COST CONSTRAINED SENSOR SELECTION AND DESIGN FOR BINARY HYPOTHESIS TESTING

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

ELECTRICAL AND ELECTRONICS ENGINEERING

By

Berkay Oymak

January 2020

Cost Constrained Sensor Selection and Design for Binary Hypothesis Testing

By Berkay Oymak

January 2020

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Sinan Gezici (Advisor)

Orhan Arıkan (Co-Advisor)

Berkan Dülek

Tolga Mete Duman

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

COST CONSTRAINED SENSOR SELECTION AND DESIGN FOR BINARY HYPOTHESIS TESTING

Berkay Oymak
M.S. in Electrical and Electronics Engineering
Advisor: Sinan Gezici
Co-Advisor: Orhan Arıkan
January 2020

We consider a sensor selection problem for binary hypothesis testing with cost-constrained measurements. Random observations 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.

Keywords: Detection, sensor selection, cost constraint.

ÖZET

İKİLİ HİPOTEZ TESTİ İÇİN BÜTÇE KISITLI SENSÖR SEÇİMİ VE TASARIMI

Berkay Oymak
Elektrik ve Elektronik Mühendisliği, Yüksek Lisans
Tez Danışmanı: Sinan Gezici
İkinci Tez Danışmanı: Orhan Arıkan
Ocak 2020

Bütçe kısıtlı ölçümlere dayanan ikili hipotez testi için sensör seçme problemi incelenmektedir. Ölçümlenen rastsal gözlemler, ilgili parametre vektörüne bağlı olarak Gauss gürültülü doğrusal bir sistem tarafından oluşturulmaktadır. Kısıtlı sayıda sensör tarafından yapılan ölçümlere dayanarak, parametre vektörünün durumunun tespit edilmesi amaçlanmaktadır. Her bir sensör ölçümünün maliyeti, sensörün güvenilir olarak ayırt edebildiği genlik seviyesi sayısıyla orantılıdır. Toplam ölçüm bütçesine ve kullanılabilecek sensör sayısına üst limit getirilerek, ikili hipotez testinin tespit başarımını maksimize etmek üzere sensör seçme problemi kurgulanmaktadır. Bu optimizasyon probleminin doğrusal gevşetmeye uğramış halinin çözümünün karakterize edilmesiyle en iyiye yakın başarıma ulaşan ve etkili bir şekilde hesaplanabilen bir çözüm algoritması önerilmektedir. Ölçüm maliyetlerinin sabit olduğu duruma ek olarak, bu maliyetlerin tasarım parametreleri kabul edildiği durum ele alınmaktadır. Bilhassa, tespit için kullanılan sensörlerin ölçüm hassasiyetini tasarlamak aracılığıyla toplam bütçenin kısıtlı sayıda sensöre paylaştırılması problemi incelenmektedir. Problemin en iyi sonucu kapalı formda elde edilmektedir. Önerilen çözümlerin başarımları sayısal örnekler aracılığıyla sunulmaktadır.

Anahtar sözcükler: Tespit, sensör seçme, maliyet kısıtı.

Acknowledgement

I would like to express my gratitude to my advisors Prof. Sinan Gezici and Prof. Orhan Arıkan for their great support throughout my graduate studies. Their immense expertise and guidance made this thesis possible. I am thankful to Assoc. Prof. Berkan Dülek for his invaluable contributions to this work.

I would like to thank my thesis committee member Prof. Tolga Mete Duman for his time and valuable feedback.

I am deeply grateful to my mother for standing beside me through all struggles and I want to express my appreciation for her never-ending kindness and forgiveness towards life as a whole. I am thankful and owe greatest of respects to my father for he never ceased his support and care and taught me most valuable of lessons. Blessings be my grandfathers' for setting the utmost examples of a virtuous life. Finally, I would like to express my gratitude to my grandmothers for their loving and caring presence and for helping me become who I am today.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

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, number and quality of the sensors in the network.

1.2 Related Work

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], sensor selection problem is formulated for parameter estimation under Gaussian noise. A heuristic 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 an optimization problem is formulated in which the number of sensors employed for measuring the target position is constrained. 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 KL divergence and number of sensor of each class are linearly related. Results indicate that it is optimal to choose all sensors from the class with the best performance to cost ratio. In [29], 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.

1.3 Our Contribution

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 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 [33], which is known to be NP-complete [34]. Although there exist methods to find an optimal solution to such problems, such as the branch and bound method given in [34, 35], 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 thesis, 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 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 thesis can be summarized as follows:

- The problem of sensor selection using *cost-constrained* measurements is formulated in order to determine the binary state of a parameter vector so that the corresponding Bayes risk is minimized 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.

1.4 Organization of the Thesis

The rest of this thesis is organized as follows. In Chapter 2, we present the linear system that generates the random observations and the measurement model used to acquire the samples employed for detection. In Chapter 3, an approximate solution is developed for determining which sensors (at most K out of N_s) are employed to collect the measurements when the cost of each sensor measurement is given. In Chapter 4, we analyze the problem of joint sensor selection and design. In Chapter 5, we provide numerical examples to evaluate the performance of the proposed methods. We conclude with some remarks in Chapter 6.

Chapter 2

System Model

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

$$x_i = \mathbf{h}_i^T \mathbf{\Theta} + n_i\,, \quad i = 1, \ldots, N_s \tag{2.1}$$

where n_i is the system noise in the ith observation and $\mathbf{h}_i$ is an $L \times 1$ vector representing the coefficients of the linear system related to observation i. The observations in (2.1) can be measured by N_s potential sensors as follows:

$$y_i = x_i + m_i\,, \quad i = 1, \ldots, N_s \tag{2.2}$$

where m_i is the measurement noise of the ith sensor (which can also be considered as quantization noise since a measurement error commonly occurs due to the finite number of quantization levels in a measurement device [36]).

In a more compact manner, the observations in (2.1) and the potential measurements in (2.2) can be expressed as

$$\mathbf{x} = \mathbf{H}^T \mathbf{\Theta} + \mathbf{n} \;\text{ and }\; \mathbf{y} = \mathbf{x} + \mathbf{m}\,, \tag{2.3}$$

respectively, where $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{N_s}]$ is the $L \times N_s$ system matrix, $\mathbf{x} = [x_1, x_2, \ldots, x_{N_s}]^T$, $\mathbf{n} = [n_1, n_2, \ldots, n_{N_s}]^T$, $\mathbf{y} = [y_1, y_2, \ldots, y_{N_s}]^T$, and $\mathbf{m} = [m_1, m_2, \ldots, m_{N_s}]^T$.

As in [4] and [24], the noise components are modeled as independent Gaussian random variables with zero mean, that is, $n_i \sim \mathcal{N}(0,\, \sigma^2_{n_i})$ and $m_i \sim \mathcal{N}(0,\, \sigma^2_{m_i})$ for $i = 1, \ldots, N_s$. In the vector notation, $\mathbf{n} \sim \mathcal{N}(\mathbf{0},\, \mathbf{\Sigma})$ and $\mathbf{m} \sim \mathcal{N}(\mathbf{0},\, \mathbf{\Sigma_m})$, where $\mathbf{\Sigma} = \mathrm{diag}\{\sigma^2_{n_1}, \sigma^2_{n_2}, \ldots, \sigma^2_{n_{N_s}}\}$ and $\mathbf{\Sigma}_m = \mathrm{diag}\{\sigma^2_{m_1}, \sigma^2_{m_2}, \ldots, \sigma^2_{m_{N_s}}\}$. In addition, it is assumed that the measurement noise $\mathbf{m}$ is independent of the system noise $\mathbf{n}$.

Chapter 3

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 available observations for measurement in an optimal manner, which is called the sensor selection problem in the literature [13,15,23,24, 26,31,32,37]. 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* $\mathbf{z} = [z_1, z_2, \ldots, z_{N_s}]^T$ that specifies whether the ith sensor is selected (i.e., $z_i = 1$ if the ith sensor is selected and $z_i = 0$ otherwise). We denote the number 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 $f : \{1, 2, \ldots, k\} \rightarrow \{1, 2, \ldots, N_s\}$, where $f(i)$ denotes the index of the ith selected sensor. Then, we construct a $k \times N_s$ *selection matrix* $\mathbf{Z}$, in which k of the columns are unit vectors $\mathbf{e}_{N_s,1}, \mathbf{e}_{N_s,2} \ldots, \mathbf{e}_{N_s,k}$ ($\mathbf{e}_{j,i}$ is defined as a column vector of length j and it has a 1 at the ith 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:

$$row_i(\mathbf{Z}) = \mathbf{e}^T_{k,f(i)}\,, \quad i = 1, 2, \ldots, k \tag{3.1}$$

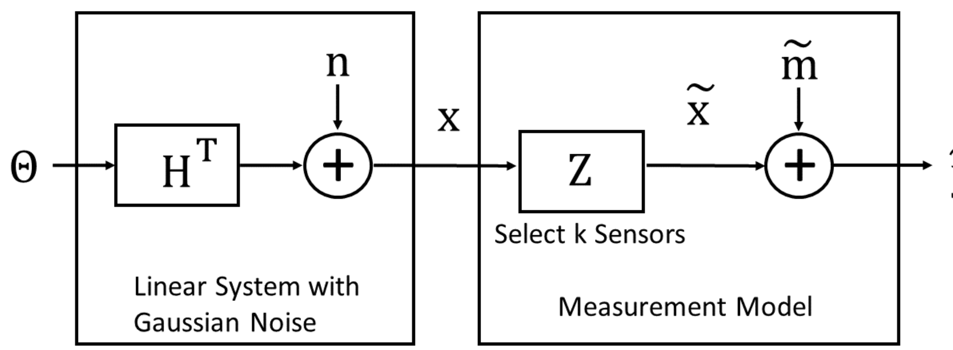

Figure 3.1: System block diagram.

where $row_i(\mathbf{Z})$ denotes the ith row of $\mathbf{Z}$. Also, $\mathbf{z}$ can be obtained from $\mathbf{Z}$ simply as $\mathbf{z} = \mathrm{diag}(\mathbf{Z}^T\mathbf{Z})$, where $\mathrm{diag}(\mathbf{Z}^T\mathbf{Z})$ 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 observations 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

$$\tilde{\mathbf{y}} \triangleq \mathbf{Z}\mathbf{y} = \mathbf{Z}\mathbf{x} + \mathbf{Z}\mathbf{m} \triangleq \tilde{\mathbf{x}} + \tilde{\mathbf{m}}\,. \tag{3.2}$$

Namely, k out of N_s observations are measured via k sensors. The resulting system and measurement model is illustrated in Fig. 3.1.

For the cost of making a sensor measurement, we employ the measurement cost model proposed in [36]. In that model, the cost of a measurement is determined by the number of quantization levels it can reliably distinguish, which is related to the ratio of system and measurement noise variances. Specifically, the cost of making a measurement via sensor i is given by [36]

$$c_i = 0.5\,\log_2\left(1 + \frac{\sigma_{n_i}^2}{\sigma_{m_i}^2}\right) \tag{3.3}$$

Similar to [2–4], we consider the expression in (3.3) 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 with observation x_i, a higher cost is associated with a more accurate measurement, y_i (see (2.2)).

