

Sensor Selection and Design for Binary Hypothesis Testing in the Presence of a Cost Constraint

Berkay Oymak, Berkan Dulek, and Sinan Gezici

Abstract—We consider a sensor selection problem for binary hypothesis testing with cost-constrained measurements. Random outputs related to a parameter vector of interest are assumed to be generated by a linear system corrupted with Gaussian noise. The aim is to decide on the state of the parameter vector based on a set of measurements collected by a limited number of sensors. The cost of each sensor measurement is determined by the number of amplitude levels that can reliably be distinguished. By imposing constraints on the total cost and the maximum number of sensors that can be employed, a sensor selection problem is formulated in order to maximize the detection performance for binary hypothesis testing. By characterizing the form of the solution corresponding to a relaxed version of the optimization problem, a computationally efficient algorithm with near optimal performance is proposed. In addition to the case of fixed sensor measurement costs, we also consider the case where they are subject to design. In particular, the problem of allocating the total cost budget to a limited number of sensors is addressed by designing the measurement accuracy (i.e., the noise variance) of each sensor to be employed in the detection procedure. The optimal solution is obtained in closed form. Numerical examples are presented to corroborate the proposed methods.²

Index Terms— Detection, sensor selection, cost constraint.

I. INTRODUCTION

With the increasing availability of sensors, performance of detection and estimation methods based on information gathered from multiple sensors has become more important. While various optimality criteria, such as Bayesian detection and estimation, Neyman-Pearson detection, and minimum variance unbiased estimation, are investigated extensively in the literature [1], additional challenges arise from practical considerations in sensor networks. These challenges are commonly related to limited resources such as power, bandwidth, and number and quality of sensors in the network.

There exist several studies in the literature that focus on the objective of maximizing detection/estimation performance in sensor networks while satisfying system-level constraints related to communication bandwidth, transmission power, and sensor costs [2]–[11]. In [2], the optimal cost allocation problem in a sensor network is investigated for centralized and decentralized detection, where it is assumed that sensors with higher costs provide less noisy measurements. Detection performance is assessed according to Bayesian, Neyman-Pearson and J -divergence criteria, and optimal cost allocation strategies are provided. The works in [3] and [4] address performance of parameter estimation with cost-constrained measurements in sensor networks. In [3], the problem of optimal cost allocation to measurement devices is investigated in order to maximize the average Fisher information about a vector parameter. A closed-form solution is obtained for the case of Gaussian noise. On the other hand, in [4], the authors

focus on the minimization of the total measurement cost while satisfying several estimation accuracy constraints. Closed form solutions are obtained when the system measurement matrix is invertible and the noise is Gaussian. Extensions that take into account the uncertainty on the system measurement matrix are also analyzed. In [5], a distributed detection problem in the presence of transmission power constraints on sensor nodes and communication bandwidth constraints between sensors and a fusion center is considered. By assuming independent and identically distributed (i.i.d.) sensor measurements, multiple and parallel access channel models are investigated under bandwidth constraints. An asymptotically optimal decision strategy is obtained for a multiple access channel, where each sensor transmits its local likelihood ratio with constant power to a fusion center. In [6], a detection problem in sensor networks is investigated, where costs due to performing measurements at each sensor as well as those due to transmissions from sensors to a fusion center are considered. The solution under such cost constraints leads to a randomized scheme that specifies when sensors should transmit data and make measurements. Examples in which the joint optimization over all sensor nodes decouples into individual optimizations at each sensor node are presented.

In addition to communication bandwidth and transmission power constraints in sensor networks, limitations on the number of actively used sensors are also important. In fact, the number of sensors activated simultaneously has direct implications on both communication bandwidth and total power consumption. Commonly, it is desirable to constrain the number of active sensors without sacrificing performance. Thus, the sensor selection problem arises naturally in resource constrained sensor networks. Some applications of sensor selection are sensor coverage [12], target localization [13], [14], discrete event systems [15], Internet of Things [16], and sensor placement [17], [18]. The information theory framework is also employed as a basis for sensor selection in [19]–[22]. To highlight main aspects and challenges in the sensor selection problem, we summarize several related papers in the literature. In [23], sensor selection is carried out to determine the most informative subset of sensors in a wireless sensor network (WSN) for a detection problem. It is shown that the sensor selection problem is NP-hard, and computationally efficient algorithms are provided to obtain near optimal solutions under Kullback-Leibler (KL) and Chernoff criteria. In [24], a sensor selection problem is formulated for parameter estimation under Gaussian noise. An intuitive method based on convex relaxation is described in order to approximately solve the problem. Numerical experiments are provided to demonstrate the proposed method. Also, additional constraints to the sensor selection problem are outlined for which the proposed method remains effective. An entropy based sensor selection approach in the context of target localization is proposed in [25]. The sensor selection problem is addressed to minimize the estimation error in target localization in [26], where the authors formulate an optimization problem with a constraint on the

B. Oymak and S. Gezici are with the Dept. of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey, Tel: +90-312-290-3139, Fax: +90-312-266-4192, E-mails: {berkayo,gezici}@ee.bilkent.edu.tr. B. Dulek is with the Dept. of Electrical and Electronics Engineering, Hacettepe University, 06800, Ankara, Turkey, E-mail: berkan@ee.hacettepe.edu.tr.

number of sensors employed for measuring the target position. An algorithm to obtain an approximate solution is presented, and it is shown that the estimation error is not higher than the twice of the minimum achievable error. (The reader is referred to [27] for commonly employed sensor selection schemes in target tracking and localization.) The study in [28] focuses on the optimal design of a WSN using different classes of sensors, where each class of sensors has a cost and measurement characteristic. The aim is to find the optimal number of sensors to choose from each class so that the detection performance based on the symmetric KL divergence is maximized. It is shown that the KL divergence and the number of sensors of each class are linearly related. The results indicate that it is optimal to choose all sensors from the class with the best performance to cost ratio. In [29], a sensor selection problem is formulated for state estimation of dynamic systems such as those found in large space structures. In the problem statement, it is required to select a measurement subsystem out of several candidates. A sensor selection policy is presented as an on-line algorithm which selects the measurement subsystem that provides the maximum information along the principal state space direction associated with the largest estimation error. The work in [30] investigates a failure diagnosis system, in which each subset of sensors can be used to make a diagnosis observation with a certain cost and failure detection probability. It is aimed to determine the cheapest combination of sensors that guarantee a certain probability of failure detection when a certain number of observations are made. A method that identifies this subset with the minimum number of trials is proposed. In [31], spectrum sensing with multiple sensors is considered. The aim is to find a subset that guarantees reliable sensing performance. It is pointed out that it is crucial to select sensors that experience uncorrelated fading; meaning that they should be spatially separated. Assuming limited knowledge on sensor positions, iterative suboptimal algorithms that are based on correlation measure, estimated sensor position, and radius information are proposed and compared with random sensor selection. In [32], a dynamic sensor selection algorithm is devised for a wearable sensor network that performs real-time activity recognition. It is shown that by utilizing the selection algorithm, a desired level of classification accuracy is sustained while increasing network lifetime significantly. In [33], the sensor selection problem is considered in the distributed detection framework. One particular application is to detect hot spots (i.e., the areas where the temperature exceeds a certain threshold) of a multi-core processor for subsequent control actions. In the theoretical formulation of the problem, the authors aim at minimizing the number of data acquired from the sensors while maintaining a desired detection performance, expressed in terms of the Bayesian probability of error as well as miss detection and false alarm probabilities. The major distinction between their work and ours is that in selecting the best subset of sensor data samples that result in a desired detection probability, they do not distinguish between the costs of different sensor measurements. An explicit sensor measurement cost function is not employed. As a result, the sensing, storage, transmission, and processing costs are assumed to be identical for all the sensors as each sensor enters into the optimization function via the ℓ_0 -quasi norm which only accounts for the presence/absence of the data from a particular sensor.

As noted from the aforementioned literature, optimal resource allocation to improve detection performance in cost constrained sensor networks is considered in various studies.

However, an in-depth analysis of the sensor selection problem under a cost constraint related to the measurement quality of the employed sensors is lacking in the literature. In this work, we propose an optimal sensor selection method that minimizes the Bayes risk while satisfying a total cost constraint related to the measurement accuracy of the sensors. As in most sensor selection problems, the corresponding optimization problem emerges as a zero-one integer linear programming problem [34], which is known to be NP-complete [35]. Although there exist methods to find an optimal solution to such problems, such as the branch and bound method given in [35], [36], they turn out to be practically ineffective in terms of the running time unless $\binom{N_s}{K}$ is small, where N_s is the number of available observations and K is the number of sensors (equivalently, the effective number of observations that can be measured by the sensors). In this paper, we first relax the binary constraint (that a sensor is either selected or not) into a linear constraint, which leads to a linearly constrained linear optimization problem. Then, the form of the solution to the relaxed problem is characterized theoretically and a numerical algorithm with reduced computational complexity is presented to obtain the solution. Based on the solution of the relaxed problem, a feasible set of sensors are selected using a local optimization approach. The effectiveness of the proposed approach is demonstrated by depicting the performance difference between the bound provided by the solution of the relaxed problem and the objective value attained by the proposed sensor selection algorithm. Also, comparisons with alternative heuristic approaches are provided to highlight the efficiency of our method. As an extension, we also consider the case where sensors (i.e., their noise variances) are subject to design, and a joint sensor selection and design method is developed. The optimal solution to this joint problem is given in closed form, where the parameter of the solution can be determined by a practical algorithm. Numerical examples are presented to illustrate the effectiveness of the proposed approach.

The main contributions of this paper can be summarized as follows:

- Based on the cost definition in [37], we propose a cost-constrained sensor selection problem for binary hypothesis testing to minimize the Bayes risk under a linear system model corrupted with Gaussian noise.
- It is shown that the solution to the linearly relaxed problem contains at most two non-integer elements and an approximate solution with near optimal performance is developed based on this observation.
- The optimal solution is obtained in closed form when the accuracy, measured by the noise variances, of individual sensors is also subject to design.

The rest of this paper is organized as follows. In Section II, we present the measurement model for the a generic linear system. In Section III, an approximate solution is developed for determining which sensors should be employed to collect the measurements when the cost of each sensor measurement is given. In Section IV, we analyze the problem of joint sensor selection and design, and characterize the optimal solution. In Section V, we provide numerical examples to evaluate the performance of the proposed methods. We conclude with some remarks in Section VI.

II. SYSTEM MODEL

Let $\Theta \in \mathbb{R}^L$ represent a parameter vector of interest. This parameter vector is processed by a noisy linear system and the

corresponding outputs are expressed as

$$x_i = \mathbf{h}_i^T \Theta + n_i, \quad i = 1, \dots, N_s \quad (1)$$

where n_i is the noise in the i th system output and \mathbf{h}_i is an $L \times 1$ vector representing the coefficients of the linear system related to output i . The output of the linear system in (1) can be measured by N_s potential sensors as follows:

$$y_i = x_i + m_i, \quad i = 1, \dots, N_s \quad (2)$$

where m_i is the measurement noise of the i th sensor.

In a more compact manner, the system outputs in (1) and the potential measurements in (2) can be expressed as

$$\mathbf{x} = \mathbf{H}^T \Theta + \mathbf{n} \quad \text{and} \quad \mathbf{y} = \mathbf{x} + \mathbf{m}, \quad (3)$$

respectively, where $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{N_s}]$ is the $L \times N_s$ system matrix, $\mathbf{x} = [x_1, x_2, \dots, x_{N_s}]^T$, $\mathbf{n} = [n_1, n_2, \dots, n_{N_s}]^T$, $\mathbf{y} = [y_1, y_2, \dots, y_{N_s}]^T$, and $\mathbf{m} = [m_1, m_2, \dots, m_{N_s}]^T$. As in [4] and [24], the noise components are modeled as independent Gaussian random variables with zero mean, that is, $n_i \sim \mathcal{N}(0, \sigma_{n_i}^2)$ and $m_i \sim \mathcal{N}(0, \sigma_{m_i}^2)$ for $i = 1, \dots, N_s$. In the vector notation, $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ and $\mathbf{m} \sim \mathcal{N}(\mathbf{0}, \Sigma_m)$, where $\Sigma = \text{diag}\{\sigma_{n_1}^2, \sigma_{n_2}^2, \dots, \sigma_{n_{N_s}}^2\}$ and $\Sigma_m = \text{diag}\{\sigma_{m_1}^2, \sigma_{m_2}^2, \dots, \sigma_{m_{N_s}}^2\}$. In addition, it is assumed that the measurement noise \mathbf{m} is independent of the system noise \mathbf{n} .

