Decentralized Hypothesis Testing
in Sensor Networks

ALLA TARIGHATI

Doctoral Thesis in Electrical Engineering
Stockholm, Sweden 2016

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen i elektro- och systemteknik onsdagen den 30 november 2016 klockan 10.00 i sal F3, Lindstedtsvägen 26, Stockholm.

To my beloved family!

Abstract

Wireless sensor networks (WSNs) play an important role in the future of Internet of Things IoT systems, in which an entire physical infrastructure will be coupled with communication and information technologies. Smart grids, smart homes, and intelligent transportation systems are examples of infrastructure that will be connected with sensors for intelligent monitoring and management. Thus, sensing, information gathering, and efficient processing at the sensors are essential.

An important problem in wireless sensor networks is that of decentralized detection. In a decentralized detection network, spatially separated sensors make observations on the same phenomenon and send information about the state of the phenomenon towards a central processor. The central processor (or the fusion center, FC) makes a decision about the state of the phenomenon, base on the aggregate received messages from the sensors. In the context of decentralized detection, the object is often to make the *best* decision at the FC. Since this decision is made based on the received messages from the sensors, it is of interest to *optimally* design decision rules at the remote sensors.

This dissertation deals mainly with the problem of designing decision rules at the remote sensors and at the FC, while the network is subject to some limitation on the communication between nodes (sensors and the FC). The contributions of this dissertation can be divided into three (overlapping) parts. First, we consider the case where the network is subject to communication rate constraint on the links connecting different nodes. Concretely, we propose an algorithm for the design of decision rules at the sensors and the FC in an arbitrary network in a person-by-person (PBP) methodology. We first introduce a network of two sensors, labeled as the *restricted model*. We then prove that the design of sensors' decision rules, in the PBP methodology, is in an arbitrary network equivalent to designing the sensors' decision rules in the corresponding restricted model. We also propose an efficient algorithm for the design of the sensors' decision rules in the restricted model.

Second, we consider the case where remote sensors share a common multiple access channel (MAC) to send their messages towards the FC, and where the MAC channel is subject to a sum rate constraint. In this situation,

the sensors compete for communication rate to send their messages. We find sufficient conditions under which allocating equal rate to the sensors, so called *rate balancing*, is an optimal strategy. We study the structure of the optimal rate allocation in terms of the Chernoff information and the Bhattacharyya distance.

Third, we consider a decentralized detection network where not only are the links between nodes subject to some communication constraints, but the sensors are also subject to some energy constraints. In particular, we study the network under the assumption that the sensors are energy harvesting devices that acquire all the energy they need to transmit their messages from their surrounding environment. We formulate a decentralized detection problem with system costs due to the random behavior of the energy available at the sensors in terms of the Bhattacharyya distance.

Sammanfattning

Trådlösa sensornätverk spelar en viktig roll i framtiden av IoT (eng. Internet of Things) system, där en hel fysisk infrastruktur kommer att koppla med kommunikations- och informationsteknik. Smarta nät, smarta hem och intelligenta transportsystem är exempel på infrastruktur som ska anslutas med sensorer för intelligent övervakning och hantering. Således är avkänning, informationsinsamling och effektiv informationsbearbetning av sensordata viktigt.

Ett centralt problem i trådlösa sensornätverk är decentraliserad detektion. I ett decentraliserad detektionsnätverk, gör spatialt separerade sensorer observationer av samma fenomen och skickar information om tillståndet för fenomenet mot en central processor. Den centrala processorn (eller fusionscentrum, FC), tar ett beslut om tillståndet hos fenomenet, med användning av de sammanlagda mottagna meddelandena från sensorerna. Inom ramen för decentraliserad detektion är syftet främst att göra det bästa beslutet i FC. Eftersom detta beslut fattas i enlighet med de mottagna meddelanden från sensorerna, är det av intresse att optimalt utforma beslutsregler för fjärrsensorerna.