Suppose that the parameter vector $\boldsymbol{\Theta}$ takes one of two possible values. Namely, there exist two hypotheses defined as $\mathcal{H}_0 : \boldsymbol{\Theta} = \boldsymbol{\Theta}_0$ and $\mathcal{H}_1 : \boldsymbol{\Theta} = \boldsymbol{\Theta}_1$, 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 (3.2) can be specified, based on the system model in Chapter 2, as

$$\tilde{\mathbf{y}} \,|\, \mathcal{H}_i \sim \mathcal{N}\left(\mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_i\,,\,\mathbf{Z}\left(\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_m\right)\mathbf{Z}^T\right) \tag{3.4}$$

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]:

$$r(\delta_B) = \pi_0 Q\left(\frac{\ln(\pi_0/\pi_1)}{d} + \frac{d}{2}\right) + \pi_1 Q\left(\frac{d}{2} - \frac{\ln(\pi_0/\pi_1)}{d}\right) \tag{3.5}$$

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

$$\begin{aligned} d \triangleq \Big(& \left(\mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_\mathbf{1} - \mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_\mathbf{0}\right)^T \left(\mathbf{Z}(\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_\mathbf{m})\mathbf{Z}^T\right)^{-1} \\ & \times \left(\mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_\mathbf{1} - \mathbf{Z}\mathbf{H}^T\boldsymbol{\Theta}_\mathbf{0}\right)\Big)^{1/2} \end{aligned} \tag{3.6}$$

The expression in (3.6) can also be written as

$$\begin{aligned} d = \Big(& (\boldsymbol{\Theta}_\mathbf{1} - \boldsymbol{\Theta}_\mathbf{0})^T\mathbf{H}\mathbf{Z}^T \left(\mathbf{Z}(\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_\mathbf{m})\mathbf{Z}^T\right)^{-1} \\ & \times \mathbf{Z}\mathbf{H}^T(\boldsymbol{\Theta}_\mathbf{1} - \boldsymbol{\Theta}_\mathbf{0})\Big)^{1/2} \end{aligned} \tag{3.7}$$

Based on the definition of the selection matrix $\mathbf{Z}$, d in (3.7) can be stated, after some manipulation, as

$$d = \sqrt{\sum_{i=1}^{N_s} z_i \frac{(\mathbf{h}_i^T(\boldsymbol{\Theta}_\mathbf{1} - \boldsymbol{\Theta}_\mathbf{0}))^2}{\sigma_{n_i}^2 + \sigma_{m_i}^2}} \tag{3.8}$$

The aim is to minimize the Bayes risk $r(\delta_B)$ under a total cost constraint by making measurements with at most K sensors. Since it is known that $r(\delta_B)$ in (3.5) 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:

$$\begin{aligned} \underset{\mathbf{z}}{\text{maximize}} \quad & \sum_{i=1}^{N_s} z_i p_i \\ \text{subject to} \quad & \sum_{i=1}^{N_s} z_i c_i \leq C_T \\ & \sum_{i=1}^{N_s} z_i \leq K \\ & z_i \in \{0,1\}, \quad i = 1,2,\ldots,N_s \end{aligned} \tag{3.9}$$

where c_i is given by (3.3) and p_i is defined as

$$p_i \triangleq \frac{\left(\mathbf{h}_i^T(\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0)\right)^2}{\sigma_{n_i}^2 + \sigma_{m_i}^2}. \tag{3.10}$$

Due to its combinatorial nature, the problem in (3.9) can be very complex to solve unless $\binom{N_s}{K}$ is small. To simplify the problem, the last constraint can be relaxed as $0 \leq z_i \leq 1$, $i = 1, \ldots, N_s$, and a suitable optimization algorithm can be employed to obtain a solution for $\mathbf{z}$. Then, the K largest elements of that solution can be used to determine the selected sensors (observations). Relaxing the last constraint, we obtain the following convex optimization problem:

$$\begin{aligned} \underset{\mathbf{z}}{\text{maximize}} \quad & \sum_{i=1}^{N_s} z_i p_i \\ \text{subject to} \quad & \sum_{i=1}^{N_s} z_i c_i \leq C_T \\ & \sum_{i=1}^{N_s} z_i \leq K \\ & 0 \leq z_i \leq 1, \quad i = 1,2,\ldots,N_s \end{aligned} \tag{3.11}$$

The problem in (3.11) is a linearly constrained linear optimization problem. Hence, it can be solved efficiently via linear/convex optimization algorithms [38] such as the simplex method [39] and the interior point method [33]. The solution to (3.11) provides a performance upper bound on the original problem in (3.9); hence, it can be used to evaluate performance of suboptimal solution methods. In addition, the solution to (3.11) can be used as an initial point for developing close-to-optimal solutions of (3.9) with low computational complexity, as discussed towards the end of this chapter.

It is possible to specify the form of an optimal solution to (3.11) based on theoretical analysis. Towards that aim, we first provide the following two lemmas:

Lemma 1. *Let $B_1, B_2, \ldots, B_{N_L}$ denote the sets of indices of K largest p_i's. Assume that there exists $j \in \{1, 2, \ldots, N_L\}$, such that $C_T \geq \sum_{i \in B_j} c_i$, where c_i is as defined in* (3.3)*. Then, $\mathbf{z}^*$ is a solution to* (3.11) *(and also to* (3.9)*), where the elements of $\mathbf{z}^*$ are given by*

$$z_i^* = \begin{cases} 0\,, & i \notin B_j \\ 1\,, & i \in B_j \end{cases} . \tag{3.12}$$

Proof: Please see Appendix A.1.

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* (3.11) *is feasible and let $B_1, B_2, \ldots, B_{N_L}$ denote the sets of indices of K largest p_i's. If $C_T < \sum_{i \in B_j} c_i$ for all $j \in \{1, 2, \ldots, N_L\}$, then there exists a solution $\mathbf{z}^*$ to* (3.11) *that satisfies*

$$\sum_{i=1}^{N_s} z_i^* c_i = C_T \tag{3.13}$$

Proof: Please see Appendix A.2.

Based on Lemma 1 and Lemma 2, the following proposition is obtained related to the solution of the relaxed problem in (3.11).

Proposition 1. *Suppose that the optimization problem in* (3.11) *is feasible. Then, there exists a solution* $\mathbf{z}^*$ *to* (3.11) *that is characterized as either of the following:*

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

$$z_i^* \in \begin{cases} \{0\}, & i \in S_0 \\ \{1\}, & i \in S_1 \\ [0,1], & i \in S_2 \end{cases}, \quad i = 1, 2, \ldots, N_s \tag{3.14}$$

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

$$\begin{aligned} &S_0 \cup S_1 \cup S_2 = \{1, 2, \ldots, N_s\}, \\ &|S_0| = N_s - K - 1, \quad |S_1| = K - 1, \quad |S_2| = 2. \end{aligned} \tag{3.15}$$

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

$$z_i^* \in \begin{cases} \{0\}, & i \in S_0 \\ \{1\}, & i \in S_1 \\ [0,1], & i \in S_2 \end{cases}, \quad i = 1, 2, \ldots, N_s \tag{3.16}$$

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

$$S_0 \cup S_1 \cup S_2 = \{1, 2, \ldots, N_s\}, \ |S_2| = 1 \tag{3.17}$$

Proof: Please see Appendix A.3.

Proposition 1 states that when the problem in (3.11) is feasible, its solution can be expressed to include at most two non-integer elements. To utilize Proposition 1 for obtaining a solution of (3.11), we first consider the following problem in which the number of selected sensors is forced to be equal to K.

$$\begin{aligned} \underset{\mathbf{z}}{\text{maximize}} \quad & \sum_{i=1}^{N_s} z_i p_i \\ \text{subject to} \quad & \sum_{i=1}^{N_s} z_i c_i \leq C_T \\ & \sum_{i=1}^{N_s} z_i = K \\ & 0 \leq z_i \leq 1, \quad i = 1, 2, \ldots, N_s \end{aligned} \tag{3.18}$$

For this problem, Proposition 1 implies that a solution $\mathbf{z}^*$ conforming to (3.14) and (3.15) 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 (3.18) in a low-complexity manner. An algorithm is proposed for this purpose, which is presented as Algorithm 1. The algorithm initially checks whether any set of sensors with K largest p_i's satisfy the cost constraint. If no such sets exist, 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 KKT conditions can be satisfied for this partition.

Algorithm 1 Solution of Problem in (3.18), Part 1

obtain $B_1, \ldots, B_{N_L}$ as sets of indices of K largest p_i's.
if $\exists\, k \in \{1, \ldots, N_L\}$ s.t. $\sum_{i \in B_k} c_i \leq C_T$ **then**
  $\hat{z}_i = 1,\ i \in B_k$
  $\hat{z}_i = 0,\ i \notin B_k$
else
  for all $\binom{N_s}{2}$ combinations of sensor indices a, b **do**
    if $c_a \neq c_b$ **then**
      init S'_0, S'_1, S'_2 as empty sets
      calculate $\mu = (p_a - p_b)/(c_a - c_b)$
      calculate $\nu = (p_b c_a - p_a c_b)/(c_a - c_b)$
      add every sensor index s that satisfies $p_s = \mu c_s + \nu$ to S'_2
      add every sensor index s that satisfies $p_s > \mu c_s + \nu$ to S'_1
      add remaining sensor indices to S'_0
      $M \leftarrow |S'_2|,\ \ N \leftarrow |S'_1|$
      if $N < K < N + M$ **then**
        let $C_{S'_2}$ consist of indices of cheapest $(K - N)$ sensors in S'_2
        let $E_{S'_2}$ consist of indices of most expensive $(K - N)$ sensors in S'_2
        if $\sum_{i \in (S'_1 \cup C_{S'_2})} c_i \leq C_T$ and
        $C_T \leq \sum_{i \in (S'_1 \cup E_{S'_2})} c_i$ **then**
          $X_0 \leftarrow C_{S'_2},\ t = 0$

Algorithm 2 Solution of Problem in (3.18), Part 2

while $\sum_{i \in (S'_1 \cup X_t)} c_i \leq C_T$ and
$t < \min\{K - N, M + N - K\}$ **do**
  let m_t be index of cheapest sensor in X_t
  let n_t be index of most expensive sensor in $S'_2 \setminus X_t$
  $X_{t+1} = (X_t \setminus \{m_t\}) \cup \{n_t\}$
  $t = t + 1$
end while
$T = t - 1$
$S'_{21} = X_T \setminus \{m_T\}$
$S'_{20} = S'_2 \setminus (X_T \cup \{n_T\})$
if $c_{m_T} = c_{n_T}$ **then**
  $\alpha = 0$
else
  $\alpha = \frac{(C_T - \sum_{i \in (S'_1 \cup S'_{21})} c_i - c_{m_T})}{(c_{n_T} - c_{m_T})}$
end if
$\hat{z}_i = 1, \quad i \in S'_1 \cup S'_{21}$
$\hat{z}_i = 0, \quad i \in S'_0 \cup S'_{20}$
$\hat{z}_i = \alpha, \quad i = n_T$
$\hat{z}_i = 1 - \alpha, \quad i = m_T$
break
end if
end if
end if
end for
end if

Remark. *If the number of sets that contain sensor indices with K largest p_i's is large, the computational complexity of Algorithm 1 can be high. However, such a scenario is not practical since p_i's are commonly distinct for different sensors as they depend on system parameters and noise levels (see (3.10)). For example, $\mathbf{h}_i$ can represent the channel for the ith observation.*

Although Algorithm 1 can be used to solve (3.18), it is not directly applicable to the relaxed problem in (3.11). However, we argue that, with a suitable change of parameters, a solution to (3.11) can be obtained by applying Algorithm 1 on an equivalent problem in the form of (3.18). To this aim, we define the following optimization problem:

$$\begin{aligned} \underset{\bar{\mathbf{z}}}{\text{maximize}} \quad & \sum_{i=1}^{\bar{N}_s} \bar{z}_i \bar{p}_i \\ \text{subject to} \quad & \sum_{i=1}^{\bar{N}_s} \bar{z}_i \bar{c}_i \leq C_T \\ & \sum_{i=1}^{\bar{N}_s} \bar{z}_i = K \\ & 0 \leq \bar{z}_i \leq 1, \quad i = 1, 2, \ldots, \bar{N}_s \end{aligned} \tag{3.19}$$

where

$$\begin{aligned} \bar{N}_s &\triangleq N_s + K \\ \bar{c}_i &\triangleq \begin{cases} c_i\,, & i = 1, 2, \ldots, N_s \\ 0\,, & i = N_s + 1, N_s + 2, \ldots, N_s + K \end{cases} \\ \bar{p}_i &\triangleq \begin{cases} p_i\,, & i = 1, 2, \ldots, N_s \\ 0\,, & i = N_s + 1, N_s + 2, \ldots, N_s + K \end{cases} \end{aligned} \tag{3.20}$$

with c_i and p_i being defined in (3.3) and (3.10), respectively. A solution $\mathbf{z}^*$ of the problem in (3.11) can be obtained from a solution $\bar{\mathbf{z}}^*$ of (3.19) as follows (please see Appendix A.4 for details):

$$z_i^* = \bar{z}_i^*,\ i = 1, 2, \ldots, N_s \tag{3.21}$$

It is important to note that the optimization problem in (3.19) can be solved via Algorithm 1 as it is in the same form as that in (3.18).