We place no restrictions on the system matrix \mathbf{H} . For example, in wave optics, \mathbf{H} may represent the real equivalent of the fractional Fourier transform matrix, which provides a convenient approximation of the Fresnel diffraction integral that relates the measured field to the unknown field [37]. \mathbf{H} is dictated by the physics of the underlying phenomenon and \mathbf{n} is an intrinsic part of the relation between \mathbf{x} and Θ , which we have no control over. On the other hand, \mathbf{m} is the noise associated with the employed measurement device and depends on our choices.

III. SENSOR SELECTION FOR BINARY HYPOTHESIS TESTING

Since N_s can be very large in various scenarios, it is an important problem to choose a subset of the N_s system outputs for measurement in an optimal manner, which is called the sensor selection problem in the literature [13], [15], [23], [24], [26], [31], [32], [38]. In particular, the aim is to optimize a certain performance metric while making measurements with at most K out of N_s potential sensors. To represent the selection operation, we define a *selection vector* $\mathbf{z} = [z_1, z_2, \dots, z_{N_s}]^T$ that specifies whether the i th sensor is selected (i.e., $z_i = 1$ if the i th sensor is selected and $z_i = 0$ otherwise). We denote the number of selected sensors as k , that is, $\mathbf{1}^T \mathbf{z} = k$, where $\mathbf{1}$ represents a column vector of ones and $k \leq K$. For notational convenience, we also introduce an injective function $f: \{1, 2, \dots, k\} \rightarrow \{1, 2, \dots, N_s\}$, where $f(i)$ denotes the index of the i th selected sensor. Then, we construct a $k \times N_s$ *selection matrix* \mathbf{Z} , in which k of the columns are unit vectors $\mathbf{e}_{N_s,1}, \mathbf{e}_{N_s,2}, \dots, \mathbf{e}_{N_s,k}$ ($\mathbf{e}_{j,i}$ is defined as a column vector of length j and it has a 1 at the i th position and 0 elsewhere), and the other columns are zero vectors. In the selection matrix \mathbf{Z} , the column indices of the unit vectors specify the selected sensors. It is noted that \mathbf{Z} can be constructed from \mathbf{z} and f as follows:

$$\text{row}_i(\mathbf{Z}) = \mathbf{e}_{k,f(i)}^T, \quad i = 1, 2, \dots, k \quad (4)$$

where $\text{row}_i(\mathbf{Z})$ denotes the i th row of \mathbf{Z} . Also, \mathbf{z} can be obtained from \mathbf{Z} simply as $\mathbf{z} = \text{diag}(\mathbf{Z}^T \mathbf{Z})$, where $\text{diag}(\mathbf{Z}^T \mathbf{Z})$

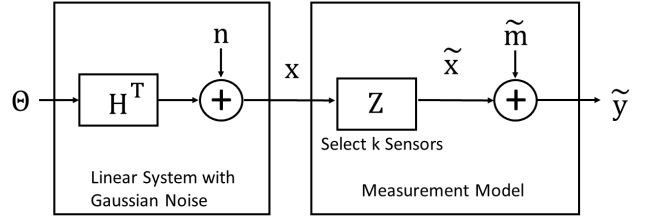


Fig. 1: System block diagram.

represents a column vector consisting of the diagonal elements of $\mathbf{Z}^T \mathbf{Z}$. As an example, for $N_s = 4$, when the second and third system outputs are selected, we have $k = 2$, $\mathbf{z} = [0, 1, 1, 0]^T$, $f(1) = 2$, $f(2) = 3$ and we construct the selection matrix as $\mathbf{Z} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$.

Based on the selection matrix \mathbf{Z} , the sensor selection operation can be expressed as

$$\tilde{\mathbf{y}} \triangleq \mathbf{Z}\mathbf{y} = \mathbf{Z}\mathbf{x} + \mathbf{Z}\mathbf{m} \triangleq \tilde{\mathbf{x}} + \tilde{\mathbf{m}}. \quad (5)$$

Namely, k out of N_s system outputs are measured via k sensors. The resulting system and measurement model is illustrated in Fig. 1.

For the cost of making a sensor measurement, we employ the measurement cost model proposed in [37]. Specifically, the cost of making a measurement via sensor i is given by [37]

$$c_i = 0.5 \log_2 \left(1 + \frac{\sigma_{n_i}^2}{\sigma_{m_i}^2} \right) \quad (6)$$

Similar to [2]–[4], we consider the expression in (6) as the cost of making a measurement with sensor i in our problem formulation. The important properties of this cost model are that it is nonnegative, monotonically decreasing, and convex with respect to $\sigma_{m_i}^2$. Considering sensor i associated with system output x_i , a higher cost means a more accurate measurement, y_i (see (2)). The sensor measurement model specified by (2) and (6) is first introduced in [37], where the physical problem of measuring the propagating wave field at a certain number of points and estimating the values of the field at other locations is considered. The aim was to recover the wave field as economically as possible based on a trade-off between the estimation accuracy and the cost of performing measurements. The plausibility of the proposed measurement cost function is discussed in meticulous detail in [37, Section III]. It is assumed that the ranges of measurement devices can be chosen freely to match any interval (similar to scaling the range of a multimeter) and the cost of the measurements is solely determined by the number of quantization levels it can reliably distinguish. The connections of the measurement problem with communication and rate-distortion theories via the Shannon's formula for the capacity of a Gaussian noise channel were also discussed.

Suppose that the parameter vector Θ takes one of two possible values. Namely, there exist two hypotheses defined as $\mathcal{H}_0: \Theta = \Theta_0$ and $\mathcal{H}_1: \Theta = \Theta_1$, where the prior probability of \mathcal{H}_i is denoted by π_i . The conditional probability distribution of the selected measurements $\tilde{\mathbf{y}}$ in (5) can be specified, based on the system model in Section II, as

$$\tilde{\mathbf{y}} | \mathcal{H}_i \sim \mathcal{N}(\mathbf{Z}\mathbf{H}^T \Theta_i, \mathbf{Z}(\Sigma + \Sigma_m)\mathbf{Z}^T) \quad (7)$$

for $i \in \{0, 1\}$. To determine the true hypothesis, we employ the Bayes rule, denoted by $\delta_B(\tilde{\mathbf{y}})$, which minimizes the

Bayes risk among all possible decision rules [1]. Assuming uniform cost assignment (UCA), the Bayes rule reduces to the maximum *a posteriori* probability (MAP) decision rule, which achieves the following Bayes risk (equivalently, the average probability of error) [2]:

$$r(\delta_B) = \pi_0 Q\left(\frac{\ln(\pi_0/\pi_1)}{d} + \frac{d}{2}\right) + \pi_1 Q\left(\frac{d}{2} - \frac{\ln(\pi_0/\pi_1)}{d}\right) \quad (8)$$

where π_0 and π_1 denote the prior probabilities of \mathcal{H}_0 and \mathcal{H}_1 , respectively, and

$$d \triangleq \left((\mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_1 - \mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_0)^T (\mathbf{Z}(\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_m)\mathbf{Z}^T)^{-1} \times (\mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_1 - \mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_0) \right)^{1/2} \quad (9)$$

The expression in (9) can also be written as

$$d = \left((\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0)^T \mathbf{H}\mathbf{Z}^T (\mathbf{Z}(\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_m)\mathbf{Z}^T)^{-1} \times \mathbf{Z}\mathbf{H}^T (\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0) \right)^{1/2} \quad (10)$$

Based on the definition of the selection matrix \mathbf{Z} , d in (10) can be stated, after some manipulation, as

$$d = \sqrt{\sum_{i=1}^{N_s} z_i \frac{(\mathbf{h}_i^T (\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0))^2}{\sigma_{n_i}^2 + \sigma_{m_i}^2}} \quad (11)$$

The aim is to minimize the Bayes risk $r(\delta_B)$ under a total cost constraint, specified by C_T , by making measurements with at most K sensors. Since it is known that $r(\delta_B)$ in (8) is a monotonically decreasing function of d [2], maximizing d is equivalent to minimizing the Bayes risk. Therefore, we propose the following sensor selection problem for binary hypothesis-testing:

$$\begin{aligned} & \underset{\mathbf{z}}{\text{maximize}} && \sum_{i=1}^{N_s} z_i p_i \\ & \text{subject to} && \sum_{i=1}^{N_s} z_i c_i \leq C_T \\ & && \sum_{i=1}^{N_s} z_i \leq K \\ & && z_i \in \{0, 1\}, \quad i = 1, 2, \dots, N_s \end{aligned} \quad (12)$$

where c_i is given by (6) and p_i is defined as

$$p_i \triangleq \frac{(\mathbf{h}_i^T (\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0))^2}{\sigma_{n_i}^2 + \sigma_{m_i}^2}. \quad (13)$$

Due to its combinatorial nature, the problem in (12) can be very complex to solve unless $\binom{N_s}{K}$ is small. To simplify the problem, the last constraint can be relaxed as $0 \leq z_i \leq 1$, $i = 1, \dots, N_s$, and a suitable optimization algorithm can be employed to obtain a solution for \mathbf{z} . Then, the elements of that solution can be used to determine the selected sensors.

Relaxing the last constraint in (12), we obtain the following convex optimization problem:

$$\begin{aligned} & \underset{\mathbf{z}}{\text{maximize}} && \sum_{i=1}^{N_s} z_i p_i \\ & \text{subject to} && \sum_{i=1}^{N_s} z_i c_i \leq C_T \\ & && \sum_{i=1}^{N_s} z_i \leq K \\ & && 0 \leq z_i \leq 1, \quad i = 1, 2, \dots, N_s \end{aligned} \quad (14)$$

The problem in (14) is a linearly constrained linear optimization problem. Hence, it can be solved efficiently via linear/convex optimization algorithms [39] such as the simplex method [40] and the interior point method [34]. Since the feasible region of (12) is contained in that of (14), the solution of (14) leads to an equal or higher objective value and provides a performance upper bound on the original problem in (12). Hence, (14) can be used to evaluate performance of suboptimal solution methods. In addition, the solution to (14) can be used as an initial point for developing close-to-optimal solutions of (12) with low computational complexity, as discussed towards the end of this section.

It is possible to specify the form of an optimal solution to (14) based on theoretical analysis. Towards that aim, we first provide the following two lemmas:

Lemma 1. *Let N_L denote the number of distinct sets that consist of indices of sensors having K largest p_i 's, and let B_1, B_2, \dots, B_{N_L} represent these sets. Assume that there exists $j \in \{1, 2, \dots, N_L\}$, such that $C_T \geq \sum_{i \in B_j} c_i$, where c_i is as defined in (6). Then, \mathbf{z}^* is a solution to (14) (and also to (12)), where the elements of \mathbf{z}^* are given by*

$$z_i^* = \begin{cases} 0, & i \notin B_j \\ 1, & i \in B_j \end{cases}. \quad (15)$$

Proof: Please see Appendix A.

To clarify the definition of the sets in Lemma 1, consider an example in which $N_s = 5$, $K = 3$, and $[p_1, p_2, p_3, p_5, p_5] = [20, 18, 22, 5, 18]$. Then, the sets in the lemma are obtained as $B_1 = \{1, 2, 3\}$ and $B_2 = \{1, 3, 5\}$ with $N_L = 2$. Basically, Lemma 1 states that if the cost budget allows the use of any best K sensors, it is optimal to select them.