Denna avhandling behandlar främst med problemet att utforma beslutsregler på fjärrsensorer och FC, medan nätverket är föremål för någon begränsning av kommunikationen mellan noderna (sensorerna och FC). Det vetenskapliga bidraget i denna avhandling kan delas in i tre (överlappande) delar. Först, är fallet där nätverket är föremål för kommunikationshastighetbegränsningar på länkarna som förbinder olika noder. Konkret föreslår vi en algoritm för utformningen av beslutsregler vid sensorerna och FC i ett godtyckligt nätverk i en person-för-person (PBP) metodik. Vi introducerar ett förenklat nätverk av två sensorer, en så kallat *restricted model*. Vi bevisar att utformningen av sensorernas beslutsregler, under en PBP metodik, i ett godtyckligt nät motsvarar utforma sensorernas beslutsregler i motsvarande restricted model. Sedan så föreslår vi en effektiv algoritm för konstruktion av sensorernas beslutsregler i det förenklade nätverket.

Sedan betraktar vi scenariot där fjärrsensorer har en gemensam multiple access channel (MAC) för att skicka sina meddelanden till FC, och där MAC-kanalen är föremål för en begränsning av summan av de individuella bithastigheterna. I den situationen, konkurrerar sensorerna om

kommunikationshastigheten för att skicka sina meddelanden. Vi definierar tillräckliga villkor under vilka det är en optimal strategi att fördela lika hastighet till sensorerna. Vi studerar strukturen för en optimal fördelning av hastigheterna i termer av Chernoff-information och Bhattacharyya-avstånd.

Sist så betraktar vi ett decentraliserat detektionsnätverk där inte bara länkarna mellan noderna är föremål för kommunikationsbegränsningar, utan där även sensorerna är också föremål energibegränsningar. Mer precist så studerar vi nätverket under antagandet att sensorerna är energiinsamblande och förvärvar alla den energi de behöver för att sända sina meddelanden från sin omgivning. Vi formulerar det decentraliserade detektionsproblemet med systemkostnader som modellerar det slumpmässiga beteendet av energin i de individuella sensorerna i termer av Bhattacharyya-avståndet.

Nyckelord: Trådlösa sensornätverk, decentraliserad detektion, person-för-person optimering, flera åtkomstkanaler, fördelningshastighet, Chernoff-information, Bhattacharyya-avstånd, energiinsamblande.

List of Papers

The thesis is based on the following papers:

[A] A. Tarighati, and J. Jaldén, "Bayesian design of decentralized hypothesis testing under communication constraints", in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, May 2014, pp. 7624–7628.

[B] A. Tarighati, and J. Jaldén, "Bayesian design of tandem networks for distributed detection with multi-bit sensor decisions," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1821–1831, April 2015.

[C] A. Tarighati, and J. Jaldén, "A general method for the design of tree networks under communication constraints", in *Proc. 17th Int. Conf. Information Fusion (FUSION)*, July 2014, pp. 1–7.

[D] A. Tarighati, and J. Jaldén, "Rate allocation for decentralized detection in wireless sensor networks", in *Proc. 16th IEEE Int. Workshop Signal Process. Advances in Wireless Commun. (SPAWC)*, June 2015, pp. 341–345.

[E] A. Tarighati, and J. Jaldén, "Optimality of rate balancing in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3735–3749, July 2016.

[F] A. Tarighati, J. Gross, and J. Jaldén, "Decentralized detection in energy harvesting wireless sensor networks", in *Proc. European Signal Process. Conf. (EUSIPCO)*, Aug. 2016.

[G] A. Tarighati, J. Gross, and J. Jaldén, "Decentralized hypothesis testing in energy harvesting wireless sensor networks", *IEEE Trans. Signal Process.*, to be submitted, available: arXiv:1605.03316.

v

Acknowledgements

This dissertation could not have been finished without the help and support from many professors, colleagues, friends and my family. It is my great pleasure to acknowledge people who have given me guidance, help and encouragement.