The main idea behind the problem formulation in (3.19) is to introduce K virtual observations, which induce no cost and no performance gain, in addition to the N_s actual observations. In this way, solving the new problem with $N_s + K$ observations by choosing exactly K sensors becomes equivalent to solving the relaxed problem in (3.11) by choosing less or equal to K sensors. This conclusion mainly comes from the introduction of slack variables to the problem in (3.11) as explained in Appendix A.4.

Based on our results related to the relaxed optimization problem in (3.11), we propose a suboptimal solution procedure for the original optimization problem in (3.9) as follows:

Proposed Suboptimal Solution to* (3.9)*:

1. Obtain the equivalent relaxed problem in the form of (3.19). Calculate its solution as $\bar{\mathbf{z}}^*$ via Algorithm 1.

2. Generate a new vector $\hat{\mathbf{z}}$ from $\bar{\mathbf{z}}^*$ by setting the K largest entries of $\bar{\mathbf{z}}^*$ to one and other entries to zero.

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 $\tilde{\mathbf{z}}$ to (3.9) from $\hat{\mathbf{z}}'$ by using the relation

$$\tilde{z}_i = \hat{z}'_i,\ i = 1, 2, \ldots, N_s \tag{3.22}$$

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 (3.9) (i.e., the solution of (3.9) must be a binary vector). However, setting the K largest entries of $\bar{\mathbf{z}}^*$ to one may lead to a violation of the total cost constraint ($\sum_{i=1}^{N_s} \hat{z}_i c_i > C_T$). Therefore, a local optimization algorithm is applied for both generating a feasible solution and improving the objective value achieved by feasible solutions. The local optimization algorithm starts with $\hat{\mathbf{z}}$, which is obtained in Step 2, as described above. If $\hat{\mathbf{z}}$ violates the total cost constraint, the algorithm checks for a swap between selected and unselected sensors which makes the new selection feasible. On the other hand, if $\hat{\mathbf{z}}$ is feasible, then the algorithm seeks to improve the objective value again via swaps. It terminates when no swaps can improve the objective value. Finally, the proposed suboptimal solution $\tilde{\mathbf{z}}$ is constructed from the first N_s entries of $\hat{\mathbf{z}}'$, which is obtained in Step 3. The pseudo-code of the local optimization algorithm is provided in Algorithm 3. The terms 'objective value of S' or 'cost of S' denote the objective value or the cost corresponding to the solution in which the selected sensors are the ones that have their indices in set S.

Algorithm 3 Local Optimization

```
get S_1                                   ▷ set of selected sensors
get S_0                                   ▷ set of unselected sensors
thisCost = cost of S_1
if thisCost > C_T then
    feasibility = 0
else
    feasibility = 1
end if
thisObjVal = objective value of S_1
top:
for i = 1 to K do
    for j = 1 to N_s − K do
        S'_1 ← S_1,    S'_0 ← S_0
        Δcost = cost of jth element in S_0 −
        cost of ith element in S_1
        if thisCost + Δcost ≤ C_T  then
            exchange ith element of S'_1 with
            jth element of S'_0
            newObjVal = objective value of S'_1
            if feasibility = 0 then
                S_1 ← S'_1,    S_0 ← S'_0
                feasibilty = 1
                thisCost = thisCost + Δcost
                thisObjVal = newObjVal
                goto top
            else
                if newObjVal > thisObjVal then
                    S_1 ← S'_1,    S_0 ← S'_0
                    thisCost = thisCost + Δcost
                    thisObjVal = newObjVal
                    goto top.
                end if
            end if
        end if
    end for
end for
```

Chapter 4

Joint Sensor Selection and Design for Binary Hypothesis Testing

In Chapter 3, the sensor selection problem is investigated under a cost constraint in order to minimize the Bayes risk for a given binary hypothesis testing problem by considering fixed measurement noise variances for the sensors; i.e., fixed values of $\sigma^2_{m_1}, \sigma^2_{m_2}, \ldots, \sigma^2_{m_{N_s}}$. In this chapter, 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 $\boldsymbol{\sigma^2_m}$ denote the vector of measurement noise variances, defined as

$$\boldsymbol{\sigma^2_m} \triangleq \left[\sigma^2_{m_1}, \sigma^2_{m_2}, \ldots, \sigma^2_{m_{N_s}}\right]^T . \quad (4.1)$$

Since the aim is to optimize the selection vector $\mathbf{z}$ and $\boldsymbol{\sigma^2_m}$ jointly, we extend the sensor selection problem in (3.9) (also see (3.3) and (3.10)) as follows:

$$
\begin{aligned}
\underset{\mathbf{z},\boldsymbol{\sigma}_{\boldsymbol{m}}^{\mathbf{2}}}{\text{maximize}} \quad & \sum_{i=1}^{N_s} z_i \frac{\left(\mathbf{h}_i^T(\boldsymbol{\Theta}_1-\boldsymbol{\Theta}_0)\right)^2}{\sigma_{n_i}^2+\sigma_{m_i}^2} \\
\text{subject to} \quad & 0.5\sum_{i=1}^{N_s} z_i \log_2\left(1+\frac{\sigma_{n_i}^2}{\sigma_{m_i}^2}\right) \le C_T \\
& \sum_{i=1}^{N_s} z_i \le K \\
& z_i \in \{0,1\}, \quad i=1,2,\ldots,N_s
\end{aligned} \tag{4.2}
$$

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

Before investigating the solution of the optimization problem in (4.2), we first consider the problem for a fixed $\mathbf{z}$ and present the following optimization problem over $\boldsymbol{\sigma}_{\boldsymbol{m}}^{\mathbf{2}}$ (called the measurement noise variance design problem):

$$
\begin{aligned}
\underset{\boldsymbol{\sigma}_{\boldsymbol{m}}^{\mathbf{2}}}{\text{maximize}} \quad & \sum_{i\in Z_1} \frac{\left(\mathbf{h}_i^T(\boldsymbol{\Theta}_1-\boldsymbol{\Theta}_0)\right)^2}{\sigma_{n_i}^2+\sigma_{m_i}^2} \\
\text{subject to} \quad & 0.5\sum_{i\in Z_1} \log_2\left(1+\frac{\sigma_{n_i}^2}{\sigma_{m_i}^2}\right) \le C_T \\
& \sigma_{m_i}^2 = \infty, \quad i \in Z_0
\end{aligned} \tag{4.3}
$$

where sets Z_0 and Z_1 are defined as

$$
Z_0 = \{i \in \{1,2,\ldots,N_s\} \mid z_i = 0\}\,, \tag{4.4}
$$

$$
Z_1 = \{i \in \{1,2,\ldots,N_s\} \mid z_i = 1\}\,. \tag{4.5}
$$

The problem in (4.3) is analyzed in [2]. It is shown that since a convex function is maximized over a convex set, the solution of (4.3) lies at the boundary. Namely, the solution can be obtained by an iterative algorithm that can be outlined as in Algorithm 4, where the following definitions are used for the simplicity of the expressions:

$$\mu_i^2 \triangleq \left(\mathbf{h}_i^T(\boldsymbol{\Theta}_1 - \boldsymbol{\Theta}_0)\right)^2, \tag{4.6}$$

$$\boldsymbol{\mu^2} \triangleq \left[\mu_1^2, \mu_2^2, \ldots, \mu_{N_s}^2\right]^T, \tag{4.7}$$

$$\boldsymbol{\sigma_n^2} \triangleq \left[\sigma_{n_1}^2, \sigma_{n_2}^2, \ldots, \sigma_{n_{N_s}}^2\right]^T. \tag{4.8}$$

Algorithm 4 Optimal Variance Design [2]

get $\mathbf{z}, \boldsymbol{\mu^2}, \boldsymbol{\sigma_n^2}$
$Z_1 = \{i \in \{1, \ldots, N_s\} \mid z_i = 1\}$
while $Z_1 \neq \emptyset$ **do**
  $\alpha = \left(2^{2C_T} \prod_{i \in Z_1} \frac{\sigma_{n_i}^2}{\mu_i^2}\right)^{\frac{1}{|Z_1|}}$
  $S_{inf} = \left\{i \in \{1, \ldots, N_s\} \mid (i \in Z_1) \,\&\, \left(\sigma_{n_i}^2 \geq \alpha \mu_i^2\right)\right\}$
  if $S_{inf} \neq \emptyset$ **then**
    $i = \arg\min_{j \in S_{inf}} \sigma_{n_j}^2$
    $Z_1 \leftarrow Z_1 \setminus \{i\}$
  else
    break
  end if
end while
$\sigma_{m_i}^2 = \begin{cases} \frac{\sigma_{n_i}^4}{\mu_i^2 \alpha - \sigma_{n_i}^2}, & i \in Z_1 \\ \infty, & \text{else} \end{cases} \qquad i = 1, 2, \ldots, N_s$

Algorithm 3 will be useful for obtaining the solution of the joint optimization problem in (4.2), as discussed in the following.

Based on (3.3), the measurement noise variance of the ith 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 (4.2) can equivalently be expressed as

$$
\begin{aligned}
\underset{\mathbf{z},\mathbf{c}}{\text{maximize}} \quad & \sum_{i=1}^{N_s} z_i \frac{\mu_i^2 \left(2^{2c_i}-1\right)}{\sigma_{n_i}^2 2^{2c_i}} \\
\text{subject to} \quad & \mathbf{z}^T\mathbf{c} \le C_T \\
& \sum_{i=1}^{N_s} z_i \le K \\
& z_i \in \{0,1\}, \quad i = 1,2,\ldots,N_s \\
& c_i \ge 0, \quad i = 1,2,\ldots,N_s
\end{aligned} \tag{4.9}
$$

where

$$
\mathbf{c} \triangleq [c_1, c_2, \ldots, c_{N_s}] \ . \tag{4.10}
$$

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

The solution of (4.9) 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 $i = 1, 2, \ldots, N_s$ (break ties arbitrarily). Then, a solution to the joint optimization problem in (4.9) is $(\mathbf{z}^*, \mathbf{c}^*)$, where the elements of $\mathbf{z}^*$ are given by*

$$
z_i^* = \begin{cases} 1\,, & i \in \tilde{B} \\ 0\,, & else \end{cases} \ , \quad i = 1, 2, \ldots, N_s \tag{4.11}
$$

and $\mathbf{c}^*$ is an optimizer of the problem in (4.9) when $\mathbf{z}$ is fixed as $\mathbf{z} = \mathbf{z}^*$. (Namely, $\mathbf{c}^*$ can be obtained via Algorithm 4 and (3.3) by setting $\mathbf{z} = \mathbf{z}^*$ in (4.9).)

Proof: Please see Appendix A.5.

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 $i = 1, 2, \ldots, N_s$. 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 chapter leads to a simpler sensor selection solution than the sensor selection problem considered in Chapter 3 for sensors with fixed measurement noise variances. In addition, it is noted that the solution of (4.2) 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.

Chapter 5

Numerical Examples

In this chapter, we provide examples for both the sensor selection problem in Chapter 3 and the joint sensor selection and design problem in Chapter 4. All the examples are carried out in the same simulation setting, which is described as follows. We consider the linear system in Fig. 3.1 with $N_s = 100$ potential sensor measurements. Parameter vector $\boldsymbol{\Theta}$ is a vector of length 20, which is equal to $\boldsymbol{\Theta_0}$ under hypothesis H_0 and equal to $\boldsymbol{\Theta_1}$ under hypothesis H_1. The entries of $\boldsymbol{\Theta_0}$ and $\boldsymbol{\Theta_1}$ 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 20×100 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, $\boldsymbol{\sigma_n^2}$, and the measurement noise variance vector, $\boldsymbol{\sigma_m^2}$, also come from a uniform distribution in the interval $[0.05, 1]$.