Lemma 2. *Suppose that the optimization problem in (14) is feasible and let B_1, B_2, \dots, B_{N_L} denote the sets of indices of K largest p_i 's. If $C_T < \sum_{i \in B_j} c_i$ for all $j \in \{1, 2, \dots, N_L\}$, then there exists a solution \mathbf{z}^* to (14) that satisfies*

$$\sum_{i=1}^{N_s} z_i^* c_i = C_T \quad (16)$$

Proof: Please see [41, Sec. A.2].

Based on Lemma 1 and Lemma 2, the following proposition is obtained related to the solution of the relaxed problem in (14).

Proposition 1. *Suppose that the optimization problem in (14) is feasible. Then, there exists a solution \mathbf{z}^* to (14) that is characterized as either of the following:*

a) $\sum_{i=1}^{N_s} z_i^* = K$ with

$$z_i^* \in \begin{cases} \{0\}, & i \in S_0 \\ \{1\}, & i \in S_1, \quad i = 1, 2, \dots, N_s \\ [0, 1], & i \in S_2 \end{cases} \quad (17)$$

where S_0 , S_1 , and S_2 are disjoint sets of indices such that

$$\begin{aligned} S_0 \cup S_1 \cup S_2 &= \{1, 2, \dots, N_s\}, \\ |S_0| &= N_s - K - 1, \quad |S_1| = K - 1, \quad |S_2| = 2. \end{aligned} \quad (18)$$

b) $\sum_{i=1}^{N_s} z_i^* < K$ with

$$z_i^* \in \begin{cases} \{0\}, & i \in S_0 \\ \{1\}, & i \in S_1, \quad i = 1, 2, \dots, N_s \\ [0, 1], & i \in S_2 \end{cases} \quad (19)$$

where S_0 , S_1 , and S_2 are disjoint sets of indices such that

$$S_0 \cup S_1 \cup S_2 = \{1, 2, \dots, N_s\}, \quad |S_2| = 1 \quad (20)$$

Proof: Please see Appendix B.

Proposition 1 states that when the problem in (14) is feasible, a solution can be expressed to include at most two non-integer elements. To utilize Proposition 1 for obtaining a solution of (14), we first consider the following problem in which the number of selected sensors is forced to be equal to K .

$$\begin{aligned} & \underset{\mathbf{z}}{\text{maximize}} && \sum_{i=1}^{N_s} z_i p_i \\ & \text{subject to} && \sum_{i=1}^{N_s} z_i c_i \leq C_T \\ & && \sum_{i=1}^{N_s} z_i = K \\ & && 0 \leq z_i \leq 1, \quad i = 1, 2, \dots, N_s \end{aligned} \quad (21)$$

For this problem, Proposition 1 implies that a solution \mathbf{z}^* conforming to (17) and (18) can be found. In particular, $K-1$ or K elements of such a solution are one, and $N_s - K - 1$ or $N_s - K$ elements are zero. This characterization is helpful for obtaining the solution of (21) in a low-complexity manner. An algorithm is proposed for this purpose, which is presented as Algorithm 1 (please see Section III-A for the complexity analysis of Algorithm 1).

The main idea behind Algorithm 1 can be explained as follows: The algorithm initially checks whether any set of sensors with K largest p_i 's satisfies the cost constraint (cf. Lemma 1). If no such set exists, the algorithm searches for the two possibly non-integer components of the solution by enumerating all $\binom{N_s}{2}$ combinations of sensor indices. For each combination, all sensor indices are partitioned into three disjoint sets (a set for which $z_i = 1$, another set for which $z_i = 0$, and finally a set for which $z_i \in [0, 1]$). Finally, it is checked whether the Karush-Kuhn-Tucker (KKT) conditions can be satisfied for this partition.

Although Algorithm 1 can be used to solve (21), it is not directly applicable to the relaxed problem in (14). However, we argue that, with a suitable change of parameters, an exact solution to (14) can be obtained by applying Algorithm 1 on

Algorithm 1 Proposed Numerical Algorithm for the Solution of (21)

```

1: obtain  $B_1, \dots, B_{N_L}$  as sets of indices of  $K$  largest  $p_i$ 's.
2: if  $\exists k \in \{1, \dots, N_L\}$  s.t.  $\sum_{i \in B_k} c_i \leq C_T$  then
3:    $z_i^* = 1, i \in B_k$ 
4:    $z_i^* = 0, i \notin B_k$ 
5: else
6:   for all  $\binom{N_s}{2}$  combinations of sensor indices  $a, b$  do
7:     if  $c_a \neq c_b$  then
8:       init  $S_0, S_1, S_2$  as empty sets
9:       calculate  $\mu = (p_a - p_b)/(c_a - c_b)$ 
10:      calculate  $\nu = (p_b c_a - p_a c_b)/(c_a - c_b)$ 
11:      add every sensor index  $s$  that satisfies
12:         $p_s = \mu c_s + \nu$  to  $S_2$ 
13:      add every sensor index  $s$  that satisfies
14:         $p_s > \mu c_s + \nu$  to  $S_1$ 
15:      add remaining sensor indices to  $S_0$ 
16:       $M = |S_2|, N = |S_1|$ 
17:      if  $N < K < N + M$  then
18:        let  $C_{S_2}$  consist of indices of cheapest
19:           $(K - N)$  sensors in  $S_2$ 
20:        let  $E_{S_2}$  consist of indices of most expensive
21:           $(K - N)$  sensors in  $S_2$ 
22:        if  $\sum_{i \in (S_1 \cup C_{S_2})} c_i \leq C_T$  and
23:           $C_T \leq \sum_{i \in (S_1 \cup E_{S_2})} c_i$  then
24:             $X_0 = C_{S_2}, t = 0$ 
25:            while  $\sum_{i \in (S_1 \cup X_t)} c_i \leq C_T$  and
26:               $t < \min\{K - N, M + N - K\}$  do
27:              let  $m_t$  be index of cheapest sensor in  $X_t$ 
28:              let  $n_t$  be index of most expensive sensor
29:              in  $S_2 \setminus X_t$ 
30:               $X_{t+1} = (X_t \setminus \{m_t\}) \cup \{n_t\}$ 
31:               $t = t + 1$ 
32:            end while
33:             $T = t - 1$ 
34:             $S_{21} = X_T \setminus \{m_T\}$ 
35:             $S_{20} = S_2 \setminus (X_T \cup \{n_T\})$ 
36:            if  $c_{m_T} = c_{n_T}$  then
37:               $\alpha = 0$ 
38:            else
39:               $\alpha = \frac{(C_T - \sum_{i \in (S_1 \cup S_{21})} c_i - c_{m_T})}{(c_{n_T} - c_{m_T})}$ 
40:            end if
41:             $z_i^* = 1, i \in S_1 \cup S_{21}$ 
42:             $z_i^* = 0, i \in S_0 \cup S_{20}$ 
43:             $z_i^* = \alpha, i = n_T$ 
44:             $z_i^* = 1 - \alpha, i = m_T$ 
45:            break
46:          end if
47:        end if
48:      end if
49:    end for
50:  end if

```

an equivalent problem in the form of (21). To this aim, we define the following optimization problem:

$$\begin{aligned}
& \underset{\bar{\mathbf{z}}}{\text{maximize}} && \sum_{i=1}^{\bar{N}_s} \bar{z}_i \bar{p}_i \\
& \text{subject to} && \sum_{i=1}^{\bar{N}_s} \bar{z}_i \bar{c}_i \leq C_T \\
& && \sum_{i=1}^{\bar{N}_s} \bar{z}_i = K \\
& && 0 \leq \bar{z}_i \leq 1, \quad i = 1, 2, \dots, \bar{N}_s
\end{aligned} \tag{22}$$

where

$$\begin{aligned}
\bar{N}_s &\triangleq N_s + K \\
\bar{c}_i &\triangleq \begin{cases} c_i, & i = 1, 2, \dots, N_s \\ 0, & i = N_s + 1, N_s + 2, \dots, N_s + K \end{cases} \\
\bar{p}_i &\triangleq \begin{cases} p_i, & i = 1, 2, \dots, N_s \\ 0, & i = N_s + 1, N_s + 2, \dots, N_s + K \end{cases}
\end{aligned} \tag{23}$$

with c_i and p_i being defined in (6) and (13), respectively. A solution \mathbf{z}^* of the problem in (14) can be obtained from a solution $\bar{\mathbf{z}}^*$ of (22) as follows:

$$z_i^* = \bar{z}_i^*, \quad i = 1, 2, \dots, N_s \tag{24}$$

It is important to note that the optimization problem in (22) can be solved via Algorithm 1 as it is in the same form as that in (21).

The main idea behind the problem formulation in (22) is to introduce K hypothetical outputs, which induce no cost and no performance gain, in addition to the N_s actual system outputs. In this way, solving the new problem with $N_s + K$ outputs by choosing exactly K sensors becomes equivalent to solving the relaxed problem in (14) by choosing less than or equal to K sensors. This conclusion mainly comes from the introduction of slack variables to the problem in (14).

Based on our results related to the relaxed optimization problem in (14), we propose a suboptimal solution procedure for the original optimization problem in (12) as follows:

Proposed Suboptimal Solution to (12):

- 1) Obtain the equivalent relaxed problem in the form of (22). Calculate its solution as $\bar{\mathbf{z}}^*$ via Algorithm 1. (Due to Proposition 1-a), $\bar{\mathbf{z}}^*$ either has all integer entries or contains exactly two non-integer entries.)
- 2) If $\bar{\mathbf{z}}^*$ has two non-integer entries, we generate a new vector $\hat{\mathbf{z}}$ by setting the non-integer entry of $\bar{\mathbf{z}}^*$ with the lower associated cost to one and the other to zero. If $\bar{\mathbf{z}}^*$ has all integer entries, $\hat{\mathbf{z}}$ is equal to $\bar{\mathbf{z}}^*$. Notice that

$$\sum_{i=1}^{N_s} \hat{z}_i c_i \leq \sum_{i=1}^{N_s} \bar{z}_i^* c_i \leq C_T. \tag{25}$$

- 3) Run the local optimization algorithm on $\hat{\mathbf{z}}$ (Algorithm 2), and denote the resulting selection vector as $\hat{\mathbf{z}}'$.
- 4) Obtain the proposed suboptimal solution $\bar{\mathbf{z}}$ to (12) from $\hat{\mathbf{z}}'$ by using the relation

$$\bar{z}_i = \hat{z}'_i, \quad i = 1, 2, \dots, N_s \tag{26}$$

It should be noted that the main aim of the second step is to modify $\bar{\mathbf{z}}^*$ in such a way that its components satisfy the last constraint in (12) (i.e., the solution of (12) must be a binary

vector). After the second step, the entries corresponding to the hypothetical outputs (indices from $N_s + 1$ to $N_s + K$) could be dropped from $\hat{\mathbf{z}}$ (as in (26)) to obtain a selection vector that is in the feasible region of (12). Instead, we attempt to increase the objective value further by considering swaps between selected and unselected sensors starting from the selection vector $\hat{\mathbf{z}}$. This approach is similar to that in [24], where swaps are performed to improve the suboptimal solution obtained from the formulation of a relaxed selection problem. (The main difference is that we only consider swaps that do not violate the total cost constraint.) The swapping algorithm can be named as ‘local optimization’ since it starts with the selection vector denoted by $\hat{\mathbf{z}}$ (i.e., a certain point in the feasible region of (12)) and performs a search by iterating through adjacent points. The local optimization algorithm starts with $\hat{\mathbf{z}}$, which is obtained in Step 2, as described above. The algorithm seeks to improve the objective value via swaps that do not violate the total cost constraint. It terminates when no such swaps can improve the objective value. (The pseudo-code of the local optimization algorithm is provided in Algorithm 2.) Finally, in the last step, the proposed suboptimal solution $\bar{\mathbf{z}}$ is constructed from the first N_s entries of $\hat{\mathbf{z}}'$, which is obtained in Step 3.