First and foremost, I would like to thank my main advisor Assoc. Prof. Joakim Jaldén for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me throughout my research and writing of this dissertation. Without his breadth of knowledge in the area of signal processing, finishing my dissertation would have remained a distant dream.

I would like to thank my co-advisor Prof. Mats Bengtsson for giving me the opportunity to pursue my Ph.D studies at the department of Signal Processing, KTH. I would like to thank you also for your essential help in Latex.

I would like to express my sincere appreciation to DeWiNe research members, specially Assoc. Prof. Tobias Oechtering who engaged in valuable discussion with me during the first years of my research on distributed detection, and also acted as the internal advanced reviewer of this dissertation. I would like to thank my co-author Assoc. Prof. James Gross for helping me to transform one of my research interests into an article, and sharing those joyful moments watching football games together.

I am always indebted to my undergraduate advisor Assoc. Prof. Kamran Kazemi. His sincere dedication to his students was both admirable and inspirational. My interest in research began when I started working with him. I feel honored to have had a chance to work with him.

It has been my pleasure to work with colleagues in the departments of signal processing and communications theory, KTH. I would like to thank Majid Nasiri, Majid Gerami, M. Reza Gholami, Hamed Farhadi, Amirpasha Shirazinia, Nafiseh Shariati, Shahab Nazari, Farshad Naghibi, Serveh Shalmashi, Ehsan Olfat, Hossein Shokri, Arash Owrang, and Nima Najari Moghadam for making my lunch/fika breaks such an amazing times. I would like to thank Marie Maros, Arun Venkitaraman, Klas Magnusson, Rasmus Brandt, Pol Del Aguila Pla, Ehsan Olfat, Arash Owrang, Martin Sundin, and Nima Najari Moghadam (monthly-fika-group members) for having great

Contents

Acronyms

WSN	Wireless sensor network
IoT	Internet of Things
i.i.d.	Independent and identically distributed
MAP	Maximum a-posteriori
SNR	Signal-to-noise ratio
PBP	Person-by-person
FC	Fusion center
DM	Decision maker
LRQ	Likelihood-ratio quantizer
ROC	Receiver operating characteristic
OOK	On-off keying
MAC	Multiple access channel
BAC	Binary asymmetric channel
LLR	Log-likelihood ratio
LRT	Likelihood-ratio test

Part I

Summary

1 Introduction

1.1 Decentralized detection in wireless sensor networks

Wireless sensor networks have gained worldwide attention over the past decade [1–3]. Small sensors with limited processing, computing resources and communication capabilities are used in wireless sensor networks. These sensors are able to sense, measure, and gather information from their surrounding environment. The sensors are also able, based on local decision processes, to transmit the sensed data [4–6].

Statistical inferences about the environment such as detection, parameter estimation and tracking, are some of the emerging applications of wireless sensor networks. In all these applications we are faced with a decision-making problem where information is distributed across the sensors in the network. In all these group decision-making problems we need to select a particular course of action based on our observations regarding a certain phenomenon. If the course of action is from a set of possible options, our problem is a decentralized detection problem [7–11].

In a decentralized detection system, spatially separated sensors observe a common phenomenon. If the sensors are able to communicate all their data to a central processor for data processing, the sensors only act as data collectors and no processing at the sensors is needed. In this situation, the data processing is centralized in nature and optimal algorithms can be implemented. On the other hand, if there are communication constraints on the sensors, some preliminary data processing at each sensor needs to be carried out and a compressed or quantized version of the received data is then given as the sensor output. According to the network arrangement, the output of each sensor is sent either to another sensor or to a fusion center (FC) where the received information is appropriately combined to yield the final inference [12–14].

Decentralized detection problems are found in many practical situations. As an example, consider a radar detection system, where spatially separated radars are observing an area. According to their observations, they send some data towards a central processor or FC. Based on the received data from the radars, the FC makes a decision regarding the presence or absence of a target in that area. As another example, consider a digital communication system in which one of several possible waveforms is transmitted and the goal is to determine the transmitted symbol according to the noisy observations at the sensors.