In order to present statistically meaningful results, we obtain 10000 realizations for the described random variables $\boldsymbol{\Theta_0}$, $\boldsymbol{\Theta_1}$, $\mathbf{H}$, $\boldsymbol{\sigma_n^2}$, and $\boldsymbol{\sigma_m^2}$. For each realization, we solve the corresponding optimization problem with the described methods and obtain the values of the objective function. We then average out the objective values for different realizations to provide the final results. Note that for the joint sensor selection and design problem, the realization of $\boldsymbol{\sigma_m^2}$ is irrelevant since it is considered as an optimization variable, and determined via the solution method.

5.1 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 (3.9). We investigate the performance of the proposed suboptimal solution to (3.9) in Chapter 3. We also present two different sensor selection strategies for comparison purposes, which are described as follows:

- **Simple Selection Strategy**: In this strategy, the sensors are sorted in a descending order according to their p_i values (please see the definition in (3.10)). Then, the top K sensors are selected. If the total cost of the selected sensors exceeds the cost budget, the most expensive selected sensor is exchanged with the cheapest unselected sensor. This procedure is repeated until the budget constraint is satisfied.

- **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 Figures 5.1 to 5.7, 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 (3.11), which is the relaxed version (3.9). Hence, the curves labeled as ‘Relaxed’ provide performance bounds in the considered scenarios.

In Figures 5.1 to 5.4, the performance of the considered strategies is presented versus the normalized total cost parameter (C_T divided by the cost of cheapest

K sensors) for 4 different values of K. For the performance metric, the objective value in (3.9) achieved by each strategy is employed, which corresponds to d^2, with d being given by (3.8). From Figures 5.1 to 5.4, it is observed that the performance of all the strategies improves as the total cost budget 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 and achieves very close performance to the performance bound ('Relaxed') especially for high values of C_T. The selection with only local optimization strategy outperforms the simple selection strategy, which has the worst performance. Although the gap between the performance bound and the selection with only local optimization strategy is significant for all values of the cost budget in case of low K values, it becomes quite small for higher values of K in the region of high cost budgets.

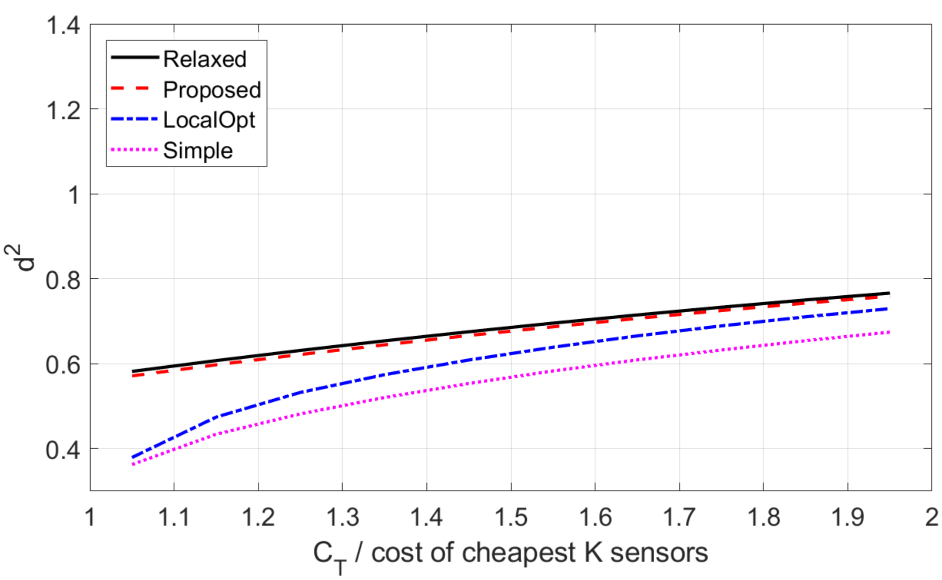

Figure 5.1: Performance of different strategies versus normalized cost, together with the performance bound obtained from the relaxed problem in (3.11), $K = 20$.

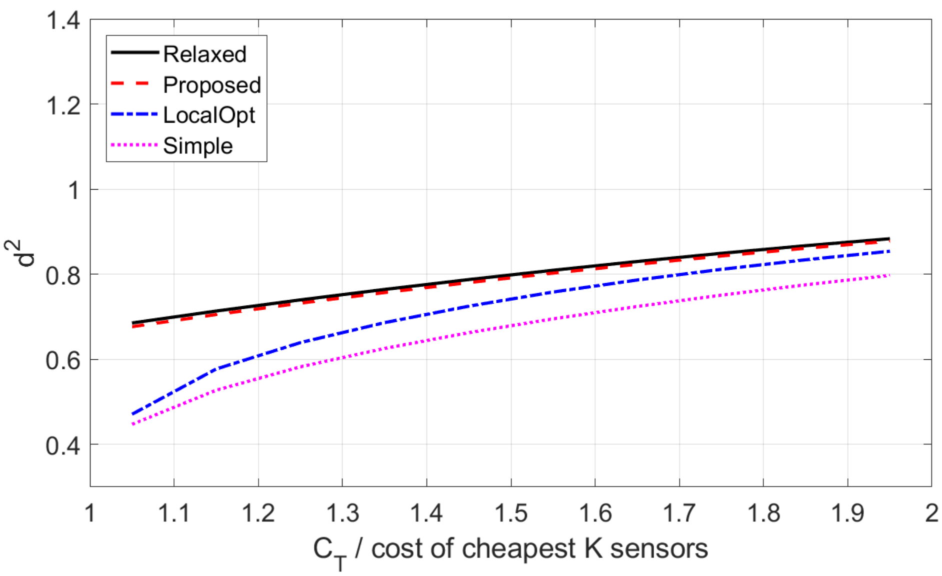

Figure 5.2: Performance of different strategies versus normalized cost, together with the performance bound obtained from the relaxed problem in (3.11), $K = 25$.

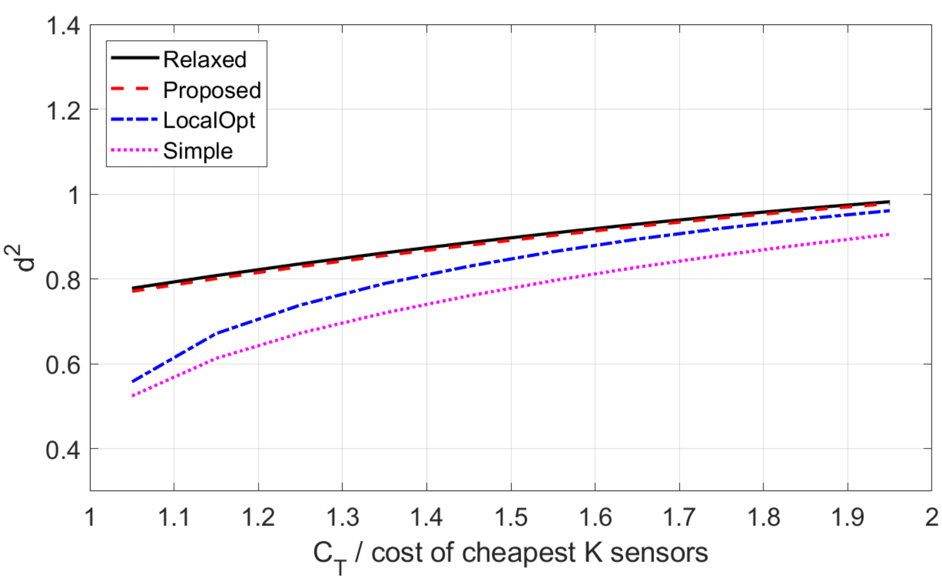

Figure 5.3: Performance of different strategies versus normalized cost, together with the performance bound obtained from the relaxed problem in (3.11), $K = 30$.

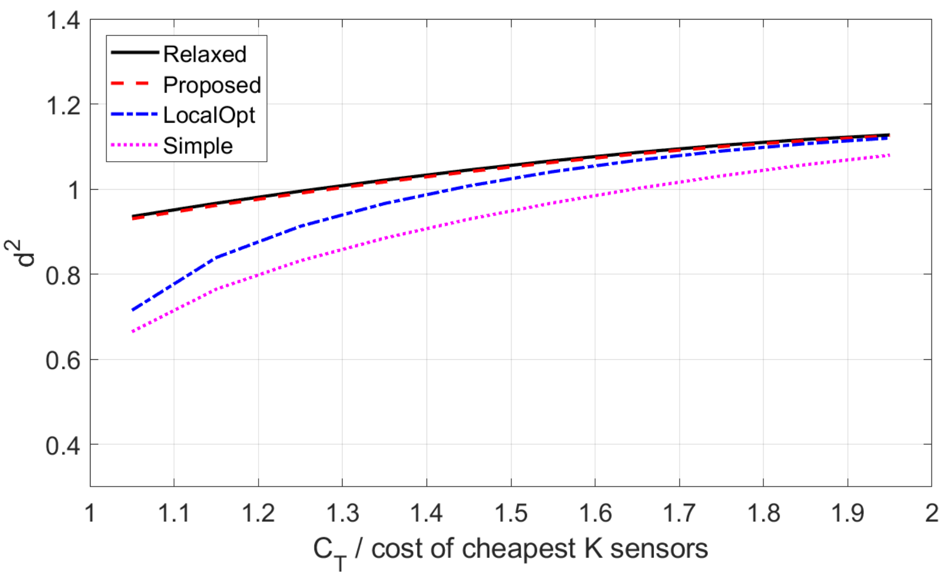

Figure 5.4: Performance of different strategies versus normalized cost, together with the performance bound obtained from the relaxed problem in (3.11), $K = 40$.

In Figures 5.5 to 5.7, the performance of the strategies are plotted versus K for two different cost budgets. Similar conclusions to those in Figures 5.1 to 5.4 are made. Namely, the proposed strategy achieves the best performance, which is close to the upper bound. It is also noted that that as K increases, the gap between the proposed strategy and the performance bound decreases, which is a desirable property. A similar argument holds for the gap between the performance bound and the selection with only local optimization strategy when C_T is equal to 1.45 or 1.85 times the cost of the cheapest K sensors. However, when C_T is equal to 1.05 times the cost of the cheapest K sensors, the corresponding gap does not reduce with K.

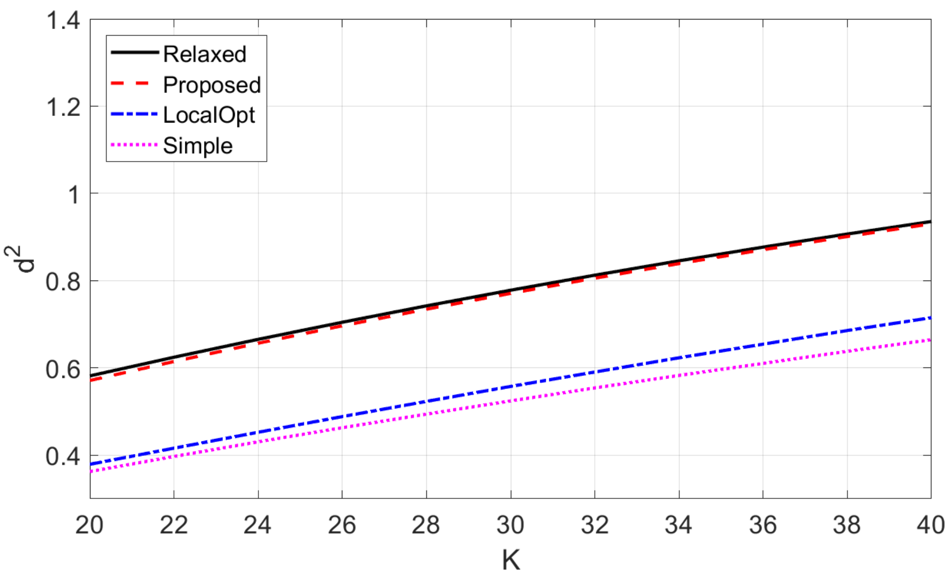

Figure 5.5: Performance of different strategies versus K, together with the performance bound obtained from the relaxed problem in (3.11), $C_T = 1.05$ times cost of cheapest K sensors.