Algorithm 2 Local Optimization

```

1: get  $S_1$  from  $\hat{\mathbf{z}}$  ▷ set of selected sensors
2: get  $S_0$  from  $\hat{\mathbf{z}}$  ▷ set of unselected sensors
3: thisCost = cost of  $S_1$ 
4: top:
5: for  $i = 1$  to  $K$  do
6:   for  $j = 1$  to  $N_s - K$  do
7:     Δcost = cost of  $j$ th element in  $S_0$  –
       cost of  $i$ th element in  $S_1$ 
8:     if thisCost + Δcost ≤  $C_T$  then
9:       ΔobjValue = obj. value of  $j$ th element in  $S_0$  –
        obj. value of  $i$ th element in  $S_1$ 
10:      if ΔobjValue > 0 then
11:        exchange  $i$ th element of  $S_1$  with
         $j$ th element of  $S_0$ 
12:        thisCost = thisCost + Δcost
13:      goto top.
14:    end if
15:  end if
16: end for
17: end for
18: construct  $\hat{\mathbf{z}}$  from  $S_1, S_0$ 

```

Regarding the computational complexity of Algorithm 2, it is first noted that the termination of Algorithm 2 is guaranteed since there exists a finite number of swaps that can improve the objective value starting from a selection vector. The number of iterations in Algorithm 2 can be large in theory since the outer *for* loop in line 5 is re-initialized after each swap. As proposed in [24], an upper limit can be imposed on the number of iterations of the inner *for* loop in line 6. We can choose the limit such that in our proposed suboptimal solution technique, the local optimization stage does not dominate Algorithm 1 in terms of the order of growth associated with the running time. From Algorithm 2, it is observed that the operations performed at each individual iteration require a constant time. Therefore, the number of iterations directly determines the complexity of the algorithm. Hence, an iteration limit that grows no faster than N_s^3 can be chosen as the complexity order of Algorithm 1 is $\mathcal{O}(N_s^3)$ as discussed next.

A. Complexity Analysis of Algorithm 1

In this subsection, we analyze the computational complexity of Algorithm 1 for solving (22) under realistic settings. We obtain an asymptotical upper bound on the order of growth of runtime. Numerical results about the runtime of Algorithm 1 and its comparisons against the simplex and interior point methods are presented in Section V-A.

Algorithm 1 starts with the computation of the total cost arising from the selection of best K sensors. Since there can exist sensors i and j with $p_i = p_j$, there can be multiple sets of best K sensors. If such a set has a total cost not violating the total cost constraint, C_T , then selecting this set of sensors is a solution and the algorithm terminates (Lemma 1). When there is no set of best K sensors with a total cost not exceeding the total cost constraint, the algorithm proceeds with the *else* condition in line 5. Algorithm 1 consists of two main functions. The first one is the enumeration of best K sensors and the computation of the cost of selecting them. The second one is to find the sensor pair a and b for which the KKT conditions hold.

The function for finding the sensor pair a and b for which the KKT conditions hold is related to the *for* loop in line 6. In the worst case, this loop iterates $\binom{N_s}{2}$ times. This occurs when the *break* statement in line 36 is executed in the $\binom{N_s}{2}$ th iteration of the *for* loop. On the other hand, the *while* loop in line 21 runs for at most $\min\{K - N, M + N - K\}$ iterations. For the discussion here, it suffices to consider that $N_s \geq \min\{K - N, M + N - K\}$. Apparently, this *while* loop is executed only in the last iteration of the outer *for* loop (see the *break* statement in line 36). Focusing on the execution time of the other lines in the *for* loop, we have comparison operations that divide the sensor indices to disjoint sets. The complexities of these operations are tied linearly to the input size, N_s . Since there are no other operations with higher order complexity, we can multiply the linear computational effort of the individual iterations with the quadratically growing iteration count to specify the overall complexity. As a result, the asymptotical upper bound on the function relating the runtime to the input size, N_s , is $\mathcal{O}(N_s^3)$.

In the worst case, the enumeration process of the sets with K largest p_i 's (called as best sets) has complexity $\mathcal{O}(\binom{N_s}{K})$. In other words, the number of such sets, N_L , can be as large as $\binom{N_s}{K}$, which is the case when all the elements in vector \mathbf{p} are equal. Such a case would render the enumeration process of best sets impractical. However, in realistic scenarios, the elements of vector \mathbf{p} are distinct since they depend on system parameters and noise levels (see (13)). For example, \mathbf{h}_i can represent the channel for the i th system output. Therefore, in practice, the complexity of the enumeration process is mainly related to sorting the elements of vector \mathbf{p} , which has a complexity of $\mathcal{O}(N_s^2)$ or $\mathcal{O}(N_s \log N_s)$ depending on the employed sorting algorithm.¹

Overall, for practical scenarios, the complexity of Algorithm 1 can be specified as $\mathcal{O}(N_s^3)$.

IV. JOINT SENSOR SELECTION AND DESIGN FOR BINARY HYPOTHESIS TESTING

In Section III, the sensor selection problem is investigated under a cost constraint to minimize the Bayes risk for a given binary hypothesis testing problem by considering fixed

measurement noise variances $(\sigma_{m_1}^2, \sigma_{m_2}^2, \dots, \sigma_{m_{N_s}}^2)$ for the sensors, which corresponds to using sensors with fixed/given costs. In this section, we focus on the joint selection and design of sensors by optimally determining both the number of sensors and their measurement noise variances (i.e., costs). To that aim, let $\boldsymbol{\sigma}_m^2$ denote the vector of measurement noise variances, defined as

$$\boldsymbol{\sigma}_m^2 \triangleq [\sigma_{m_1}^2, \sigma_{m_2}^2, \dots, \sigma_{m_{N_s}}^2]^T. \quad (27)$$

Since the aim is to optimize the selection vector \mathbf{z} and $\boldsymbol{\sigma}_m^2$ jointly, we extend the sensor selection problem in (12) (also see (6) and (13)) as follows:

$$\begin{aligned} & \underset{\mathbf{z}, \boldsymbol{\sigma}_m^2}{\text{maximize}} && \sum_{i=1}^{N_s} z_i \frac{(\mathbf{h}_i^T (\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0))^2}{\sigma_{n_i}^2 + \sigma_{m_i}^2} \\ & \text{subject to} && 0.5 \sum_{i=1}^{N_s} z_i \log_2 \left(1 + \frac{\sigma_{n_i}^2}{\sigma_{m_i}^2} \right) \leq C_T \\ & && \sum_{i=1}^{N_s} z_i \leq K \\ & && z_i \in \{0, 1\}, \quad i = 1, 2, \dots, N_s \end{aligned} \quad (28)$$

In other words, the Bayes risk is to be minimized over both \mathbf{z} and $\boldsymbol{\sigma}_m^2$ under the cost constraint.

Before investigating the solution of the optimization problem in (28), we first consider the problem for a fixed \mathbf{z} and present the following optimization problem over $\boldsymbol{\sigma}_m^2$ (called the measurement noise variance design problem):

$$\begin{aligned} & \underset{\boldsymbol{\sigma}_m^2}{\text{maximize}} && \sum_{i \in Z_1} \frac{(\mathbf{h}_i^T (\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0))^2}{\sigma_{n_i}^2 + \sigma_{m_i}^2} \\ & \text{subject to} && 0.5 \sum_{i \in Z_1} \log_2 \left(1 + \frac{\sigma_{n_i}^2}{\sigma_{m_i}^2} \right) \leq C_T \\ & && \sigma_{m_i}^2 = \infty, \quad i \in Z_0 \end{aligned} \quad (29)$$

where sets Z_0 and Z_1 are defined as

$$Z_0 = \{i \in \{1, 2, \dots, N_s\} \mid z_i = 0\}, \quad (30)$$

$$Z_1 = \{i \in \{1, 2, \dots, N_s\} \mid z_i = 1\}. \quad (31)$$

It is noted that the optimization variables $\{\sigma_{m_i}^2\}_{i \in Z_0}$ do not affect the values of the objective function and the cost constraint since they correspond to unselected sensor measurements. However, they are included for keeping the generality of the formulation.

The problem in (29) is analyzed in [2]. It is shown that since a convex function is maximized over a convex set, the solution of (29) lies at the boundary. Namely, the solution can be obtained by an iterative algorithm that can be outlined as in Algorithm 3, where the following definitions are used for the simplicity of the expressions:

$$\mu_i^2 \triangleq (\mathbf{h}_i^T (\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0))^2, \quad (32)$$

$$\boldsymbol{\mu}^2 \triangleq [\mu_1^2, \mu_2^2, \dots, \mu_{N_s}^2]^T, \quad (33)$$

$$\boldsymbol{\sigma}_n^2 \triangleq [\sigma_{n_1}^2, \sigma_{n_2}^2, \dots, \sigma_{n_{N_s}}^2]^T. \quad (34)$$

The complexity of Algorithm 3 is governed by the maximum number of iterations that the *while* loop in line 3 can have, and the number of operations required for the

¹The hypothetical outputs, which are introduced to construct (22), do not affect the number of sets with K largest p_i 's since their p_i 's are zero by definition.

Algorithm 3 Optimal Variance Design [2]

```

1: get  $\mathbf{z}, \mu^2, \sigma_n^2$ 
2:  $Z_{fin} = \{i \in \{1, \dots, N_s\} \mid z_i = 1\}$ 
3: while  $Z_{fin} \neq \emptyset$  do
4:    $\alpha = \left(2^{2C_T} \prod_{i \in Z_{fin}} \frac{\sigma_{n_i}^2}{\mu_i^2}\right)^{\frac{1}{|Z_{fin}|}}$ 
5:    $S_{inf} = \{i \in \{1, \dots, N_s\} \mid (i \in Z_1) \& (\sigma_{n_i}^2 \geq \alpha \mu_i^2)\}$ 
6:   if  $S_{inf} \neq \emptyset$  then
7:      $i = \arg \min_{j \in S_{inf}} \sigma_{n_j}^2$ 
8:      $Z_{fin} = Z_{fin} \setminus \{i\}$ 
9:   else
10:    break
11:  end if
12: end while
13:  $\sigma_{m_i}^2 = \begin{cases} \frac{\sigma_{n_i}^4}{\mu_i^2 \alpha - \sigma_{n_i}^2}, & i \in Z_{fin} \\ \infty, & \text{else} \end{cases} \quad i = 1, 2, \dots, N_s$ 

```

computation of α in each iteration (line 4). They are both in linear relationship with the number of elements in Z_{fin} , which is always less than or equal to N_s . Hence, a quadratic function asymptotically bounds the order of growth of its runtime ($\mathcal{O}(N_s^2)$).

Algorithm 3 will be useful for obtaining the solution of the joint optimization problem in (28), as discussed in the following.

Based on (6), the measurement noise variance of the i th sensor can be stated in terms of its cost as $\sigma_{m_i}^2 = \sigma_{n_i}^2 / (2^{2c_i} - 1)$. Then, the joint optimization problem in (28) can equivalently be expressed as

$$\begin{aligned}
& \underset{\mathbf{z}, \mathbf{c}}{\text{maximize}} && \sum_{i=1}^{N_s} z_i \frac{\mu_i^2 (2^{2c_i} - 1)}{\sigma_{n_i}^2 2^{2c_i}} \\
& \text{subject to} && \mathbf{z}^T \mathbf{c} \leq C_T \\
& && \sum_{i=1}^{N_s} z_i \leq K \\
& && z_i \in \{0, 1\}, \quad i = 1, 2, \dots, N_s \\
& && c_i \geq 0, \quad i = 1, 2, \dots, N_s
\end{aligned} \tag{35}$$

where

$$\mathbf{c} \triangleq [c_1, c_2, \dots, c_{N_s}]. \tag{36}$$

Remark. Setting either $c_i = 0$ or $z_i = 0$ effectively results in not selecting the sensor with index i .

The solution of (35) is specified by the following proposition.

Proposition 2. Let \tilde{B} denote the set of indices corresponding to K largest values of $\mu_i^2 / \sigma_{n_i}^2$ for $i = 1, 2, \dots, N_s$ (break ties arbitrarily). Then, a solution to the joint optimization problem in (35) is $(\mathbf{z}^*, \mathbf{c}^*)$, where the elements of \mathbf{z}^* are given by

$$z_i^* = \begin{cases} 1, & i \in \tilde{B} \\ 0, & \text{else} \end{cases}, \quad i = 1, 2, \dots, N_s \tag{37}$$

and \mathbf{c}^* is an optimizer of the problem in (35) when \mathbf{z} is fixed as $\mathbf{z} = \mathbf{z}^*$. (Namely, \mathbf{c}^* can be obtained via Algorithm 3 and (6) by setting $\mathbf{z} = \mathbf{z}^*$ in (35).)