In the context of decentralized hypothesis testing, each sensor is an intelligent unit, and is therefore often referred to as a decision maker (DM). In what follows, we shall discuss decentralized hypothesis testing problems in different topologies.
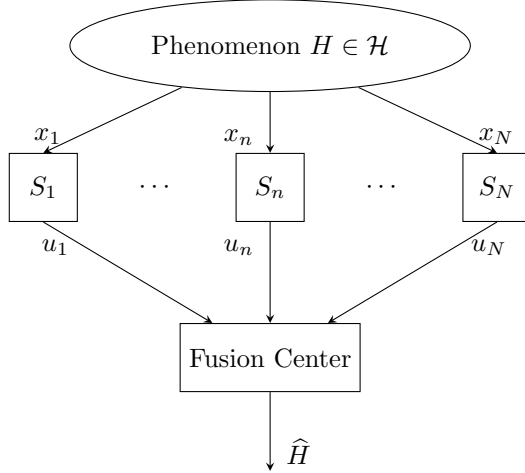
Figure 1: Decentralized hypothesis testing scheme in a parallel network.

1.2 Decentralized hypothesis testing problems

In a decentralized detection problem, we may be faced with a binary problem in which we are looking for a *yes* or *no* type of answer. For example, in the radar detection system described previously (see Section 1.1), our goal is to determine if the target is present (*yes*) or not (*no*). These problems are known as binary hypothesis testing problems and the two hypotheses are usually denoted by H_0 and H_1. We may also be faced with a more general situation with multiple hypotheses. An example of this problem is the digital communication system also described previously (see Section 1.1), where the set of possible waveforms consists of $M > 2$ different symbols. In the M-ary hypothesis testing problems, we denote the hypotheses by[1] $H_0, H_1, \ldots, H_{M-1}$.

The design of decentralized hypothesis testing algorithms depends on the underlying sensor network topology (or configuration), i.e., the arrangement of the sensors in a network [4]. The most common topology is the parallel topology, which is extensively considered in the literature (see [9,15–20] and references therein) and is shown in Figure 1. In this configuration, a sensor $S_n, n = 1, \ldots, N$, makes a private observation, denoted by x_n, of the common phenomenon H and sends a message u_n towards the FC. The FC after combining all the messages $u_1, \ldots, u_N$ received from the sensors $S_1, \ldots, S_N$

---

[1]The definition of a hypothesis set for M-ary problems for notational simplicity in some publications is as $\{H_1, \ldots, H_M\}$.
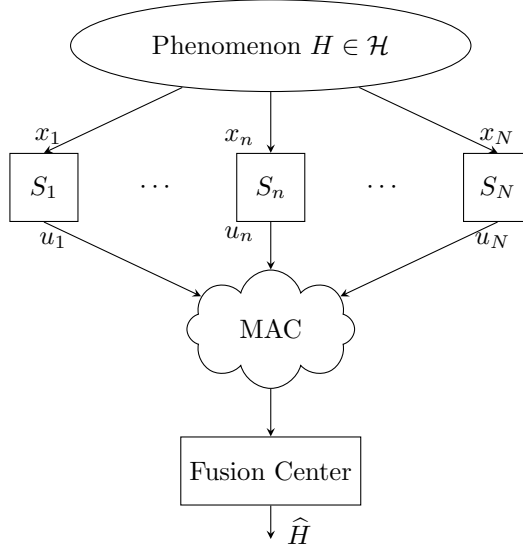
Figure 2: Decentralized hypothesis testing scheme in a network, where the sensors send their data through a MAC channel to the FC.

makes a global decision $\widehat{H}$ in favor of a hypothesis from hypothesis set $\mathcal{H}$. In this configuration there is no communication between the sensors and each sensor-to-FC channel is a one-way channel from the sensor.

Although a large body of research in decentralized detection has been devoted to the case where the sensors transmit their messages towards the FC through parallel access channels, in wireless sensor networks the wireless medium is typically shared among the sensors. In other words, the sensor-to-FC channels are modeled as common multiple access channels (MAC) [21–26]. This configuration is shown in Figure 2.