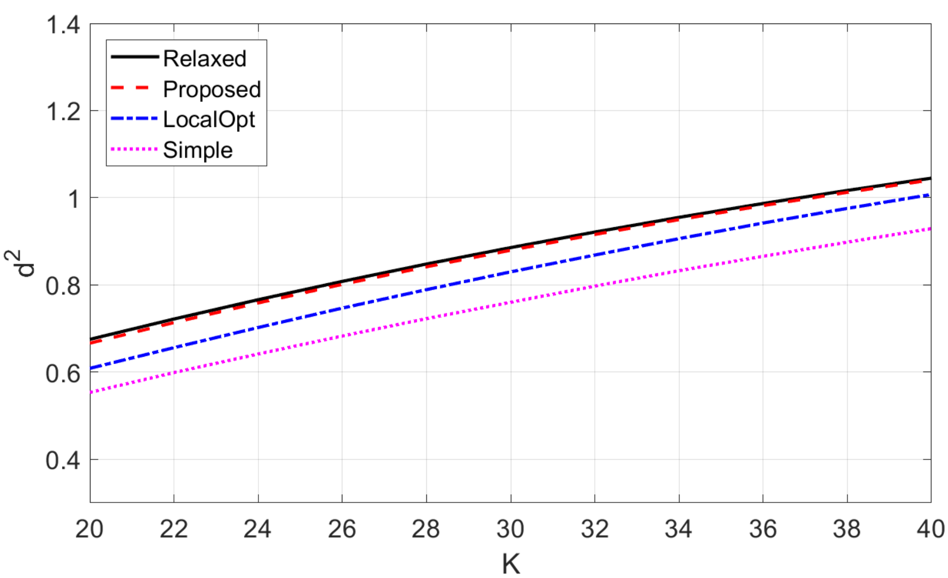

Figure 5.6: Performance of different strategies versus K, together with the performance bound obtained from the relaxed problem in (3.11), $C_T = 1.45$ times cost of cheapest K sensors.

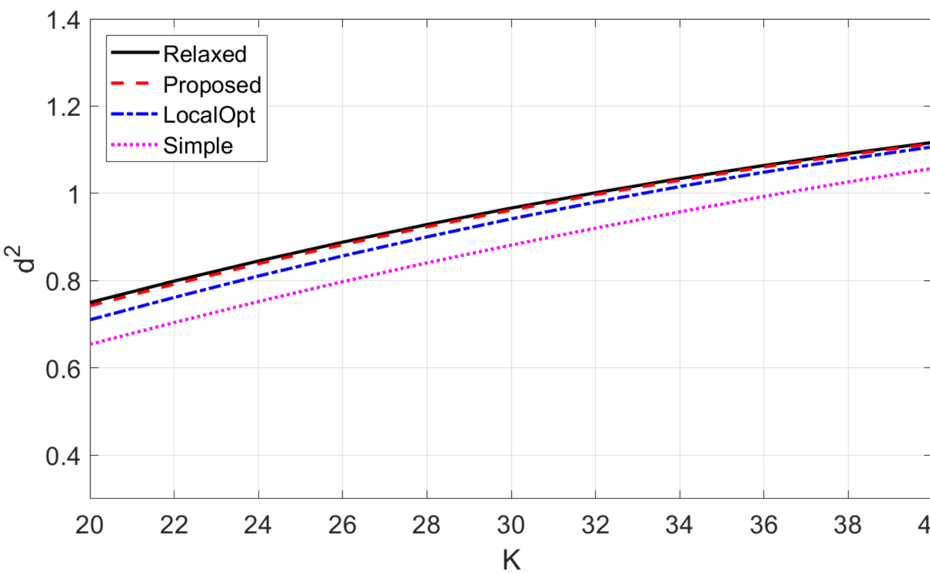

Figure 5.7: Performance of different strategies versus K, together with the performance bound obtained from the relaxed problem in (3.11), $C_T = 1.85$ times cost of cheapest K sensors.

5.2 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 (4.2). To obtain the proposed optimal solution to (4.2), we utilize the approach described in Proposition 2. In addition to the proposed optimal solution, we also present results for two suboptimal 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

$$\sigma_{m_j}^2 = \frac{\sigma_{n_j}^2}{2^{2C_T/K} - 1} \,. \tag{5.1}$$

- **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

$$\sigma_{m_j}^2 = \frac{\sigma_{n_j}^2}{2^{2C_T} - 1} \,. \tag{5.2}$$

In Figures 5.8 to 5.10, the performance of the proposed strategy (labeled as 'Optimal') is evaluated by plotting d^2 against the total cost budget C_T for various values of K. 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 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 Figures 5.8 to 5.10 that the detection performance achieved by allocating all the cost to the best sensor quickly reaches a constant level as C_T increases since only one sensor is utilized all the time. 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. From the three plots in Figures 5.8 to 5.10, it is also noted that as K increases, the optimal strategy offers more significant benefits with respect to the other strategies. Moreover, the convergence of the strategy that allocates equal costs to the best K sensors to the optimal strategy occurs at higher values of C_T as K increases (which is related to the fact that a larger value of K leads to the distribution of the cost budget among more sensors).

In Figures 5.11 to 5.13, the detection performance of the considered strategies is plotted with respect to K for some fixed cost budgets; namely, $C_T = 1$, $C_T = 5$, and $C_T = 10$. 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. (This is possible since zero cost can be assigned to a sensor in the problem formulation in (4.9).) In addition, the value of K after which the optimal strategy has constant detection performance increases as the cost budget 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, Figures 5.8 to 5.13 illustrate the advantages of the proposed optimal strategy in various scenarios.

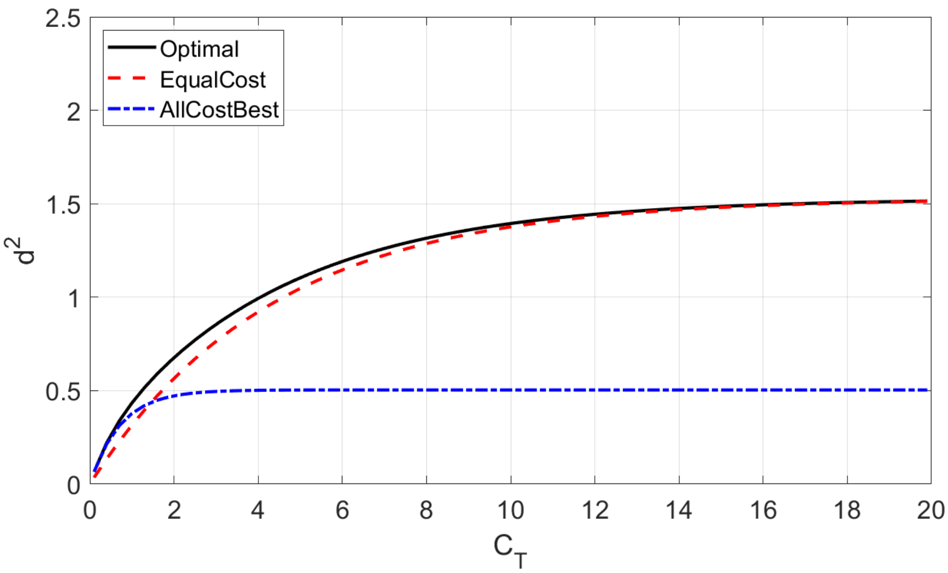

Figure 5.8: Performance of different strategies versus C_T, $K = 6$.

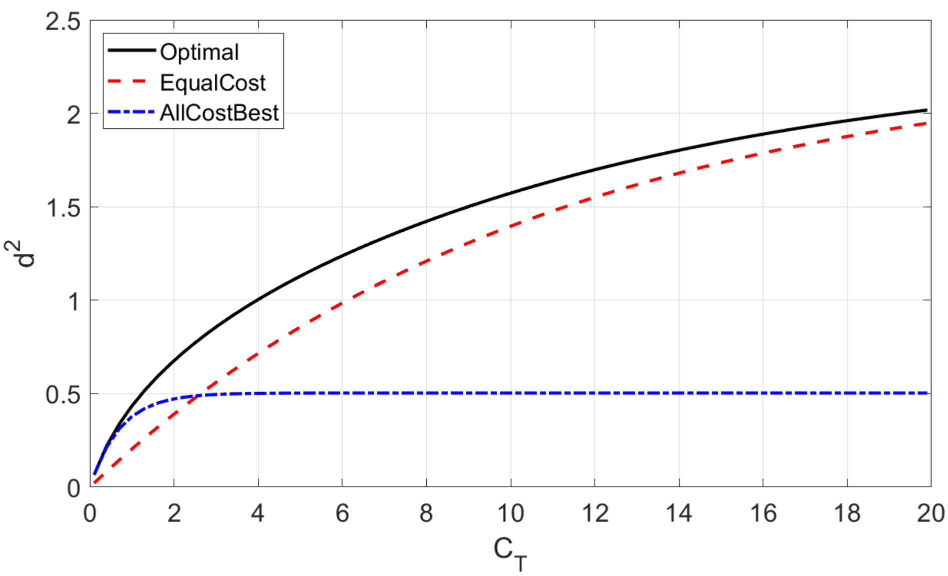

Figure 5.9: Performance of different strategies versus C_T, $K = 15$.

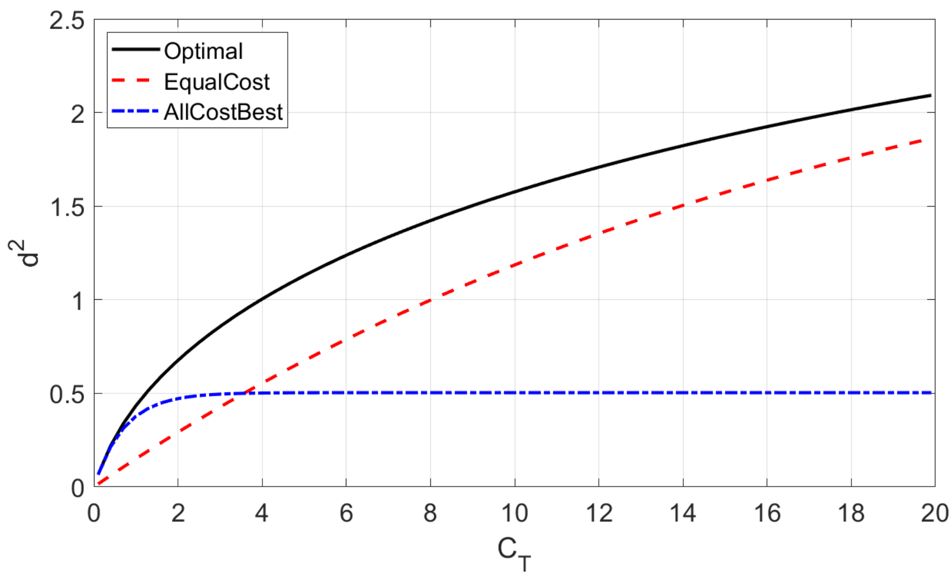

Figure 5.10: Performance of different strategies versus C_T, $K = 25$.

