{T}_{t-1}$, more specifically, $\tilde{\boldsymbol{\theta}}_{t-1} = \arg \min_{\boldsymbol{\theta}_{t-1} \in \mathcal{T}_{t-1}} \lambda_{\min}(\mathbf{J}_{t-1}(\boldsymbol{\theta}_{t-1}))$.

4. ADAPTIVE SENSOR SELECTION

In this section, we provide an algorithm to solve the proposed optimization problem and also model the evolution of \mathbf{w}_t in time.

4.1. Relaxed sensor selection

The optimization problem in (9) is non-convex due to the ℓ_0 (-quasi) norm cost function and the Boolean constraint. We use the traditional best convex surrogate for the ℓ_0 (-quasi) norm based on the ℓ_1 -norm heuristic, and the Boolean constraint is relaxed to the convex box constraint $[0, 1]^M$. Due to the box constraint (9c), the ℓ_1 -norm will simply be an affine function $\mathbf{1}^T \mathbf{w}_t$. The relaxed optimization problem is given as

$$\hat{\mathbf{w}}_t = \arg \min_{\mathbf{w}_t \in [0, 1]^M} \mathbf{1}^T \mathbf{w}_t \quad (10a)$$

$$\text{s.t. } \mathbf{J}_{t-1}(\tilde{\boldsymbol{\theta}}_{t-1}) + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}(\boldsymbol{\theta}_t) \succeq \lambda \mathbf{I}_N, \quad \forall \boldsymbol{\theta}_t \in \mathcal{T}_t. \quad (10b)$$

Remark 2 (Relaxed active sensor selection). *The relaxed active sensor selection problem takes the same form as in (10). In fact, minimizing the ℓ_1 -norm in active sensor selection minimizes the overall network resources (e.g., overall network ranging energy).*

The relaxed optimization problem is a standard SDP problem that can be solved efficiently in polynomial time using interior-point methods or off-the-shelf solvers like SeDuMi [17]. The computational complexity of solving (10) is of the order of $O(M^3)$ [13].

4.2. Smoothness model

We now model the evolution of \mathbf{w}_t in time. A smooth evolution of the selection vector is important for mobile sensing (or multi-static

target tracking for instance). The evolution of the selection vector is modeled as a Gauss-Markov recursion [18]

$$\mathbf{w}_t = \mathbf{B}_t \mathbf{w}_{t-1} + \mathbf{e}_t, \quad (11)$$

where $\mathbf{B}_t \in \mathbb{R}^{M \times M}$ is a (right-)stochastic transition matrix, and $\mathbf{e}_t = [e_{t,1}, \dots, e_{t,M}]^T \in \mathbb{R}^M$ is the process noise vector. The smoothness depends on the design of the matrix \mathbf{B}_t . Initializing the recursion with an arbitrary $\mathbf{w}_0 \neq \mathbf{0}$, and running (11) for a large enough t , leads to non-sparse \mathbf{w}_t . In order to incorporate the smoothing effect of the selection vector between subsequent time instances, we use the sparse estimate $\hat{\mathbf{w}}_{t-1}$ instead of \mathbf{w}_{t-1} . The optimization problem taking into account the smoothness is given as

$$\hat{\mathbf{w}}_{t,\text{sm}} = \arg \min_{\mathbf{w}_t \in [0, 1]^M} \|\mathbf{w}_t\|_1 + \mu \|\mathbf{w}_t - \mathbf{B}_t \hat{\mathbf{w}}_{t-1,\text{sm}}\|_2^2 \quad (12a)$$

$$\text{s.t. } \mathbf{J}_{t-1}(\tilde{\boldsymbol{\theta}}_{t-1}) + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}(\boldsymbol{\theta}_t) \succeq \lambda \mathbf{I}_N, \quad \forall \boldsymbol{\theta}_t \in \mathcal{T}_t, \quad (12b)$$

where μ is the smoothness controlling parameter. Smoothness in the selection vector ensures an easy hand-off between the selected sensors. The approximate Boolean solution can be then recovered by randomization rounding [13].

5. SIMULATIONS

To test the proposed algorithms, we use CVX [19] which internally calls SeDuMi [17]. We illustrate the problem of *adaptive sensor selection* for target tracking based on the distance measurements. Using the selected sensors, the state vector $\boldsymbol{\theta}_t = [\mathbf{x}_t^T, \dot{\mathbf{x}}_t^T]^T \in \mathbb{R}^{4 \times 1}$ is estimated at a discrete time instance t . Here, $\mathbf{x}_t \in \mathbb{R}^{2 \times 1}$ is the target position vector, and $\dot{\mathbf{x}}_t \in \mathbb{R}^{2 \times 1}$ is the velocity vector. The non-linear measurement model is given by

$$h_{t,m}(\boldsymbol{\theta}_t) := d_{t,m} = \|\mathbf{x}_t - \mathbf{a}_m\|_2, \quad m = 1, 2, \dots, M,$$

where $\mathbf{a}_m \in \mathbb{R}^{2 \times 1}$ is the known position of the m th sensor.