In both configurations shown in Figures 1 and 2, the sensors make private observations on the underlying phenomenon and transmit their output, known as local decisions, directly to the FC through wireless channels. In another popular structure, known as the tandem (or serial) topology, sensors are connected in series [27–34]. In this configuration, each sensor is connected to two neighboring sensors: its predecessor and its successor. Two exceptions to this rule are the first sensor, which does not have any predecessor, and the last sensor which does not have any successor. A network of sensors arranged in tandem is also shown in Figure 3. In this configuration, each sensor $S_n, n = 2, \ldots, N - 1$ makes an observation x_n
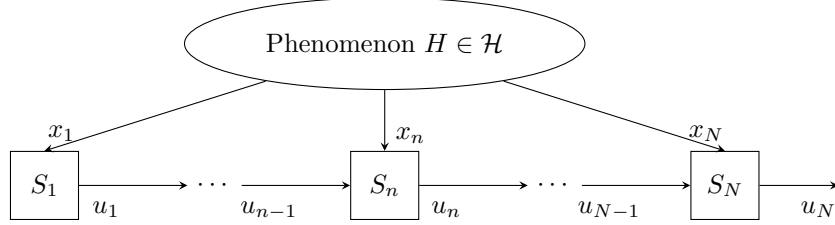
Figure 3: Decentralized hypothesis testing scheme in a tandem network.

and receives the output u_{n-1} of its predecessor as its inputs, and makes a decision u_n to send to its successor S_{n+1}. Sensor S_1 makes a decision u_1 only based on its private observation x_1, and the decision of the last sensor, S_N, is the final decision of the network.

Distributed detection in the tandem topology is closely connected to decision making in *social networks* [35–40]. In social networks, each agent chooses an action/decision by optimizing its local utility function and subsequent agents then choose their actions/decisions using their private observations together with the actions/decisions of previous agents [38, 41]. In social networks, the agents typically do not reveal their raw private observations on the underlying state of nature, while they do reveal their actions/decisions (e.g., votes and recommendations). This can be viewed as a low resolution (or in the context of decentralized detection as *quantized*) version of their observations and their interaction with other agents in the network.

In contrast to decentralized detection networks where the agents' decisions are made in such a way that a global performance measure is optimized, in social networks the agents' decisions are made in such a way that their local performance measure is optimized. In other words, in social networks agents *selfishly* try to optimize their outputs [42].

In addition to the extreme cases of parallel and tandem networks, there are other configurations such as tree topologies which are combinations of these cases [43–48]. An example of such a configuration is shown in Figure 4. In this situation, some of the nodes make observations while other nodes act as relays which combine the input messages and forward their output to another node. All the observed data eventually arrive at the FC, where the final decision about the present hypothesis is made.

In the following section, state of the art problems for each of the aforementioned configurations will be discussed. However, before discussing the problems, let us define the parameters that will be repeatedly used through-
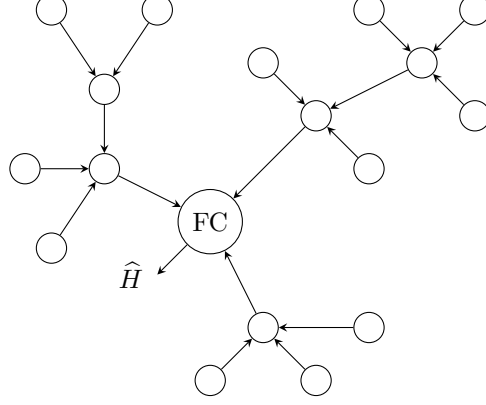
Figure 4: Decentralized hypothesis testing scheme in a tree network. Each sensor is shown by a circle and the observations are not shown.