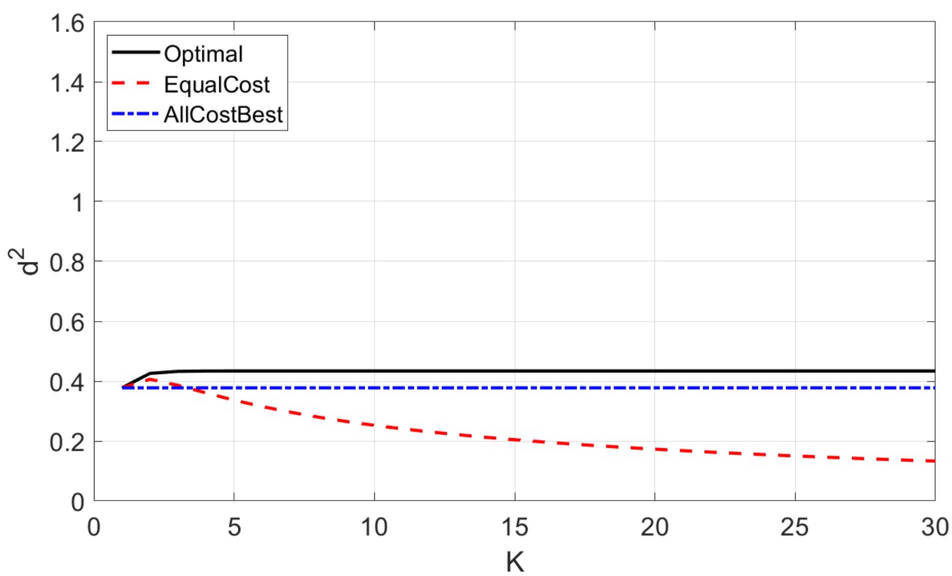

Figure 5.11: Performance of different strategies versus K, $C_T = 1$.

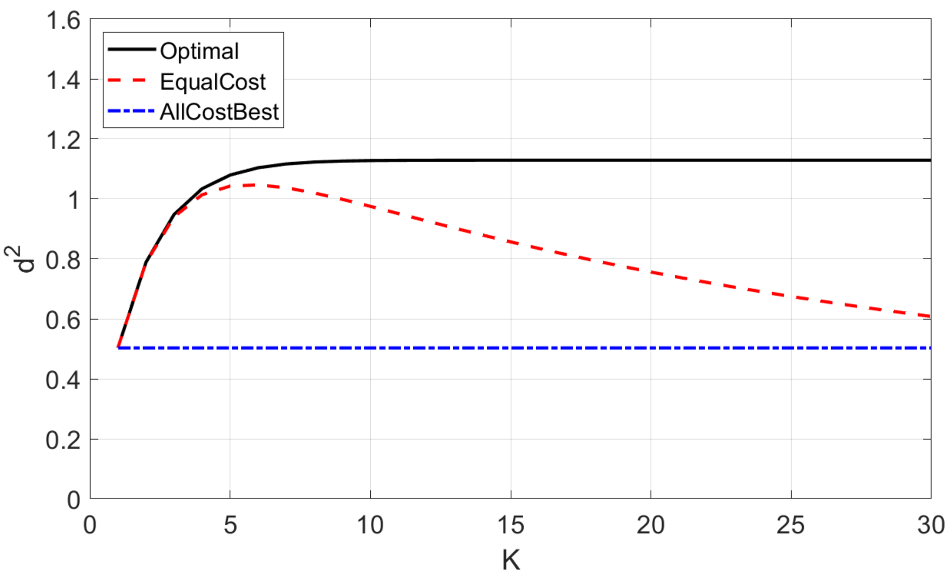

Figure 5.12: Performance of different strategies versus K, $C_T = 5$.

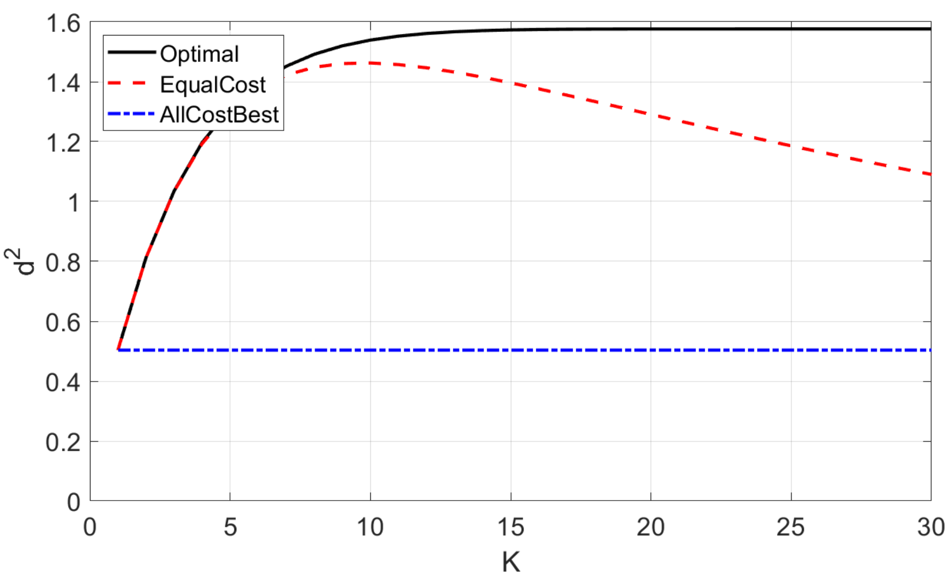

Figure 5.13: Performance of different strategies versus K, $C_T = 10$.

Chapter 6

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 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. We have also observed that when the cost constraint is strict, utilizing only local optimization produces solutions that are quite close to those obtained via linear relaxation followed by local optimization.

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.

Bibliography

[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] B. Dulek and S. Gezici, "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, pp. 314–322, IEEE, 2013.

[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*, pp. 405–416, Springer, 2003.

[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*, pp. 36–45, 2004.

[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*, p. 20, IEEE Press, 2005.

[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, p. 65621A, 2007.

[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*, pp. 1–11, IEEE, 2008.

[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*, pp. 17–33, Springer, 2008.

[33] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. MIT Press, 2009.

[34] A. Schrijver, *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.

[35] E. L. Lawler and D. E. Wood, "Branch-and-bound methods: A survey," *Operations Research*, vol. 14, no. 4, pp. 699–719, 1966.

[36] A. Ozcelikkale, H. M. Ozaktas, and E. Arikan, "Signal recovery with cost-constrained measurements," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3607–3617, 2010.

[37] F. Bian, D. Kempe, and R. Govindan, "Utility based sensor selection," in *Proceedings of the 5th International Conference on Information Processing in Sensor Networks*, pp. 11–18, 2006.

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

[39] G. B. Dantzig, *Linear Programming and Extensions*. Princeton University Press, 1998.

Appendix A

Proofs of Lemmas and Propositions

A.1 Proof of Lemma 1

Consider the optimization problem in (3.11) in the absence of the cost constraint. Then, it is easy to verify that $\mathbf{z}^*$ defined in the lemma is a solution to (3.11) 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 (3.11) 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 (3.9).

$\square$

A.2 Proof of Lemma 2

Assume that $\mathbf{z}'$ is a solution to (3.11) with the objective value $v' = \sum_{i=1}^{N_s} z'_i p_i$, where $\sum_{i=1}^{N_s} z'_i c_i < C_T$. Then, $\mathbf{z}'$ satisfies one of the following conditions:

a) $\sum_{i=1}^{N_s} z'_i = K$

b) $\sum_{i=1}^{N_s} z'_i < K$ and

$\sum_{i=1}^{N_s} z'_i p_i < \sum_{i \in B_j} p_i, \ \forall j \in \{1, 2, \ldots, N_L\}$

c) $\sum_{i=1}^{N_s} z'_i < K$ and

$\exists j \in \{1, 2, \ldots, N_L\}$ such that $\sum_{i=1}^{N_s} z'_i p_i = \sum_{i \in B_j} p_i$

Suppose (a) holds. Then, we prove by contradiction that $\mathbf{z}'$ is not a solution to (3.11). It can be shown that, for any B_j with $j \in \{1, \ldots, N_L\}$, there exist indices a and b such that $a \in B_j$, $b \notin B_j$, $z'_a < 1$, $z'_b > 0$, $p_a > p_b$, and $c_a > c_b$ since $\sum_{i \in B_j} c_i > C_T$ and $\sum_{i=1}^{N_s} z'_i = 1$. (This can also be proved by contradiction, which is not included for brevity.) Then, $\mathbf{z}''$ can be constructed from $\mathbf{z}'$ by choosing a sufficiently small positive ϵ as follows:

$$z''_i = \begin{cases} z'_i \,, & i \neq a, b \\ z'_i + \epsilon \,, & i = a \\ z'_i - \epsilon \,, & i = b \end{cases} \quad . \tag{A.1}$$

The objective value v'' achieved by $\mathbf{z}''$ satisfies the following relation:

$$v'' = \sum_{i=1}^{N_s} z''_i p_i = \sum_{i=1}^{N_s} z'_i p_i + \epsilon(p_a - p_b) > \sum_{i=1}^{N_s} z'_i p_i = v' \tag{A.2}$$

In addition, by choosing ϵ as

$$0 < \epsilon \leq \min \left\{ \frac{C_T - \sum_{i=1}^{N_s} z'_i c_i}{c_a - c_b} \,, 1 - z'_a \,, z'_b \right\} , \tag{A.3}$$

it is guaranteed that $\mathbf{z}''$ in (A.1) satisfies the constraints in (3.11); i.e.,

$$\sum_{i=1}^{N_s} z''_i c_i = \sum_{i=1}^{N_s} z'_i c_i + \epsilon(c_b - c_a) \leq C_T \tag{A.4}$$

$$0 \leq z''_i \leq 1 \,, \ \ i = 1, \ldots, N_s \,. \tag{A.5}$$

Since $\mathbf{z}''$ achieves a higher objective value than $\mathbf{z}'$ and also satisfies the constraints in (3.11), it is deduced that $\mathbf{z}'$ cannot be a solution to (3.11). That is, in case (a), a selection vector $\mathbf{z}'$ with $\sum_{i=1}^{N_s} z_i' c_i < C_T$ cannot be a solution.

Suppose (b) holds. Similarly to (a), we prove that a higher objective value can be attained by constructing a feasible selection vector $\mathbf{z}''$. For any B_j with $j \in \{1, \ldots, N_L\}$, there exists an index a such that $z_a' < 1$, $p_a > 0$ and $a \in B_j$. Then, consider

$$z_i'' = \begin{cases} z_i' \,, & i \neq a \\ z_i' + \epsilon \,, & i = a \end{cases} . \tag{A.6}$$

where ϵ is sufficiently small; i.e.,

$$0 < \epsilon \leq \min \left\{ \frac{C_T - \sum_{i=1}^{N_s} z_i' c_i}{c_a} \,, 1 - z_a' \,, K - \sum_{i=1}^{N_s} z_i' \right\} . \tag{A.7}$$

The resulting objective value associated with $\mathbf{z}''$ satisfies

$$v'' = \sum_{i=1}^{N_s} z_i'' p_i = \sum_{i=1}^{N_s} z_i' p_i + \epsilon(p_a) > \sum_{i=1}^{N_s} z_i' p_i = v' \tag{A.8}$$

$\mathbf{z}''$ is feasible since

$$\sum_{i=1}^{N_s} z_i'' c_i = \sum_{i=1}^{N_s} z_i' c_i + \epsilon(c_a) \leq C_T \tag{A.9}$$

$$\sum_{i=1}^{N_s} z_i'' = \sum_{i=1}^{N_s} z_i' + \epsilon \leq K \tag{A.10}$$

$$0 \leq z_i'' \leq 1 \,, \;\; i = 1, \ldots, N_s \,. \tag{A.11}$$

Since $\mathbf{z}''$ achieves a higher objective value, $\mathbf{z}'$ is not a solution. That is, in case (b), a selection vector $\mathbf{z}'$ with $\sum_{i=1}^{N_s} z_i' c_i < C_T$ cannot be a solution.