Proof: Please see Appendix C.

Proposition 2 states that it is optimal to allocate all the cost budget to K sensors with largest values of $\mu_i^2 / \sigma_{n_i}^2$ ratios among indices $i = 1, 2, \dots, N_s$. Intuitively, these ratios can be regarded as the SNR values of the sensors; hence, the sensors with highest SNRs are selected. It is also interesting to note that the joint problem considered in this section leads to a simpler sensor selection solution than the sensor selection problem considered in Section III for sensors with fixed measurement noise variances. In addition, it is noted that the solution of (28) includes cases in which measurement noise variances of some sensors are set to infinity, which corresponds to assigning no cost to those sensors. In fact, this is equivalent to not selecting (using) those sensors at all.

V. NUMERICAL EXAMPLES

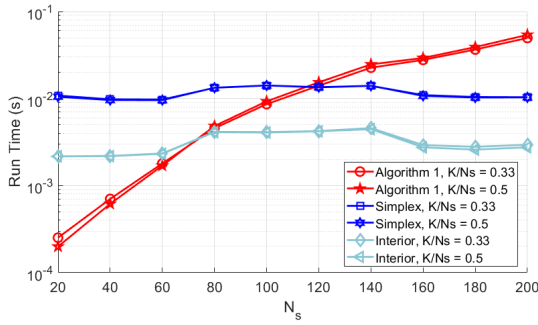
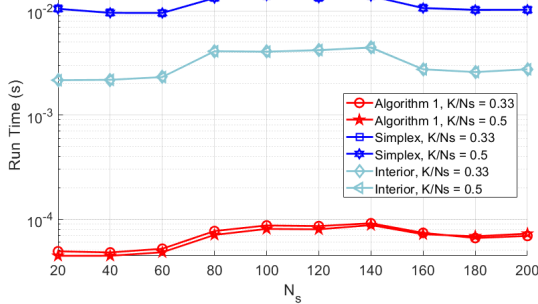
In this section, we first compare Algorithm 1 with the simplex method and the interior point method in terms of runtime for solving the relaxed problem in (14). Then, we provide examples for both the sensor selection problem in Section III and the joint sensor selection and design problem in Section IV. In the simulations, a linear system as in Fig. 1 is considered, where the number of potential sensors (N_s), the number of sensors to select (K), the total cost constraint (C_T) and the length of Θ (which is L) are taken as the simulation parameters (to be specified in the related sections). Also, parameter vector Θ is equal to Θ_0 under hypothesis \mathcal{H}_0 and equal to Θ_1 under hypothesis \mathcal{H}_1 . The entries of Θ_0 and Θ_1 are i.i.d. with each component being uniformly distributed in the closed interval of $[0, 1]$. \mathbf{H} is a system matrix of size $L \times N_s$ and is considered to be known in advance for the considered problems. The entries of the system matrix \mathbf{H} are i.i.d. random variables that are uniformly distributed in the interval $[-0.1, 0.1]$. The entries of the system noise variance vector, σ_n^2 , and the measurement noise variance vector, σ_m^2 , also come from a uniform distribution in the interval $[0.05, 1]$.

A. Runtime Simulation of Algorithm 1

In this part, we consider the linear optimization problem in (14) and solve it via three different methods. The first one is based on Algorithm 1. Specifically, the problem in (14) is converted into the problem in (22), and the solution of (22) is computed via Algorithm 1. Then, by dropping the indices of the hypothetical outputs, we obtain the solution of (14) (see (24)). The other methods are the simplex and interior point methods, which are commonly employed for solving linear optimization problems.

In the literature, the simplex method is investigated in detail and its implementation is optimized so that it performs well in most applications [42], [43]. Despite its effectiveness in practice, it is proven that the simplex method has an exponential complexity with respect to input size (number of optimization variables) [42]. This is due to the fact that in the worst case, the simplex method visits all vertices of the feasible region.² On the other hand, the interior point methods address this issue and are constructed to perform in polynomial time. As the name implies, in such methods, optimization variables are moved to the interior of the feasible region at each iteration. Although the interior point method is asymptotically more efficient, special care must be taken when evaluating optimization methods for a specific problem, which is (14) in this case.

²By randomizing its inputs, the simplex method can have polynomial time complexity [43].

(a) $C_T = 1.4$ times the cost of the cheapest K sensors.(b) $C_T = 4$ times the cost of the cheapest K sensors.Fig. 2: Average running time of different methods versus N_s for solving (14).

To obtain statistically meaningful results for the running times of the three solution methods, we obtain 10000 realizations for the previously described random variables Θ_0 , Θ_1 , \mathbf{H} , σ_n^2 , and σ_m^2 . For each realization, we solve the corresponding optimization problem with the described methods and obtain the running time of each method. In Fig. 2, we plot the average running time of each method versus N_s for two different values of the numbers of sensors to select (K) and for two different cost constraints. The length of Θ (which is denoted as L) is taken as $\lceil N_s/2 \rceil$. Also, the total cost constraint (C_T) is defined as a multiple of the sum of the costs of the cheapest K sensor measurements. The simulations are performed on an Intel Core i7 2.6GHz PC with 16GB of physical memory using MATLAB R2018a on a Windows 10 operating system. The simplex and interior point methods are implemented via the “linprog” function in the Optimization Toolbox of MATLAB.

In Fig. 2a, the total cost constraint, C_T , is set to 1.4 times the cost of the cheapest K sensors. It is seen that for $N_s < 80$, it takes a shorter amount of average time for Algorithm 1 to solve the relaxed problem in (14) than the simplex and interior point methods. However, as N_s gets larger, the runtime performance of Algorithm 1 deteriorates. Also, it is noted that the number of sensors to select (K) does not have significant effects on runtime when it is changed from one third to one half of N_s . In addition, the running times of the simplex and interior point methods are not affected significantly from N_s or K . In Fig. 2b, the total cost constraint is set to 4 times the cost of the cheapest K sensors. In this case, the cost constraint is not strict and in most cases it becomes feasible to select K sensor measurements with highest contributions to the objective value. Since Algorithm 1 evaluates the feasibility of best sets of K sensors at the beginning, it achieves significantly lower running times than the other methods in this scenario. The execution times of the simplex and interior point methods

are not affected by this fact as they are not specifically designed for the problem in (14).

Overall, it is concluded that when the cost constraint is not strict or N_s is not very large, the proposed solution employing Algorithm 1 and Algorithm 2 can be used to perform sensor selection. On the other hand, if the cost constraint is strict and N_s is very large, then the interior point method can be used instead of Algorithm 1.

B. Sensor Selection for Binary Hypothesis Testing

In this part, we consider the sensor selection problem for the described binary hypothesis testing problem and focus on the formulation in (12). We investigate the performance of the proposed suboptimal solution to (12) in Section III. We consider the linear system in Fig. 1, where the value of N_s is set to 100 and L is taken to be 20. Again, 10000 realizations are obtained for the previously described random variables Θ_0 , Θ_1 , \mathbf{H} , σ_n^2 , and σ_m^2 . For each realization, we solve the optimization problem in (12) via the proposed method (which is described in Section III in items 1-4 preceding (26)) and obtain the resulting values of the objective function. We then average out the objective values for different realizations to provide the average performance results.

We also present two different sensor selection strategies, along with the proposed solution method, for comparison purposes, which are described as follows:

- **Simple Selection Strategy:** In this strategy, sensors are sorted in descending order according to their p_i values (please see the definition in (13)). Then, starting from the top of the sorted list, sensors are added to the set of selected sensors one by one until no remaining sensors can be selected (because either selecting any of the remaining sensors results in the violation of the cost constraint or the constraint on the number of sensors to select, K , is achieved).

- **Selection with only Local Optimization:** In this strategy, the cheapest K sensors are selected and the local optimization algorithm (Algorithm 2) is executed based on this initial selection.

In Figs. 3 and 4, the proposed solution (based on relaxation and local search), the simple selection strategy, and the selection with only local optimization strategy are labeled as ‘Proposed’, ‘Simple’, and ‘LocalOpt’, respectively. In addition, ‘Relaxed’ denotes the objective value achieved by the solution of the linear optimization problem in (14), which is the relaxed version (12). Hence, the curves labeled as ‘Relaxed’ provide performance bounds in the considered scenarios.

In Fig. 3, the performance of the considered strategies is presented versus the normalized total cost parameter (C_T divided by the cost of the cheapest K sensors) for two different values of K . For the performance metric, the objective value in (12) achieved by each strategy is employed, which corresponds to d^2 , with d being given by (11). From Fig. 3, it is observed that the performance of all the strategies improves as the total cost constraint C_T and/or the number of selected measurements, K , increase. Also, it is noted that the rate of performance improvement decreases as C_T increases; hence, there is a diminishing return in increasing the cost budget. In addition, it is noted that the proposed strategy has the best performance, which is very close to the performance bound (‘Relaxed’). Moreover, the simple selection strategy achieves higher (lower) objective values than the selection with only local optimization strategy when C_T is lower (higher) than a certain value for both $K = 20$ and $K = 40$. Although the gap

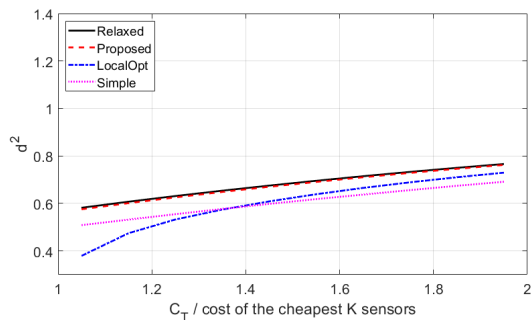
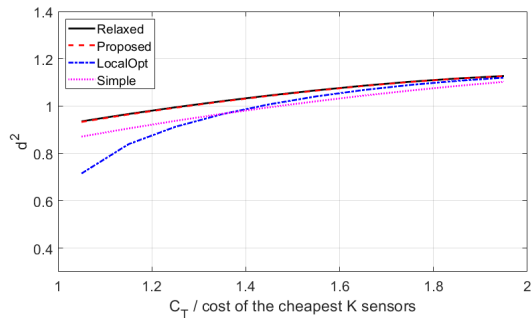
(a) $K = 20$ (b) $K = 40$

Fig. 3: Performance of different strategies versus normalized cost, together with the performance bound obtained from the relaxed problem in (14).

between the performance bound and the selection with only local optimization strategy is significant for all values of the total cost constraint in case of low K values, it becomes quite small for higher values of K in the region of high total cost constraints.

In Fig. 4, the performance of the strategies are plotted versus K for two different cost budgets. The proposed strategy achieves the best performance, which is close to the upper bound. The simple selection strategy performs better than the selection with only local optimization strategy when the total cost constraint, C_T , is equal to 1.05 times the cost of the cheapest K sensors. As the cost constraint becomes less strict (i.e., when C_T is equal to 1.45 times the cost of the cheapest K sensors), the local optimization starts to perform better than simple selection method. When C_T is equal to 1.45 times the cost of the cheapest K sensors, the gap between the performance bound and the selection with only local optimization strategy decreases as K increases. However, when C_T is equal to 1.05 times the cost of the cheapest K sensors (i.e., for low total cost constraints), the corresponding gap does not reduce with K . A similar observation is also valid for the gap between the simple selection strategy and the performance bound.

C. Sensor Selection and Design for Binary Hypothesis Testing

In this part, we provide numerical results for the joint sensor selection and design problem given in (28). The same simulation setup as in Section V-B is used. However, it should be noted that, for the joint sensor selection and design problem, the realization of σ_m^2 is irrelevant since it is considered as an optimization variable, and determined via the solution method.