out this thesis. Let X_n and U_n, for some $1 \leq n \leq N$, be random variables corresponding to observation x_n and output u_n of sensor S_n, respectively. The phenomenon H is also modeled as a random variable drawn from a set $\mathcal{H} \triangleq \{H_0, H_1, \ldots, H_{M-1}\}$ with corresponding a-priori probabilities $\pi_0, \pi_1, \ldots, \pi_{M-1}$ for a general M-ary hypothesis testing problem. We also define $f_{\underline{X}|H}(x_1, \ldots, x_N|h)$, for $h \in \mathcal{H}$, as the joint conditional distribution of the observations at sensors S_1 to S_N, where by definition $\underline{X} \triangleq X_1, \ldots, X_N$. When the observations at the sensors, conditioned on the hypothesis, are independent the joint conditional distribution decouples as

$$f_{\underline{X}|H}(x_1, \ldots, x_N|h) = f_{X_1|H}(x_1|h) \ldots f_{X_N|H}(x_N|h).$$

We also define $\mathcal{X}_n$ as the observation space of X_n, and $\mathcal{U}_{r_n}$ as the message space of output u_n.

1.3   State-of-the-art problems in distributed detection

For the parallel topology shown in Figure 1, two main problems need to be considered: the design of decision rules at the sensors, and the design of decision rules at the FC. To optimize a performance metric of the network, the decision rules at the FC and the local decision rules at the sensors should be designed jointly. Assuming perfect knowledge of the system parameters, the optimum design of the FC rule is conceptually a straightforward task [4, 9]. However, even for small-sized networks, the design of decision rules at the sensor nodes is a formidable task [49].

Given a realization x_n, sensor S_n in a parallel topology makes a decision u_n according to its own private observation x_n using a decision function $\gamma_n(\cdot)$, i.e.,

$$\gamma_n(x_n) = u_n, \tag{1}$$

and the FC makes the final decision $\widehat{h}$ in favor of a hypothesis from the set $\mathcal{H}$ by combining all the received messages from the sensors using a decision function $\gamma_{\mathrm{FC}}(\cdot)$, i.e.,

$$\gamma_{\mathrm{FC}}(u_1, \ldots, u_N) = \widehat{h}. \tag{2}$$

Note that, in this situation the sensor makes a decision only based on its own observation and not the observations at the other sensors, since there is no communication between the sensors. In what follows, we shall first discuss the structure of optimal sensors' decision rules and then formulate state-of-the-art problems in distributed detection networks.

If there is no communication constraint on the channel from the sensor to the FC, the sensor gives its own observation as its output, i.e., $u_n = x_n$. However, in practical situations the sensor-to-FC channel is subject to some communication constraints. Let us assume that this channel is an error-free but rate-constrained channel of rate r_n, for a positive integer r_n. In other words, this channel is capable of reliably carrying r_n bits to the FC. In this situation, each sensor S_n is a quantizer and its decision rule is a mapping from the observation space $\mathcal{X}_n$ to the message space $\mathcal{U}_{r_n} \triangleq \{1, \ldots, 2^{r_n}\}$, i.e., $\gamma_n : \mathcal{X}_n \to \mathcal{U}_{r_n}$. We are then led to the problem of finding quantization strategies at the sensors with respect to optimizing a performance measure. This problem was extensively considered in the quantization literature [50–55] and decentralized detection literature [9,56–58]. It is well known that for several specific observation distribution models, when observations at the sensors are independent given the hypothesis, likelihood-ratio quantizers (LRQ) are optimal. Therefore the problem of finding optimal decision rules at the sensors reduces to that of quantization thresholds [9, 15, 16, 59–61].