Suppose (c) holds. Then, $\mathbf{z}'$ is a solution to (3.11). In this case, we argue that there exists another solution $\mathbf{z}''$ which satisfies $\sum_{i=1}^{N_s} z_i'' c_i = C_T$. To that aim, define $S \triangleq \{a : a \in B_j, p_a > 0\}$. Since $\sum_{i=1}^{N_s} z_i' p_i = \sum_{i \in B_j} p_i$, it implies that $z_i' = 1 \; \forall i \in S$. Then, using the inequality $\sum_{i=1}^{N_s} z_i' < K$, it is inferred that

$|S| < K$ and $p_i = 0\ \forall i \in B_j \backslash S$. Then, we have

$$\sum_{i \in B_j} c_i = \sum_{i \in S} c_i + \sum_{i \in B_j \backslash S} c_i > C_T > \sum_{i=1}^{N_s} z_i' c_i \geq \sum_{i \in S} c_i \tag{A.12}$$

Rearranging the terms, we get

$$\sum_{i \in B_j \backslash S} c_i > C_T - \sum_{i \in S} c_i > 0 \tag{A.13}$$

Evidently, we can find $\{w_i\}_{i \in B_j \backslash S}$ such that

$$\sum_{i \in B_j \backslash S} w_i c_i = C_T - \sum_{i \in S} c_i \tag{A.14}$$

where $0 \leq w_i \leq 1$, $i \in B_j \backslash S$. Then, we construct a new solution $\mathbf{z}''$ as follows:

$$z_i'' = \begin{cases} 1\,, & i \in S \\ w_i\,, & i \in B_j \backslash S\,. \\ 0\,, & \text{else} \end{cases} \tag{A.15}$$

$\mathbf{z}''$ achieves the same objective value as $\mathbf{z}'$ as shown below:

$$\begin{aligned} v'' &= \sum_{i=1}^{N_s} z_i'' p_i = \sum_{i \in S} p_i + \sum_{i \in B_j \backslash S} w_i p_i \\ &= \sum_{i \in S} p_i + \sum_{i \in B_j \backslash S} p_i = \sum_{i \in B_j} p_i = \sum_{i=1}^{N_s} z_i' p_i = v' \end{aligned} \tag{A.16}$$

Also, $\mathbf{z}''$ is feasible by (A.14) as noted in the following:

$$\sum_{i=1}^{N_s} z_i'' c_i = \sum_{i \in S} c_i + \sum_{i \in B_j \backslash S} w_i c_i = C_T \tag{A.17}$$

$$\sum_{i=1}^{N_s} z_i'' = |S| + \sum_{i \in B_j \backslash S} w_i < |S| + |B_j \backslash S| = K \tag{A.18}$$

Therefore, $\mathbf{z}''$ is a solution that satisfies $\sum_{i=1}^{N_s} z_i'' c_i = C_T$. In other words, in case (c), for any solution $\mathbf{z}'$ with $\sum_{i=1}^{N_s} z_i' c_i < C_T$, there exists an alternative solution $\mathbf{z}''$ with $\sum_{i=1}^{N_s} z_i'' c_i = C_T$.

Overall, it is shown that when $C_T < \sum_{i \in B_j} c_i$ for all $j \in \{1, \ldots, N_L\}$, either a solution must satisfy (3.13) (in cases (a) and (b)), or there exists a solution satisfying (3.13) (in case (c)).

□

A.3 Proof of Proposition 1

Let $B_1, B_2, \ldots, B_{N_L}$ 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 (3.11) can be expressed as

$$z_i^* = \begin{cases} 0\,, & i \notin B_j \\ 1\,, & i \in B_j \end{cases} \tag{A.19}$$

which conforms to the characterization in (3.14) and (3.15).

Consider the case of $C_T < \sum_{i \in B_j} c_i$ for all $j \in \{1, 2, \ldots, N_L\}$. In this case, there exists a solution $\mathbf{z}^*$ to (3.11) that satisfies $C_T = \sum_{i=1}^{N_s} z_i^* c_i$ by Lemma 2. $\mathbf{z}^*$ should satisfy the following Karush-Kuhn-Tucker (KKT) conditions with an equality constraint for the total cost:

$$\sum_{i=1}^{N_s} z_i^* - K \leq 0 \tag{A.20}$$

$$\sum_{i=1}^{N_s} z_i^* c_i - C_T = 0 \tag{A.21}$$

$$0 \leq z_i^* \leq 1,\ i = 1, 2, \ldots, N_s \tag{A.22}$$

$$\nu \left(\sum_{i=1}^{N_s} z_i^* - K \right) = 0 \tag{A.23}$$

$$\lambda_i \geq 0,\ i = 1, 2, \ldots, 2N_s \tag{A.24}$$

$$\lambda_i z_i^* = 0,\ i = 1, 2, \ldots, N_s \tag{A.25}$$

$$\lambda_{N_s+i}(z_i^* - 1) = 0,\ i = 1, 2, \ldots, N_s \tag{A.26}$$

$$-p_i - \lambda_i + \lambda_{N_s+i} + \mu c_i + \nu = 0,\ i = 1, 2, \ldots, N_s \tag{A.27}$$

where $\lambda_1, \lambda_2, \dots, \lambda_{2N_s}$, μ, and ν are the KKT multipliers. From (A.23)–(A.27), it is observed that if $z_i^* \in (0,1)$ we get

$$\begin{aligned} \sum_{i=1}^{N_s} z_i^* = K &\implies p_i = \mu c_i + \nu \\ \sum_{i=1}^{N_s} z_i^* < K &\implies p_i = \mu c_i \end{aligned} \tag{A.28}$$

Suppose that there exists a solution $\mathbf{z}'$ to (3.11) ($\sum_{i=1}^{N_s} z_i' c_i = C_T$), where $\sum_{i=1}^{N_s} z_i' = K$ and $\mathbf{z}'$ does not satisfy the property in (3.14) and (3.15), 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 (3.11) that satisfies (3.14) and (3.15). Define sets of indices S_0', S_1' and S_2' as

$$\begin{aligned} S_0' &\triangleq \{i : z_i' = 0,\ i = 1, 2, \dots, N_s\} \\ S_1' &\triangleq \{i : z_i' = 1,\ i = 1, 2, \dots, N_s\} \\ S_2' &\triangleq \{i : z_i' = (0,1)\,,\ i = 1, 2, \dots, N_s\} \end{aligned} \tag{A.29}$$

and let $N \triangleq |S_1'|$. Then, we have

$$|S_2'| = M > 2 \tag{A.30}$$

$$|S_0'| = N_s - M - N \tag{A.31}$$

$$0 \leq N < K < N + M \leq N_s \tag{A.32}$$

$$\sum_{i \in S_2'} z_i' = K - N. \tag{A.33}$$

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

$$\sum_{i \in C_{S_2'}} c_i \leq \sum_{i \in S_2'} z_i' c_i \leq \sum_{i \in E_{S_2'}} c_i\,. \tag{A.34}$$

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

$$\begin{aligned} X_{t+1} &= (X_t \setminus \{m_t\}) \cup \{n_t\} \\ m_t &= \arg\min_i \, c_i, \ i \in X_t \\ n_t &= \arg\max_i \, c_i, \ i \in S'_2 \setminus X_t \end{aligned} \tag{A.35}$$

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:

$$\sum_{i \in X_T} c_i \leq \sum_{i \in S'_2} z'_i c_i \leq \sum_{i \in X_{T+1}} c_i \,. \tag{A.36}$$

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

$$\sum_{i \in X_T \setminus \{m_T\}} c_i + (1-\alpha) c_{m_T} + \alpha c_{n_T} = \sum_{i \in S'_2} z'_i c_i \,. \tag{A.37}$$

In particular,

$$\alpha = \begin{cases} 0\,, & c_{n_T} = c_{m_T} \\ \frac{\sum_{i \in S'_2} z'_i c_i - \sum_{i \in X_T} c_i}{c_{n_T} - c_{m_T}}\,, & c_{n_T} > c_{m_T} \end{cases} . \tag{A.38}$$

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

$$z^*_i = \begin{cases} 0\,, & i \in S'_0 \cup S'_{20} \\ 1\,, & i \in S'_1 \cup S'_{21} \\ \alpha\,, & i = n_T \\ 1-\alpha\,, & i = m_T \end{cases} . \tag{A.39}$$

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

$$S'_2 = S'_{20} \cup S'_{21} \cup \{m_T\} \cup \{n_T\} \tag{A.40}$$

Also it is noted that

$$|S'_{21}| = K - N - 1\,. \tag{A.41}$$

In the following, it is shown that $\mathbf{z}^*$ in (A.39) satisfies the condition in (A.21); i.e., the cost constraint.

$$\sum_{i=1}^{N_s} z_i^* c_i = \sum_{i \in S'_1 \cup S'_{21}} c_i + \alpha\, c_{n_T} + (1-\alpha) c_{m_T} \tag{A.42}$$

$$= \sum_{i \in S'_1} c_i + \sum_{i \in X_T \setminus \{m_T\}} c_i + \alpha\, c_{n_T} + (1-\alpha) c_{m_T} \tag{A.43}$$

$$= \sum_{i \in S'_1} c_i + \sum_{i \in S'_2} z'_i c_i \tag{A.44}$$

$$= \sum_{i=1}^{N_s} z'_i c_i \tag{A.45}$$

$$= C_T \tag{A.46}$$

where (A.42) follows from (A.39), (A.43) is due to the definition of S'_{21}, (A.44) is based on (A.37), (A.45) follows from definitions in (A.29) and finally (A.46) is due to Lemma 2.

To prove that $\mathbf{z}^*$ is a solution that satisfies the property in (3.14) and achieves the same objective value as $\mathbf{z}'$, consider following equalities:

$$v' = \sum_{i=1}^{N_s} z'_i p_i \tag{A.47}$$

$$= \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z'_i p_i \tag{A.48}$$

$$= \sum_{i \in S'_1} p_i + \mu \sum_{i \in S'_2} z'_i c_i + \nu \sum_{i \in S'_2} z'_i \tag{A.49}$$

$$= \sum_{i \in S'_1} p_i + \mu \left(\sum_{i \in X_T \setminus \{m_T\}} c_i + (1-\alpha) c_{m_T} + \alpha c_{n_T} \right)$$
$$+ \nu \,(K - N) \tag{A.50}$$

$$= \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z_i^* \,(\mu c_i + \nu) \tag{A.51}$$

$$= \sum_{i \in S'_1} p_i + \sum_{i \in S'_2} z_i^* p_i \tag{A.52}$$

$$= \sum_{i=1}^{N_s} z_i^* p_i = v^* \tag{A.53}$$

where (A.48) follows from the definitions in (A.29), (A.49) is due to (A.28), (A.50) is based on (A.29) and (A.37), (A.51) follows from (A.39)–(A.41), (A.52) is due to (A.28), and finally (A.53) is based on (A.39) and (A.40).