To obtain the proposed optimal solution to (28), we utilize the approach described in Proposition 2. In addition to the

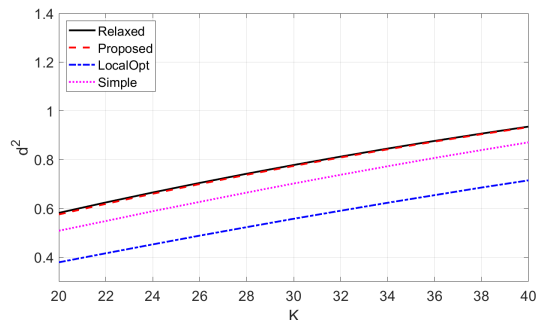
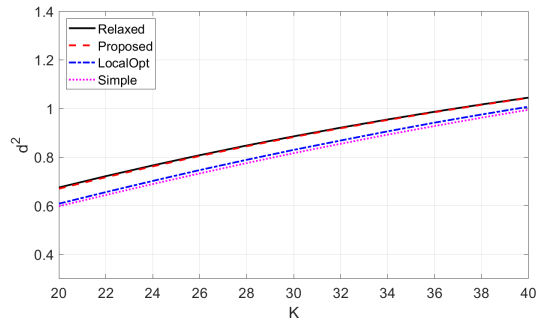
(a) $C_T = 1.05$ times the cost of the cheapest K sensors(b) $C_T = 1.45$ times the cost of the cheapest K sensors

Fig. 4: Performance of different strategies versus K together with the performance bound obtained from the relaxed problem in (14).

proposed optimal solution, we also present results for two sub-optimal sensor selection and design strategies for comparison purposes. These strategies are explained as follows:

- **Allocate Equal Cost to Best K Sensors:** In this strategy, the sensors are sorted in a descending order according to the values of $\mu_i^2/\sigma_{n_i}^2$. Then, the top K sensors are selected and a cost of C_T/K is allocated to each of them. Therefore, the measurement noise variance for a selected sensor (call sensor j) becomes

$$\sigma_{m_j}^2 = \frac{\sigma_{n_j}^2}{2C_T/K - 1}. \quad (38)$$

- **Allocate All Cost to Best Sensor:** As in the previous strategy, the sensors are sorted in a descending order of $\mu_i^2/\sigma_{n_i}^2$ ratios, and the top K sensors are selected. Then, all the cost is allocated to the sensor with the highest $\mu_i^2/\sigma_{n_i}^2$ ratio, and the other sensors are allocated zero cost (i.e., infinite measurement noise variance). If the best sensor has index j , then its measurement noise variance is given by

$$\sigma_{m_j}^2 = \frac{\sigma_{n_j}^2}{2C_T - 1}. \quad (39)$$

In Fig. 5, the performance of the proposed strategy (labeled as ‘Optimal’) is evaluated by plotting d^2 against the total cost constraint C_T for $K = 15$. In addition, the performance of the ‘allocate equal cost to best K sensors’ strategy (labeled as ‘EqualCost’) and the ‘allocate all cost to best sensor’ strategy (labeled as ‘AllCostBest’) is presented in the same figure. It is observed that allocating all the cost to the best sensor achieves the same performance as the proposed optimal strategy for very small values of C_T . However, as C_T increases its detection performance diverges significantly from the optimal performance. The main reason for this is that, for very low

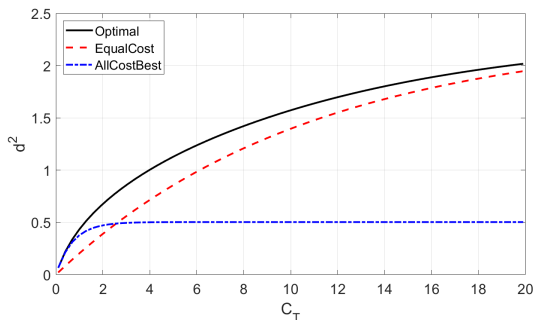


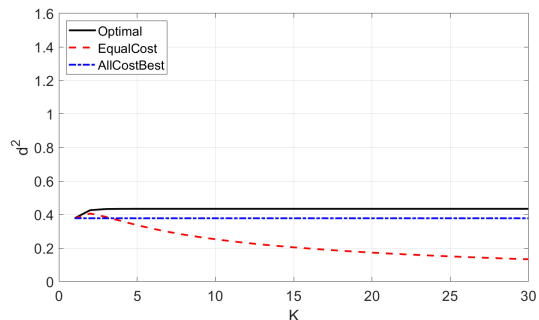
Fig. 5: Performance of different strategies versus C_T for $K = 15$.

cost budgets, the optimal strategy assigns non-zero cost only to the best sensor. As the cost budget increases, the optimal approach requires assigning non-zero costs to multiple sensors to benefit from the diversity of sensor measurements. It is also noted from Fig. 5 that the detection performance achieved by allocating all the cost to the best sensor quickly increases with C_T for small values of C_T . However, further increases in C_T do not result in any significant performance increase since the performance of this strategy is upper bounded by $\mu_j^2/\sigma_{n_j}^2$, where j is the index of the best sensor (please see the objective value of the joint sensor selection and design problem in (35) for verification). On the other hand, the strategy that allocates equal costs to the best K sensors yields a close performance to the proposed optimal strategy at high cost budgets; however, its performance becomes the worst at low values of C_T .

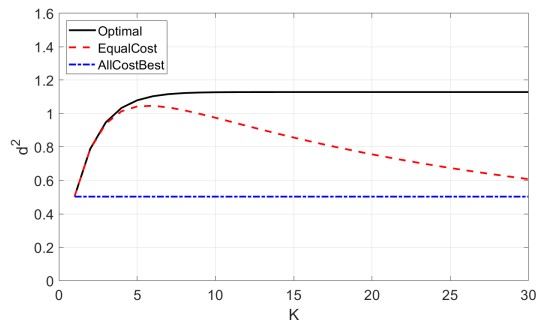
In Fig. 6, the detection performance of the considered strategies is plotted with respect to K for some fixed cost budgets; namely, $C_T = 1$ and $C_T = 5$. From the figure, it is first noted that the strategy of allocating all the cost to the best sensor achieves a constant performance with respect to K since it only employs one sensor. It is also observed that the performance of the optimal strategy improves with K up to a certain value. After that value, the optimal strategy does not allocate any positive cost to new sensors but rather keeps the previously selected sensors. In addition, the value of K after which the optimal strategy has constant detection performance increases as the total cost constraint C_T gets larger. On the other hand, the performance of the strategy that allocates equal costs to the best K sensors first increases and then decreases with respect to K . The increasing part occurs since allocating the cost budget C_T to a larger set of sensors is beneficial up to some point due to the diversity in the sensor measurements. However, after some value of K , distributing C_T among a large number of sensors equally becomes unfavorable since each sensor starts getting a low cost of C_T/K , which corresponds to low quality sensor measurements. Moreover, it is noted that the value of K after which the performance starts degrading gets larger as the cost budget increases. Overall, Fig. 5 and Fig. 6 illustrate the advantages of the proposed optimal strategy in various scenarios.

VI. CONCLUSION

We have formulated and investigated a sensor selection problem for binary hypothesis testing in order to minimize the Bayes risk via sensor selection in the presence of a constraint on the total cost of sensors. Due to the combinatorial nature of the problem, we have first performed linear relaxation of the selection vector and obtained a relaxed version of the original



(a) $C_T = 1$



(b) $C_T = 5$

Fig. 6: Performance of different strategies versus K .

problem. For calculating the solution of the relaxed problem, a low complexity algorithm has been developed based on some theoretical results. Then, a local search algorithm has been used to generate a solution to the original problem. Via numerical examples, we have showed that linear relaxation along with local optimization proves to be a practical method to provide close-to-optimal solutions for the proposed cost constrained sensor selection problem.

As an extension, we have regarded the measurement noise variances of sensors as additional optimization variables, and proposed a joint sensor selection and design problem. Based on theoretical results, a practical approach has been proposed to obtain an optimal solution to this joint problem. Numerical examples have been presented to evaluate the proposed approaches and to provide comparisons with other techniques.

APPENDIX

A. Proof of Lemma 1

Consider the optimization problem in (14) in the absence of the cost constraint. Then, it is easy to verify that \mathbf{z}^* defined in the lemma is a solution to (14) as it corresponds to K largest p_i 's. Since it is assumed that $C_T \geq \sum_{i \in B_j} c_i$, the cost constraint is already satisfied for \mathbf{z}^* . Hence, \mathbf{z}^* is a solution to (14) in the presence of the cost constraint, as well. As the elements of \mathbf{z}^* are either zero or one, it also becomes the solution of (12).

B. Proof of Proposition 1

Let B_1, B_2, \dots, B_{N_L} denote the sets of indices of K largest p_i 's (break ties arbitrarily). Consider the case that there exists j such that $C_T \geq \sum_{i \in B_j} c_i$. Then, by Lemma 1, a solution to (14) can be expressed as

$$z_i^* = \begin{cases} 0, & i \notin B_j \\ 1, & i \in B_j \end{cases} \quad (40)$$

which conforms to the characterization in (17) and (18).

Consider the case of $C_T < \sum_{i \in B_j} c_i$ for all $j \in \{1, 2, \dots, N_L\}$. In this case, there exists a solution \mathbf{z}^* to (14) that satisfies $C_T = \sum_{i=1}^{N_s} z_i^* c_i$ by Lemma 2. \mathbf{z}^* should satisfy the following KKT conditions with an equality constraint for the total cost:

$$\sum_{i=1}^{N_s} z_i^* - K \leq 0 \quad (41)$$

$$\sum_{i=1}^{N_s} z_i^* c_i - C_T = 0 \quad (42)$$

$$0 \leq z_i^* \leq 1, \quad i = 1, 2, \dots, N_s \quad (43)$$

$$\nu \left(\sum_{i=1}^{N_s} z_i^* - K \right) = 0 \quad (44)$$

$$\lambda_i \geq 0, \quad i = 1, 2, \dots, 2N_s \quad (45)$$

$$\lambda_i z_i^* = 0, \quad i = 1, 2, \dots, N_s \quad (46)$$

$$\lambda_{N_s+i} (z_i^* - 1) = 0, \quad i = 1, 2, \dots, N_s \quad (47)$$

$$-p_i - \lambda_i + \lambda_{N_s+i} + \mu c_i + \nu = 0, \quad i = 1, 2, \dots, N_s \quad (48)$$

where $\lambda_1, \lambda_2, \dots, \lambda_{2N_s}$, μ , and ν are the KKT multipliers. From (44)–(48), it is observed that if $z_i^* \in (0, 1)$ we get

$$\begin{aligned} \sum_{i=1}^{N_s} z_i^* = K &\implies p_i = \mu c_i + \nu \\ \sum_{i=1}^{N_s} z_i^* < K &\implies p_i = \mu c_i \end{aligned} \quad (49)$$

Suppose that there exists a solution \mathbf{z}' to (14) ($\sum_{i=1}^{N_s} z'_i c_i = C_T$), where $\sum_{i=1}^{N_s} z'_i = K$ and \mathbf{z}' does not satisfy the property in (17) and (18), meaning that it has $M > 2$ non-integer components; i.e., $z'_i \in (0, 1)$. In this case, we argue that there exists another solution to (14) that satisfies (17) and (18). Define sets of indices S'_0 , S'_1 and S'_2 as

$$\begin{aligned} S'_0 &\triangleq \{i : z'_i = 0, \quad i = 1, 2, \dots, N_s\} \\ S'_1 &\triangleq \{i : z'_i = 1, \quad i = 1, 2, \dots, N_s\} \\ S'_2 &\triangleq \{i : z'_i \in (0, 1), \quad i = 1, 2, \dots, N_s\} \end{aligned} \quad (50)$$

and let $N \triangleq |S'_1|$. Then, we have

$$|S'_2| = M > 2 \quad (51)$$

$$|S'_0| = N_s - M - N \quad (52)$$

$$0 \leq N < K < N + M \leq N_s \quad (53)$$

$$\sum_{i \in S'_2} z'_i = K - N. \quad (54)$$

Also, define set $C_{S'_2}$ as the indices of $K - N$ elements of S'_2 with minimum c_i 's (i.e., cheapest sensors). Similarly, let $E_{S'_2}$ have the indices of $K - N$ elements of S'_2 with maximum c_i 's (i.e., most expensive sensors), where ties are broken arbitrarily. It is clear that

$$\sum_{i \in C_{S'_2}} c_i \leq \sum_{i \in S'_2} z'_i c_i \leq \sum_{i \in E_{S'_2}} c_i. \quad (55)$$