We do not make measurements on the velocity, hence, we constrain only the FIM related to the distance measurements. Assuming that the FIM is composed of the following submatrices $\mathbf{F}_t(\boldsymbol{\theta}_t) = \begin{bmatrix} \mathbf{F}_t^{(\text{xx})} & \mathbf{F}_t^{(\text{x}\dot{\mathbf{x}})} \\ \mathbf{F}_t^{(\dot{\mathbf{x}}\mathbf{x})} & \mathbf{F}_t^{(\dot{\mathbf{x}}\dot{\mathbf{x}})} \end{bmatrix}$, then using the Schur complement, the *a posteriori* estimate $\hat{\mathbf{x}}_{t|t}$ satisfies the following PCRB inequality $\mathbb{E}\{(\hat{\mathbf{x}}_{t|t} - \mathbf{x}_t)(\hat{\mathbf{x}}_{t|t} - \mathbf{x}_t)^T\} \geq [\mathbf{F}_t(\mathbf{x})]^{-1}$, where

$$\mathbf{F}_t(\mathbf{x}) = \mathbf{F}_t^{(\text{xx})} - \mathbf{F}_t^{(\text{x}\dot{\mathbf{x}})} [\mathbf{F}_t^{(\dot{\mathbf{x}}\dot{\mathbf{x}})}]^{-1} \mathbf{F}_t^{(\dot{\mathbf{x}}\mathbf{x})}.$$

Since, we do not make any measurements related to $\dot{\mathbf{x}}$, the FIM related to measurements $\mathbf{F}_{t,m}$, $m = 1, \dots, M$, will have a non-zero upper-left 2×2 submatrix denoted by $\mathbf{F}_{t,m}^{(\text{xx})}$, and remaining submatrices will all be zeros. Similarly, assume that the prior FIM is also

composed of the following submatrices $\mathbf{J}_t = \begin{bmatrix} \mathbf{J}_t^{(\text{xx})} & \mathbf{J}_t^{(\text{x}\dot{\mathbf{x}})} \\ \mathbf{J}_t^{(\dot{\mathbf{x}}\mathbf{x})} & \mathbf{J}_t^{(\dot{\mathbf{x}}\dot{\mathbf{x}})} \end{bmatrix}$.

Using the Schur complement and some straightforward matrix properties, we arrive at

$$\mathbf{F}_t^{(\text{xx})}(\mathbf{w}_t, \mathbf{x}_t) = \tilde{\mathbf{J}}_{t-1}(\tilde{\boldsymbol{\theta}}_{t-1}) + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}^{(\text{xx})}(\mathbf{x}_t),$$