Based on the preceding arguments, it is proved that $\mathbf{z}^*$ is also a solution to (3.11). From (A.39), it is noted that $\mathbf{z}^*$ satisfies (3.14), where

$$\begin{aligned} S_0 &= S'_0 \cup S'_{20} \\ S_1 &= S'_1 \cup S'_{21} \\ S_2 &= \{n_T, m_T\} \end{aligned} \tag{A.54}$$

Suppose there exists a solution $\mathbf{z}'$ to (3.11) ($\sum_{i=1}^{N_s} z'_i c_i = C_T$) such that $\sum_{i=1}^{N_s} z'_i < K$, where $\mathbf{z}'$ does not satisfy (3.16); i.e., $\mathbf{z}'$ has $M > 1$ non-integer components. We argue that there exists another solution satisfying (3.16). Define sets of indices S'_0, S'_1 and S'_2 as in (A.29). Then, the following relation holds:

$$0 \leq \sum_{i \in S'_2} z'_i c_i < \sum_{i \in S'_2} c_i \tag{A.55}$$

Let $X_0 = \emptyset$ and

$$\begin{aligned} X_{t+1} &= X_t \cup \{n_t\} \\ n_t &= \arg\max_i \; c_i, \;\; i \in S'_2 \setminus X_t \end{aligned} \tag{A.56}$$

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

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

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

$$\sum_{i \in X_T} c_i + \alpha c_{n_T} = \sum_{i \in S'_2} z'_i c_i \tag{A.58}$$

where

$$\alpha = \frac{\sum_{i \in S'_2} z'_i c_i - \sum_{i \in X_T} c_i}{c_{n_T}}. \tag{A.59}$$

Since $\sum_{i\in S_2'} z_i' c_i \le \sum_{i\in X_T} c_i + \left(\sum_{i\in S_2'} z_i' - T\right) c_{n_T}$, we have

$$\begin{aligned}
\alpha &\le \frac{\left(\sum_{i\in X_T} c_i + \left(\sum_{i\in S_2'} z_i' - T\right) c_{n_T}\right) - \sum_{i\in X_T} c_i}{c_{n_T}} \\
&= \frac{\left(\sum_{i\in S_2'} z_i' - T\right) c_{n_T}}{c_{n_T}} \qquad \text{(A.60)} \\
&= \sum_{i\in S_2'} z_i' - T
\end{aligned}$$

Let

$$S_{21}' = X_T, \quad S_{20}' = S_2' \setminus (X_T \cup \{n_T\}), \qquad \text{(A.61)}$$

split S_2' into four disjoint sets as

$$S_2' = S_{20}' \cup S_{21}' \cup \{n_T\} \qquad \text{(A.62)}$$

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

$$z_i^* = \begin{cases} 0\,, & i \in S_0' \cup S_{20}' \\ 1\,, & i \in S_1' \cup S_{21}' \\ \alpha\,, & i = n_T \end{cases} . \qquad \text{(A.63)}$$

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

$$\begin{aligned}
\sum_{i=1}^{N_s} z_i^* &= |S_1'| + |S_{21}'| + \alpha && \text{(A.64)} \\
&\le |S_1'| + |X_T| + \sum_{i\in S_2'} z_i' - T && \text{(A.65)} \\
&= |S_1'| + \sum_{i\in S_2'} z_i' && \text{(A.66)} \\
&= \sum_{i=1}^{N_s} z_i' \le K && \text{(A.67)}
\end{aligned}$$

Here, (A.64) follows from (A.63), (A.65) is due to (A.60), (A.66) follows from (A.56), and finally (A.67) is based on (A.29).

The total cost constraint is also satisfied since

$$\sum_{i=1}^{N_S} z_i^* c_i = \sum_{i \in S_1' \cup S_{21}'} c_i + \alpha c_{n_T} \tag{A.68}$$

$$= \sum_{i \in S_1'} c_i + \sum_{i \in X_T} c_i + \alpha c_{n_T} \tag{A.69}$$

$$= \sum_{i \in S_1'} c_i + \sum_{i \in S_2'} z_i' c_i \tag{A.70}$$

$$= \sum_{i=1}^{N_S} z_i' c_i = C_T \tag{A.71}$$

where (A.68) follows from the definition of $\hat{\mathbf{z}}$ in (A.63), (A.69) is based in (A.29), (A.70) is due to (A.58), and (A.71) follows from the definition of $\mathbf{z}'$.

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

$$v' = \sum_{i=1}^{N_S} z_i' p_i = \sum_{i \in S_1'} p_i + \sum_{i \in S_2'} z_i' p_i \tag{A.72}$$

$$= \sum_{i \in S_1'} p_i + \mu \sum_{i \in S_2'} z_i' c_i \tag{A.73}$$

$$= \sum_{i \in S_1'} p_i + \mu \left(\sum_{i \in X_T} c_i + \alpha c_{n_T} \right) \tag{A.74}$$

$$= \sum_{i \in S_1'} p_i + \sum_{i \in S_2'} z_i^* \left(\mu c_i \right) \tag{A.75}$$

$$= \sum_{i \in S_1'} p_i + \sum_{i \in S_2'} z_i^* p_i \tag{A.76}$$

$$= \sum_{i=1}^{N_S} z_i^* p_i = v^* \tag{A.77}$$

where (A.72) is based on (A.29), (A.73) and (A.76) follow from (A.28), (A.74) is due to (A.58), and (A.75) is based on (A.63) and (A.62).

Overall, it is shown that if there exists a solution to (3.11) that does not satisfy neither *a)* (i.e., (3.14) and (3.15)) nor *b)* (i.e., (3.16) and (3.17)), then there always exists an alternative solution which satisfies either *a)* or *b)* in the proposition. Thus, when the relaxed problem in (3.11) is feasible, there exists a solution satisfying either *a)* or *b)*, as claimed in the proposition.

□

A.4 Conversion of Inequality Constraint in (3.11) into Equality Constraint

Consider the relaxed optimization problem in (3.11). A slack variable s is added to convert the inequality constraint on the number of selected sensors into an equality constraint as follows:

$$\begin{aligned}
\underset{\mathbf{z},s}{\text{maximize}} \quad & \sum_{i=1}^{N_s} z_i p_i \\
\text{subject to} \quad & \sum_{i=1}^{N_s} z_i c_i \leq C_T \\
& \sum_{i=1}^{N_s} z_i + s = K \\
& 0 \leq z_i \leq 1, \quad i = 1, 2, \ldots, N_s \\
& s \geq 0
\end{aligned} \tag{A.78}$$

Then, we can substitute the slack variable s with a column vector $\mathbf{s}$ of length K, and obtain the following equivalent problem based on the definitions in (3.20):

$$\begin{aligned}
\underset{\mathbf{z},\mathbf{s}}{\text{maximize}} \quad & \sum_{i=1}^{N_s} z_i \bar{p}_i + \sum_{i=1}^{K} s_i \bar{p}_{N_s+i} \\
\text{subject to} \quad & \sum_{i=1}^{N_s} z_i \bar{c}_i + \sum_{i=1}^{K} s_i \bar{c}_{N_s+i} \leq C_T \\
& \sum_{i=1}^{N_s} z_i + \sum_{i=1}^{K} s_i = K \\
& 0 \leq z_i \leq 1, \quad i = 1, 2, \ldots, N_s \\
& \sum_{i=1}^{K} s_i \geq 0
\end{aligned} \tag{A.79}$$

The last constraint can simply be replaced with

$$0 \leq s_i \leq 1, \quad i = 1, \ldots, K. \tag{A.80}$$

Finally, letting $\bar{\mathbf{z}} \triangleq [\mathbf{z}^T,\ \mathbf{s}^T]^T$, we obtain the problem in (3.19) Hence, a solution to (3.19) can be used to obtain a solution to (3.11) by dropping the indices that correspond to the slack variables.

A.5 Proof of Proposition 2

To simplify notation, we denote the objective function in (4.9) as $F(\mathbf{z}, \mathbf{c})$, that is,

$$F(\mathbf{z}, \mathbf{c}) \triangleq \sum_{i=1}^{N_s} z_i \frac{\mu_i^2 \left(2^{2c_i} - 1\right)}{\sigma_{n_i}^2 2^{2c_i}} . \tag{A.81}$$

Let $(\mathbf{z}, \mathbf{c})$ be any feasible solution to the problem in (4.9). We argue that there exists $\mathbf{c}' = \left[c_1', c_2', \ldots, c_{N_s}'\right]^T$ such that

$$F(\mathbf{z}^*, \mathbf{c}') \geq F(\mathbf{z}, \mathbf{c}) \tag{A.82}$$

and $(\mathbf{z}^*, \mathbf{c}')$ is feasible, where $\mathbf{z}^*$ is as defined in (4.11). To justify this claim, we define $\mathbf{c}'$ in the following way: First, consider any bijection g from $\{i \,|\, z_i^* = 1, z_i = 0\}$ to $\{j \,|\, z_j^* = 0, z_j = 1\}$. Then, let

$$c_i' = \begin{cases} c_i\,, & z_i^* = 1,\ z_i = 1 \\ c_{g(i)}\,, & z_i^* = 1,\ z_i = 0 \\ 0\,, & \text{else} \end{cases} , \quad i = 1, 2, \ldots, N_s\,. \tag{A.83}$$

It is noted that $c_i' \geq 0$ for $i = 1, 2, \ldots, N_s$. Also,

$$(\mathbf{z}^*)^T \mathbf{c}' = \sum_{i \in \{k \,|\, z_k^* = 1,\, z_k = 1\}} c_i + \sum_{i \in \{k \,|\, z_k^* = 1,\, z_k = 0\}} c_{g(i)} \tag{A.84}$$

$$= \sum_{i \in \{k \,|\, z_k^* = 1,\, z_k = 1\}} c_i + \sum_{i \in \{k \,|\, z_k^* = 0,\, z_k = 1\}} c_i \tag{A.85}$$

$$= \mathbf{z}^T \mathbf{c} \leq C_T\,. \tag{A.86}$$

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:

$$
\begin{aligned}
F(\mathbf{z}^*, \mathbf{c}') &= \sum_{i \in \{k \,|\, z_k^*=1, z_k=1\}} \frac{\mu_i^2 \left(2^{2c_i}-1\right)}{\sigma_{n_i}^2 2^{2c_i}} \\
&\quad + \sum_{i \in \{k \,|\, z_k^*=1, z_k=0\}} \frac{\mu_i^2 \left(2^{2c_{g(i)}}-1\right)}{\sigma_{n_i}^2 2^{2c_{g(i)}}} &\text{(A.87)}\\
&= \sum_{i \in \{k \,|\, z_k=1\}} \frac{\mu_i^2 \left(2^{2c_i}-1\right)}{\sigma_{n_i}^2 2^{2c_i}} \\
&\quad - \sum_{i \in \{k \,|\, z_k^*=0, z_k=1\}} \frac{\mu_i^2 \left(2^{2c_i}-1\right)}{\sigma_{n_i}^2 2^{2c_i}} \\
&\quad + \sum_{i \in \{k \,|\, z_k^*=1, z_k=0\}} \frac{\mu_i^2 \left(2^{2c_{g(i)}}-1\right)}{\sigma_{n_i}^2 2^{2c_{g(i)}}} &\text{(A.88)}\\
&= F(\mathbf{z}, \mathbf{c}) - \sum_{i \in \{k \,|\, z_k^*=0, z_k=1\}} \frac{\mu_i^2 \left(2^{2c_i}-1\right)}{\sigma_{n_i}^2 2^{2c_i}} \\
&\quad + \sum_{i \in \{k \,|\, z_k^*=1, z_k=0\}} \frac{\mu_i^2 \left(2^{2c_{g(i)}}-1\right)}{\sigma_{n_i}^2 2^{2c_{g(i)}}} &\text{(A.89)}\\
&= F(\mathbf{z}, \mathbf{c}) - \sum_{i \in \{k \,|\, z_k^*=1, z_k=0\}} \frac{\mu_{g(i)}^2 \left(2^{2c_{g(i)}}-1\right)}{\sigma_{n_{g(i)}}^2 2^{2c_{g(i)}}} \\
&\quad + \sum_{i \in \{k \,|\, z_k^*=1, z_k=0\}} \frac{\mu_i^2 \left(2^{2c_{g(i)}}-1\right)}{\sigma_{n_i}^2 2^{2c_{g(i)}}} &\text{(A.90)}\\
&= F(\mathbf{z}, \mathbf{c}) + \sum_{i \in \{k \,|\, z_k^*=1, z_k=0\}} \left(\frac{2^{2c_{g(i)}}-1}{2^{2c_{g(i)}}}\right)\left(\frac{\mu_i^2}{\sigma_{n_i}^2} - \frac{\mu_{g(i)}^2}{\sigma_{n_{g(i)}}^2}\right) &\text{(A.91)}\\
&\geq F(\mathbf{z}, \mathbf{c}) &\text{(A.92)}
\end{aligned}
$$

By definition, $\frac{\mu_i^2}{\sigma_{n_i}^2} \geq \frac{\mu_{g(i)}^2}{\sigma_{n_{g(i)}}^2}$ for $i \in \{k \,|\, z_k^* = 1, z_k = 0\}$ and $g(i) = j \in \{k \,|\, z_k^* = 0, z_k = 1\}$ 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 (A.91) is larger than or equal to zero. Hence, the inequality in (A.92) follows. Finally, we have

$$F(\mathbf{z}^*, \mathbf{c}^*) \geq F(\mathbf{z}^*, \mathbf{c}') \geq F(\mathbf{z}, \mathbf{c}) \tag{A.93}$$

since $\mathbf{c}^*$ is a solution of (4.3) 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.

□