Starting with the set of indices $X_0 = C_{S'_2}$, let

$$\begin{aligned} X_{t+1} &= (X_t \setminus \{m_t\}) \cup \{n_t\} \\ m_t &= \arg \min_i c_i, \quad i \in X_t \\ n_t &= \arg \max_i c_i, \quad i \in S'_2 \setminus X_t \end{aligned} \quad (56)$$

Note that $c_{m_t} \leq c_{n_t}$. For some integer T , where $0 \leq T < \min\{(K - N), (M + N - K)\}$, the following relation holds:

$$\sum_{i \in X_T} c_i \leq \sum_{i \in S'_2} z'_i c_i \leq \sum_{i \in X_{T+1}} c_i. \quad (57)$$

It is possible to find $\alpha \in [0, 1)$ such that

$$\sum_{i \in X_T \setminus \{m_T\}} c_i + (1 - \alpha) c_{m_T} + \alpha c_{n_T} = \sum_{i \in S'_2} z'_i c_i. \quad (58)$$

In particular,

$$\alpha = \begin{cases} 0, & c_{n_T} = c_{m_T} \\ \frac{\sum_{i \in S'_2} z'_i c_i - \sum_{i \in X_T} c_i}{c_{n_T} - c_{m_T}}, & c_{n_T} > c_{m_T} \end{cases}. \quad (59)$$

Let $S'_{21} = X_T \setminus \{m_T\}$ and $S'_{20} = S'_2 \setminus (X_T \cup \{n_T\})$, and consider selection vector $\hat{\mathbf{z}}$ with

$$z_i^* = \begin{cases} 0, & i \in S'_0 \cup S'_{20} \\ 1, & i \in S'_1 \cup S'_{21} \\ \alpha, & i = n_T \\ 1 - \alpha, & i = m_T \end{cases}. \quad (60)$$

Here, we basically split up S'_2 into four disjoint sets as

$$S'_2 = S'_{20} \cup S'_{21} \cup \{m_T\} \cup \{n_T\} \quad (61)$$

Also it is noted that

$$|S'_{21}| = K - N - 1. \quad (62)$$

In the following, it is shown that \mathbf{z}^* in (60) satisfies the condition in (42); i.e., the cost constraint.

$$\sum_{i=1}^{N_s} z_i^* c_i = \sum_{i \in S'_1 \cup S'_{21}} c_i + \alpha c_{n_T} + (1 - \alpha) c_{m_T} \quad (63)$$

$$= \sum_{i \in S'_1} c_i + \sum_{i \in X_T \setminus \{m_T\}} c_i + \alpha c_{n_T} + (1 - \alpha) c_{m_T} \quad (64)$$

$$= \sum_{i \in S'_1} c_i + \sum_{i \in S'_2} z'_i c_i \quad (65)$$

$$= \sum_{i=1}^{N_s} z'_i c_i \quad (66)$$

$$= C_T \quad (67)$$

where (63) follows from (60), (64) is due to the definition of S'_{21} , (65) is based on (58), (66) follows from definitions in (50) and finally (67) is due to Lemma 2.

To prove that \mathbf{z}^* is a solution that satisfies the property in (17) and achieves the same objective value as \mathbf{z}' , consider following equalities:

$$v' = \sum_{i=1}^{N_s} z'_i p_i \quad (68)$$

$$= \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z'_i p_i \quad (69)$$

$$= \sum_{i \in S'_1} p_i + \mu \sum_{i \in S'_2} z'_i c_i + \nu \sum_{i \in S'_2} z'_i \quad (70)$$

$$= \sum_{i \in S'_1} p_i + \mu \left(\sum_{i \in X_T \setminus \{m_T\}} c_i + (1 - \alpha) c_{m_T} + \alpha c_{n_T} \right) + \nu (K - N) \quad (71)$$

$$= \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z_i^* (\mu c_i + \nu) \quad (72)$$

$$= \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z_i^* p_i \quad (73)$$

$$= \sum_{i=1}^{N_s} z_i^* p_i = v^* \quad (74)$$

where (69) follows from the definitions in (50), (70) is due to (49), (71) is based on (50) and (58), (72) follows from (60)–(62), (73) is due to (49), and finally (74) is based on (60) and (61).

Based on the preceding arguments, it is proved that \mathbf{z}^* is also a solution to (14). From (60), it is noted that \mathbf{z}^* satisfies (17), where

$$\begin{aligned} S_0 &= S'_0 \cup S'_{20} \\ S_1 &= S'_1 \cup S'_{21} \\ S_2 &= \{n_T, m_T\} \end{aligned} \quad (75)$$

Suppose there exists a solution \mathbf{z}' to (14) ($\sum_{i=1}^{N_s} z'_i c_i = C_T$) such that $\sum_{i=1}^{N_s} z'_i < K$, where \mathbf{z}' does not satisfy (19); i.e., \mathbf{z}' has $M > 1$ non-integer components. We argue that there exists another solution satisfying (19). Define sets of indices S'_0 , S'_1 and S'_2 as in (50). Then, the following relation holds:

$$0 \leq \sum_{i \in S'_2} z'_i c_i < \sum_{i \in S'_2} c_i \quad (76)$$

Let $X_0 = \emptyset$ and

$$\begin{aligned} X_{t+1} &= X_t \cup \{n_t\} \\ n_t &= \arg \max_i c_i, \quad i \in S'_2 \setminus X_t \end{aligned} \quad (77)$$

Here, X_t represents the set of indices in S'_2 with t largest c_i 's. Hence $|X_t| = t$. Then, for some integer T , where $0 \leq T \leq \lfloor \sum_{i \in S'_2} z'_i \rfloor$, we get

$$\sum_{i \in X_T} c_i \leq \sum_{i \in S'_2} z'_i c_i \leq \sum_{i \in X_{T+1}} c_i \quad (78)$$

It is possible to find $\alpha \in [0, 1]$ such that

$$\sum_{i \in X_T} c_i + \alpha c_{n_T} = \sum_{i \in S'_2} z'_i c_i \quad (79)$$

where

$$\alpha = \frac{\sum_{i \in S'_2} z'_i c_i - \sum_{i \in X_T} c_i}{c_{n_T}}. \quad (80)$$

Since $\sum_{i \in S'_2} z'_i c_i \leq \sum_{i \in X_T} c_i + (\sum_{i \in S'_2} z'_i - T) c_{n_T}$, we have

$$\begin{aligned} \alpha &\leq \frac{\left(\sum_{i \in X_T} c_i + (\sum_{i \in S'_2} z'_i - T) c_{n_T} \right) - \sum_{i \in X_T} c_i}{c_{n_T}} \\ &= \frac{(\sum_{i \in S'_2} z'_i - T) c_{n_T}}{c_{n_T}} \\ &= \sum_{i \in S'_2} z'_i - T \end{aligned} \quad (81)$$

Let

$$S'_{21} = X_T, \quad S'_{20} = S'_2 \setminus (X_T \cup \{n_T\}), \quad (82)$$

split S'_2 into four disjoint sets as

$$S'_2 = S'_{20} \cup S'_{21} \cup \{n_T\} \quad (83)$$

and consider the selection vector \mathbf{z}^* defined as

$$z_i^* = \begin{cases} 0, & i \in S'_0 \cup S'_{20} \\ 1, & i \in S'_1 \cup S'_{21} \\ \alpha, & i = n_T \end{cases} \quad (84)$$

It is noted that \mathbf{z}^* satisfies the total number of sensors constraint as shown below:

$$\sum_{i=1}^{N_s} z_i^* = |S'_1| + |S'_{21}| + \alpha \quad (85)$$

$$\leq |S'_1| + |X_T| + \sum_{i \in S'_2} z'_i - T \quad (86)$$

$$= |S'_1| + \sum_{i \in S'_2} z'_i \quad (87)$$

$$= \sum_{i=1}^{N_s} z'_i \leq K \quad (88)$$

Here, (85) follows from (84), (86) is due to (81), (87) follows from (77), and finally (88) is based on (50).

The total cost constraint is also satisfied since

$$\sum_{i=1}^{N_s} z_i^* c_i = \sum_{i \in S'_1 \cup S'_{21}} c_i + \alpha c_{n_T} \quad (89)$$

$$= \sum_{i \in S'_1} c_i + \sum_{i \in X_T} c_i + \alpha c_{n_T} \quad (90)$$

$$= \sum_{i \in S'_1} c_i + \sum_{i \in S'_2} z'_i c_i \quad (91)$$

$$= \sum_{i=1}^{N_s} z'_i c_i = C_T \quad (92)$$

where (89) follows from the definition of $\hat{\mathbf{z}}$ in (84), (90) is based in (50), (91) is due to (79), and (92) follows from the definition of \mathbf{z}' .

Also, \mathbf{z}^* achieves the same objective value as \mathbf{z}' since

$$v' = \sum_{i=1}^{N_S} z'_i p_i = \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z'_i p_i \quad (93)$$

$$= \sum_{i \in S'_1} p_i + \mu \sum_{i \in S'_2} z'_i c_i \quad (94)$$

$$= \sum_{i \in S'_1} p_i + \mu \left(\sum_{i \in X_T} c_i + \alpha c_{n_T} \right) \quad (95)$$

$$= \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z_i^* (\mu c_i) \quad (96)$$

$$= \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z_i^* p_i \quad (97)$$

$$= \sum_{i=1}^{N_S} z_i^* p_i = v^* \quad (98)$$

where (93) is based on (50), (94) and (97) follow from (49), (95) is due to (79), and (96) is based on (84) and (83).

Overall, it is shown that if there exists a solution to (14) that does not satisfy neither *a*) (i.e., (17) and (18)) nor *b*) (i.e., (19) and (20)), then there always exists an alternative solution which satisfies either *a*) or *b*) in the proposition. Thus, when the relaxed problem in (14) is feasible, there exists a solution satisfying either *a*) or *b*), as claimed in the proposition.