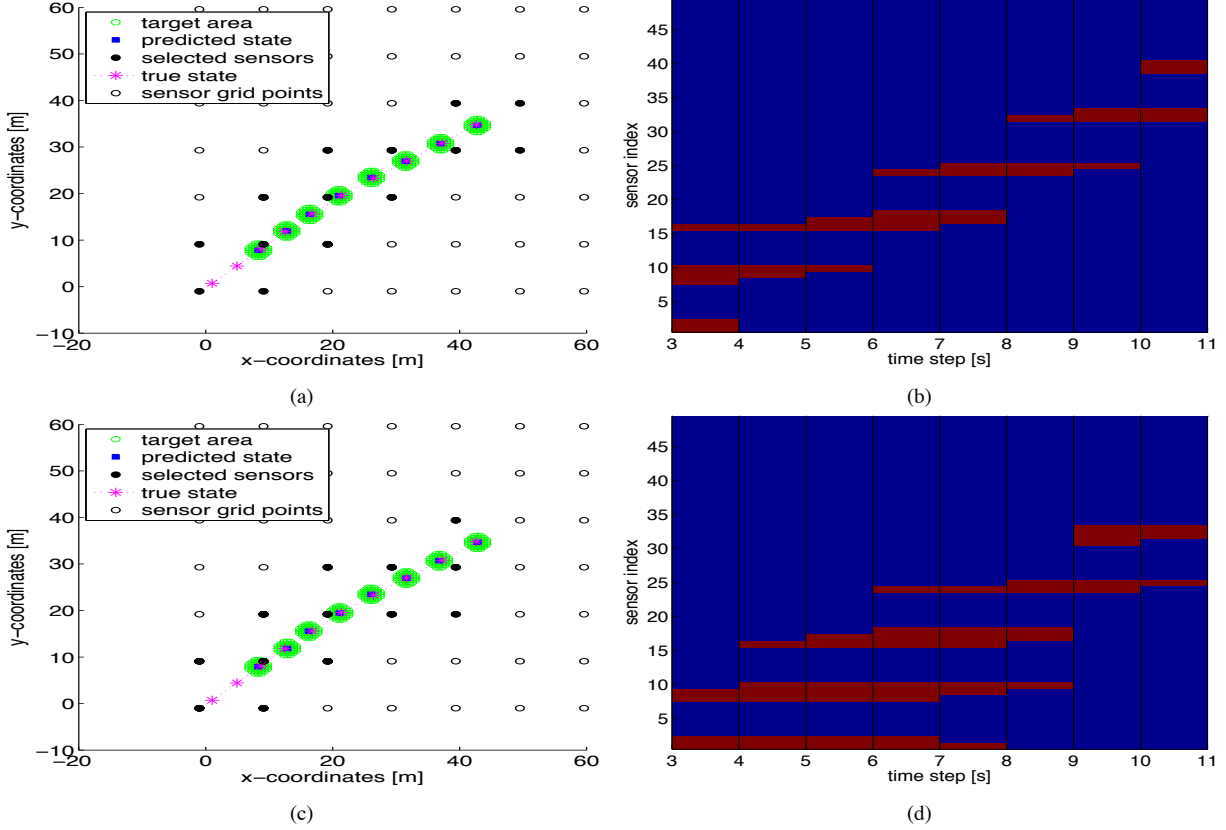


Fig. 1: Adaptive sensor selection for target tracking based on range measurements for the time interval $(3\tau_s, 10\tau_s)$. (a) and (c) trajectory of the true state for a certain realization, predicted estimate from the EKF, and the target area \mathcal{T}_t . (b) and (d) sensor activation time pattern for $\mu = 0$ (without smoothness) and $\mu = 0.5$ (with smoothness), respectively.

where $\tilde{\mathbf{J}}_{t-1} = \mathbf{J}_{t-1}^{(xx)} - \mathbf{J}_{t-1}^{(xx)}[\mathbf{J}_{t-1}^{(xx)}]^{-1}\mathbf{J}_{t-1}^{(xx)} \in \mathbb{R}^{2 \times 2}$. The LMI constraints in (10b) for the case when only some of the state parameters are measured are given by

$$\tilde{\mathbf{J}}_{t-1}(\tilde{\boldsymbol{\theta}}_{t-1}) + \sum_{m=1}^M w_{t,m} \mathbf{F}_{t,m}^{(xx)}(\mathbf{x}_t) \succeq \lambda \mathbf{I}_2, \quad \forall \mathbf{x}_t \in \mathcal{T}_t,$$

where \mathcal{T}_t is a circle with center $\hat{\mathbf{x}}_{t|t-1}$ and radius $5\sqrt{\text{tr}\{\mathbf{P}_{t|t-1}^{(xx)}\}}$.

Here, $\mathbf{P}_{t|t-1}^{(xx)}$ is the 2×2 upper-left submatrix of $\mathbf{P}_{t|t-1}$. We simply use $\tilde{\boldsymbol{\theta}}_{t-1} = \hat{\boldsymbol{\theta}}_{t-1|t-1}$. The parameters determining the accuracy are set to $R_e = 10$ cm and $P_e = 0.95$ to compute $\lambda = \frac{2}{R_e^2} \log(\frac{1}{1-P_e})$.

We consider an area of 50×50 square meter with $M = 36$ equally spaced sensor grid points as shown in Fig. 1a. We use the following parameters for simulations: observation time $T = 20$ s,

$$\mathbf{A}_t = \begin{bmatrix} 1 & 0 & \tau_s & 0 \\ 0 & 1 & 0 & \tau_s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{Q} = 10^{-2} \begin{bmatrix} \frac{\tau_s^3}{3} & 0 & \frac{\tau_s^2}{2} & 0 \\ 0 & \frac{\tau_s^3}{3} & 0 & \frac{\tau_s^2}{2} \\ \frac{\tau_s^2}{2} & 0 & \tau_s & 0 \\ 0 & \frac{\tau_s^2}{2} & 0 & \tau_s \end{bmatrix},$$

and, sampling time $\tau_s = 2.5$ s. The dynamic model is initialized with $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{1}_2, 2.778\mathbf{I}_2)$ and $\dot{\mathbf{x}}_0 \sim \mathcal{N}(2\mathbf{1}_2, 0.01\mathbf{I}_2)$ to emulate the target heading towards the north-east direction. The EKF is initialized with $\hat{\boldsymbol{\theta}}_{0|0} = \mathbf{0}$, and $\mathbf{P}_{0|0} = 1000\mathbf{I}_2$. The stochastic matrix \mathbf{B}_t is designed such that the transition to the one-hop sensor grid

points and staying in the current state takes equal probabilities. In other words, in Fig. 1a, the corner most grid point has 3 one-hop neighbors, hence, it can move to any of these one-hop neighbors each with a probability of $1/4$ or stay in the current state with a probability of $1/4$. The selection vector is initialized with four selected sensors on the corner most grid points in the south-west (see Fig. 1c). We assume a distance dependent noise model such that $n_m \sim \mathcal{N}(0, \sigma_{t,m}^2)$ with $\sigma_{t,m}^2 := \frac{\sigma^2}{d_{t,m}^2}$, where σ^2 is the nominal noise. We use $\sigma^2 = 10^{-5}$.

For a practical implementation of the algorithm, we discretize the target area \mathcal{T}_t , and each of the target grid points results in an LMI constraint. Here, we discretize \mathcal{T}_t with 25 points as shown in Fig. 1a and Fig. 1c (indicated as the target area in green color). For the sake of easy visibility, we plot the results in the time interval $(3\tau_s, 10\tau_s)$, as the target area is very large for initial estimates. Even though the predicted estimates are not necessarily on top of the true state, the true location will be within the target area with an overwhelming probability. Due to the assumed path-loss model, the sensors close to the target area are selected. The sensor activation time pattern without ($\mu = 0$) and with ($\mu > 0$) the smoothness model is shown in Fig. 1b and Fig. 1d, respectively. The Boolean solution is obtained by simply rounding all the non-zero entries of $\hat{\mathbf{w}}$ to one. The number of selected sensors with $\mu > 0$ is larger as compared to the case with $\mu = 0$. However, with the smoothness model, the sensors stay active for a longer duration ensuring a smooth hand-off.

6. 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] Y. Fu, Q. Ling, and Z. Tian, "Distributed sensor allocation for multi-target tracking in wireless sensor networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3538–3553, 2012.
- [3] M.L. Hernandez, T. Kirubarajan, and Y. Bar-Shalom, "Multisensor resource deployment using Posterior Cramer-Rao bounds," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 2, pp. 399–416, 2004.
- [4] F. Jiang, J. Chen, and A. L. Swindlehurst, "Linearly reconfigurable Kalman filtering for a vector process," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013.
- [5] Long Zuo, Ruixin Niu, and P.K. Varshney, "Posterior CRLB based sensor selection for target tracking in sensor networks," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2007, vol. 2, pp. II–1041–II–1044.
- [6] Pengcheng Zhan, D.W. Casbeer, and A.L. Swindlehurst, "Adaptive mobile sensor positioning for multi-static target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 120–132, 2010.
- [7] V. Kekatos, G.B. Giannakis, and B. Wollenberg, "Optimal placement of phasor measurement units via convex relaxation," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1521–1530, 2012.
- [8] J. Ranieri, A. Chebira, and M. Vetterli, "Near-optimal sensor placement for inverse problems," *ArXiv*, vol. abs/1305.6292, 2013.
- [9] V. Kekatos and G.B. Giannakis, "From sparse signals to sparse residuals for robust sensing," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3355–3368, 2011.
- [10] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, NY, USA, 2004.
- [11] A. Carmi, "Sensor scheduling via compressed sensing," in *Proc. of 13th Conference on Information Fusion (FUSION)*, 2010, pp. 1–8.
- [12] S.P. Chepuri, G. Leus, and A.-J. van der Veen, "Sparsity-exploiting anchor placement for localization in sensor networks," in *Proc. of the 21st Eur. Signal Process. Conf (EUSIPCO)*, Sept. 2013.
- [13] S.P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Trans. Signal Process.* (Submitted), Sep. 2013.
- [14] 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, 2012.
- [15] P. Tichavsky, C.H. Muravchik, and A. Nehorai, "Posterior Cramer-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, 1998.
- [16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [17] Jos F. Sturm, "Using SeDuMi 1.02, a matlab toolbox for optimization over symmetric cones," 1998.
- [18] S. Farahmand, G.B. Giannakis, G. Leus, and Z. Tian, "Tracking target signal strengths on a grid using sparsity," *arXiv preprint arXiv:1104.5288*, 2011.
- [19] [Online], "CVX: Matlab software for disciplined convex programming, version 2.0 beta," <http://cvxr.com/cvx>, Sep. 2012.