C. Proof of Proposition 2

To simplify notation, we denote the objective function in (35) as $F(\mathbf{z}, \mathbf{c})$, that is,

$$F(\mathbf{z}, \mathbf{c}) \triangleq \sum_{i=1}^{N_S} z_i \frac{\mu_i^2 (2^{2c_i} - 1)}{\sigma_{n_i}^2 2^{2c_i}}. \quad (99)$$

Let (\mathbf{z}, \mathbf{c}) be any feasible solution to the problem in (35). We argue that there exists $\mathbf{c}' = [c'_1, c'_2, \dots, c'_{N_s}]^T$ such that

$$F(\mathbf{z}^*, \mathbf{c}') \geq F(\mathbf{z}, \mathbf{c}) \quad (100)$$

and $(\mathbf{z}^*, \mathbf{c}')$ is feasible, where \mathbf{z}^* is as defined in (37). To justify this claim, we define \mathbf{c}' in the following way: First, consider any bijection g from $\{i \mid z_i^* = 1, z_i = 0\}$ to $\{j \mid z_j^* = 0, z_j = 1\}$. Then, let

$$c'_i = \begin{cases} c_i, & z_i^* = 1, z_i = 1 \\ c_{g(i)}, & z_i^* = 1, z_i = 0 \\ 0, & \text{else} \end{cases}, \quad i = 1, 2, \dots, N_s. \quad (101)$$

It is noted that $c'_i \geq 0$ for $i = 1, 2, \dots, N_s$. Also,

$$(\mathbf{z}^*)^T \mathbf{c}' = \sum_{i \in \{k \mid z_k^* = 1, z_k = 1\}} c_i + \sum_{i \in \{k \mid z_k^* = 1, z_k = 0\}} c_{g(i)} \quad (102)$$

$$= \sum_{i \in \{k \mid z_k^* = 1, z_k = 1\}} c_i + \sum_{i \in \{k \mid z_k^* = 0, z_k = 1\}} c_i \quad (103)$$

$$= \mathbf{z}^T \mathbf{c} \leq C_T. \quad (104)$$

Hence, $(\mathbf{z}^*, \mathbf{c}')$ is feasible. Then, we consider the objective values, $F(\mathbf{z}^*, \mathbf{c}')$ and $F(\mathbf{z}, \mathbf{c})$, and compare them as follows:

$$F(\mathbf{z}^*, \mathbf{c}') = \sum_{i \in \{k \mid z_k^* = 1, z_k = 1\}} \frac{\mu_i^2 (2^{2c_i} - 1)}{\sigma_{n_i}^2 2^{2c_i}} + \sum_{i \in \{k \mid z_k^* = 1, z_k = 0\}} \frac{\mu_i^2 (2^{2c_{g(i)}} - 1)}{\sigma_{n_i}^2 2^{2c_{g(i)}}} \quad (105)$$

$$= \sum_{i \in \{k \mid z_k = 1\}} \frac{\mu_i^2 (2^{2c_i} - 1)}{\sigma_{n_i}^2 2^{2c_i}} - \sum_{i \in \{k \mid z_k^* = 0, z_k = 1\}} \frac{\mu_i^2 (2^{2c_i} - 1)}{\sigma_{n_i}^2 2^{2c_i}} + \sum_{i \in \{k \mid z_k^* = 1, z_k = 0\}} \frac{\mu_i^2 (2^{2c_{g(i)}} - 1)}{\sigma_{n_i}^2 2^{2c_{g(i)}}} \quad (106)$$

$$= F(\mathbf{z}, \mathbf{c}) - \sum_{i \in \{k \mid z_k^* = 0, z_k = 1\}} \frac{\mu_i^2 (2^{2c_i} - 1)}{\sigma_{n_i}^2 2^{2c_i}} + \sum_{i \in \{k \mid z_k^* = 1, z_k = 0\}} \frac{\mu_i^2 (2^{2c_{g(i)}} - 1)}{\sigma_{n_i}^2 2^{2c_{g(i)}}} \quad (107)$$

$$= F(\mathbf{z}, \mathbf{c}) - \sum_{i \in \{k \mid z_k^* = 1, z_k = 0\}} \frac{\mu_{g(i)}^2 (2^{2c_{g(i)}} - 1)}{\sigma_{n_{g(i)}}^2 2^{2c_{g(i)}}} + \sum_{i \in \{k \mid z_k^* = 1, z_k = 0\}} \frac{\mu_i^2 (2^{2c_{g(i)}} - 1)}{\sigma_{n_i}^2 2^{2c_{g(i)}}} \quad (108)$$

$$= F(\mathbf{z}, \mathbf{c}) + \sum_{i \in \{k \mid z_k^* = 1, z_k = 0\}} \left(\frac{2^{2c_{g(i)}} - 1}{2^{2c_{g(i)}}} \right) \left(\frac{\mu_i^2}{\sigma_{n_i}^2} - \frac{\mu_{g(i)}^2}{\sigma_{n_{g(i)}}^2} \right) \quad (109)$$

$$\geq F(\mathbf{z}, \mathbf{c}) \quad (110)$$

By definition, $\frac{\mu_i^2}{\sigma_{n_i}^2} \geq \frac{\mu_{g(i)}^2}{\sigma_{n_{g(i)}}^2}$ for $i \in \{k \mid z_k^* = 1, z_k = 0\}$ and $g(i) = j \in \{k \mid z_k^* = 0, z_k = 1\}$ since $i \in \tilde{B}$ and $j \notin \tilde{B}$. Moreover, $\left(\frac{2^x - 1}{2^x} \right) \geq 0$ for $x \geq 0$. Therefore, each element in the summation term in (109) is larger than or equal to zero. Hence, the inequality in (110) follows. Finally, we have

$$F(\mathbf{z}^*, \mathbf{c}^*) \geq F(\mathbf{z}^*, \mathbf{c}') \geq F(\mathbf{z}, \mathbf{c}) \quad (111)$$

since \mathbf{c}^* is a solution of (29) for \mathbf{z}^* . Thus, we conclude that no feasible solution (\mathbf{z}, \mathbf{c}) produces an objective value larger than $(\mathbf{z}^*, \mathbf{c}^*)$, which makes $(\mathbf{z}^*, \mathbf{c}^*)$ an optimal solution.

REFERENCES

- [1] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [2] E. Laz and S. Gezici, "Centralized and decentralized detection with cost-constrained measurements," *Signal Processing*, vol. 132, pp. 8–18, 2017.
- [3] B. Dulek and S. Gezici, "Average fisher information maximisation in presence of cost-constrained measurements," *Electronics Letters*, vol. 47, no. 11, pp. 654–656, 2011.
- [4] —, "Cost minimization of measurement devices under estimation accuracy constraints in the presence of gaussian noise," *Digital Signal Processing*, vol. 22, no. 5, pp. 828–840, 2012.
- [5] K. Liu and A. M. Sayeed, "Optimal distributed detection strategies for wireless sensor networks," in *Proc. 42nd Annual Allerton Conference on Commun., Control and Comp.*, 2004.
- [6] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Energy-efficient detection in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 693–702, 2005.

- [7] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-part i: Gaussian case," *IEEE transactions on signal processing*, vol. 54, no. 3, pp. 1131–1143, 2006.
- [8] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, 2006.
- [9] S. Cui, J.-J. Xiao, A. J. Goldsmith, Z.-Q. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4683–4695, 2007.
- [10] J. Li and G. AlRegib, "Rate-constrained distributed estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1634–1643, 2007.
- [11] G. Thatte and U. Mitra, "Sensor selection and power allocation for distributed estimation in sensor networks: Beyond the star topology," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2649–2661, 2008.
- [12] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, 2006.
- [13] 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, 2014.
- [14] L. M. Kaplan, "Global node selection for localization in a distributed sensor network," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 113–135, 2006.
- [15] S. Jiang, R. Kumar, and H. E. Garcia, "Optimal sensor selection for discrete-event systems with partial observation," *IEEE Transactions on Automatic Control*, vol. 48, no. 3, pp. 369–381, 2003.
- [16] C. Perera, A. Zaslavsky, P. Christen, M. Compton, and D. Georgakopoulos, "Context-aware sensor search, selection and ranking model for internet of things middleware," in *2013 IEEE 14th international conference on mobile data management*, vol. 1. IEEE, 2013, pp. 314–322.
- [17] D. C. Kammer, "Sensor placement for on-orbit modal identification and correlation of large space structures," *Journal of Guidance, Control, and Dynamics*, vol. 14, no. 2, pp. 251–259, 1991.
- [18] L. Yao, W. A. Sethares, and D. C. Kammer, "Sensor placement for on-orbit modal identification via a genetic algorithm," *AIAA journal*, vol. 31, no. 10, pp. 1922–1928, 1993.
- [19] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration for tracking applications," *IEEE Signal processing magazine*, vol. 19, no. 2, pp. 61–72, 2002.
- [20] F. Zhao, L. J. Guibas, and L. Guibas, *Wireless Sensor Networks: An Information Processing Approach*. Morgan Kaufmann, 2004.
- [21] E. Ertin, J. W. Fisher, and L. C. Potter, "Maximum mutual information principle for dynamic sensor query problems," in *Information processing in sensor networks*. Springer, 2003, pp. 405–416.
- [22] M. Chu, H. Haussecker, and F. Zhao, "Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks," *The International Journal of High Performance Computing Applications*, vol. 16, no. 3, pp. 293–313, 2002.
- [23] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4938–4953, 2011.
- [24] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.
- [25] H. Wang, K. Yao, G. Pottie, and D. Estrin, "Entropy-based sensor selection heuristic for target localization," in *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*, 2004, pp. 36–45.
- [26] V. Isler and R. Bajcsy, "The sensor selection problem for bounded uncertainty sensing models," in *Proceedings of the 4th international symposium on Information processing in sensor networks*. IEEE Press, 2005, p. 20.
- [27] H. Rowaihy, S. Eswaran, M. Johnson, D. Verma, A. Bar-Noy, T. Brown, and T. La Porta, "A survey of sensor selection schemes in wireless sensor networks," in *Unattended Ground, Sea, and Air Sensor Technologies and Applications IX*, vol. 6562, 2007, p. 65621A.
- [28] M. Lázaro, M. Sánchez-Fernández, and A. Artés-Rodríguez, "Optimal sensor selection in binary heterogeneous sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1577–1587, 2009.
- [29] Y. Oshman, "Optimal sensor selection strategy for discrete-time state estimators," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 2, pp. 307–314, 1994.
- [30] R. Debouk, S. LaFortune, and D. Teneketzis, "On an optimization problem in sensor selection," *Discrete Event Dynamic Systems*, vol. 12, no. 4, pp. 417–445, 2002.
- [31] Y. Selen, H. Tullberg, and J. Kronander, "Sensor selection for cooperative spectrum sensing," in *2008 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*. IEEE, 2008, pp. 1–11.
- [32] P. Zappi, C. Lombriser, T. Stiefmeier, E. Farella, D. Roggen, L. Benini, and G. Tröster, "Activity recognition from on-body sensors: accuracy-power trade-off by dynamic sensor selection," in *European Conference on Wireless Sensor Networks*. Springer, 2008, pp. 17–33.
- [33] S. P. Chepuri and G. Leus, "Sparse sensing for distributed detection," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1446–1460, 2016.
- [34] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. MIT Press, 2009.
- [35] A. Schrijver, *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.
- [36] E. L. Lawler and D. E. Wood, "Branch-and-bound methods: A survey," *Operations Research*, vol. 14, no. 4, pp. 699–719, 1966.
- [37] A. Ozcelikkale, H. M. Ozaktas, and E. Arıkan, "Signal recovery with cost-constrained measurements," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3607–3617, 2010.
- [38] F. Bian, D. Kempe, and R. Govindan, "Utility based sensor selection," in *Proceedings of the 5th International Conference on Information Processing in Sensor Networks*, 2006, pp. 11–18.
- [39] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [40] G. B. Dantzig, *Linear Programming and Extensions*. Princeton University Press, 1998.
- [41] B. Oymak, "Cost constrained sensor selection and design for binary hypothesis testing," *M.S. Thesis, Bilkent University*, Jan. 2020. [Online]. Available: www.ee.bilkent.edu.tr/~gezici/thesis.pdf
- [42] R. Robere, "Interior point methods and linear programming," *University of Toronto*, pp. 1–15, 2012.
- [43] J. A. Kelner and D. A. Spielman, "A randomized polynomial-time simplex algorithm for linear programming," in *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing*, 2006, pp. 51–60.



Berkay Oymak received the B.S. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University, Ankara, Turkey, in 2016 and 2020, respectively. His research interests are in the areas of estimation and detection theory.



and estimation theory, and communication theory.

Berkan Dulek received the B.S., M.S. and Ph.D. degrees in electrical and electronics engineering from Bilkent University in 2003, 2006 and 2012, respectively. From 2007 to 2010, he worked at industry. From 2012 to 2013, he was a postdoctoral research associate at the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY. Since 2014, he has been with the Department of Electrical and Electronics Engineering at Hacettepe University, where he is currently an Associate Professor. His research interests are in statistical signal processing, detection



Sinan Gezici (S'03–M'06–SM'11) received the B.S. degree from Bilkent University, Turkey in 2001, and the Ph.D. degree in Electrical Engineering from Princeton University in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA. Since 2007, he has been with the Department of Electrical and Electronics Engineering at Bilkent University, where he is currently a Professor. Dr. Gezici's research interests are in the areas of detection and estimation theory, wireless communications, and localization systems. Among his publications in these areas is the book *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols* (Cambridge University Press, 2008). Dr. Gezici was an associate editor for *IEEE Transactions on Communications*, *IEEE Wireless Communications Letters*, and *Journal of Communications and Networks*.