In the simple case of binary hypothesis testing where the sensors are arranged in parallel and each sensor sends a single bit towards the FC, the optimal decision rule at each sensor, as discussed above, is a likelihood-ratio test (LRT) as

$$\frac{f_{X_n|H}\left(x_n|H_1\right)}{f_{X_n|H}\left(x_n|H_0\right)} \underset{u_n=0}{\overset{u_n=1}{\gtrless}} \alpha_n, \tag{3}$$

where α_n is a scalar constant. If the likelihood-ratio is monotone in observation x_n, (3) admits the following form

$$x_n \underset{u_n=0}{\overset{u_n=1}{\gtrless}} T_n, \tag{4}$$

for a constant T_n which depends on α_n and an observation model at the sensor. In other words, in this case, without loss of generality we can assume

that a quantizer is directly applied to the observations, rather than to the likelihood-ratio (see [58, 62] for more detailed discussion). Furthermore, we can without loss of generality assume that observations at the sensors are from the set of real numbers, i.e., $\mathcal{X}_n \in \mathbb{R}$; even for non-real observations the likelihood-ratios at the sensors are from the real set.

Returning now to the parallel configuration, let us consider the detection problem in which our aim is to design the decision rules of the sensor nodes and the FC by optimizing a performance measure, subject to the rate constraints of the sensor-to-FC channels. This problem can be formulated as:

$$
\begin{aligned}
&\text{Optimize: some performance measure,} \\
&\text{variables: } \gamma_1, \ldots, \gamma_N, \gamma_{\text{FC}}, \\
&\text{subject to: } r_1, \ldots, r_N .
\end{aligned}
\tag{5}
$$

Note that so far we have not discussed the choice of the performance metric and we postpone this issue to Section 1.4.

This problem was considered quite extensively in the literature, specially for the binary hypothesis test and when the sensor-to-FC channels are one-bit channels, and for conditionally independent observations at the sensors [9, 16]. Even under these simplified assumptions, the problem of designing decision rules at the sensors and the FC is a demanding task and in most cases, finding globally optimal decision rules at the sensors is mathematically intractable. To overcome this difficulty, we can use a person-by-person (PBP) optimization method for the design of decision rules at the sensors and the FC. In the person-by-person optimization approach, we optimize the decision rule of each sensor while assuming fixed decision rules at all other sensors and the FC. This approach is only guaranteed to reach a locally optimal solution and not necessarily a globally optimal solution [4].

This specific problem was considered by Longo *et al.* in [17] for general correlated observations at the peripheral nodes. The peripheral nodes (scalar quantizers) were to be cooperatively designed according to a system-wide measure of performance. They argued that the natural criterion of optimization—in their case the power of a Neyman-Pearson test—made the design procedure intractable. They therefore proposed to instead optimize a measure of similarity between the conditional distributions of the joint index space, i.e., the full set of quantizer outputs. Longo *et al.* obtained their final result using a cyclic person-by-person optimization algorithm for maximizing the Bhattacharyya distance [63] between the conditional probability distributions under the two hypotheses. We shall discuss this choice of performance measure in detail in Section 1.4.

The design problem is even more challenging when the sensors share a MAC channel to send their outputs towards the FC, as shown in Figure 2. In this situation the detection problem involves the design of decision rules

at the sensor nodes and the FC and, at the same time, allocation of rates
to the sensors in such a way that some communication constraints imposed
by the MAC channel are satisfied. The MAC channel model, as in [21], can
be considered as a set of error-free channels that are subject to a sum rate
constraint. We can mathematically formulate this problem as:

$$
\begin{aligned}
&\text{Optimize: some performance measure,}\\
&\text{variables: } \gamma_1, \ldots, \gamma_N, \gamma_{\text{FC}}, r_1, \ldots, r_N\\
&\text{subject to: } \sum_{n=1}^{N} r_n \leq R,
\end{aligned}
\tag{6}
$$

for a positive integer R, where R is sum rate capacity of the MAC channel.

For the case of a binary hypothesis test where the observations at the
N remote sensors are independent and identically distributed (i.i.d.) condi-
tioned on the true hypothesis H, Chamberland and Veeravalli [21] studied
the structure of an optimal sensor configuration. Their study was in terms
of the optimal number of sensors (N) and the optimal sensors rate alloca-
tion. They found sufficient conditions under which having $N = R$ one bit
(binary) sensors is optimal. Although the condition of having only rate one
sensors leads to a very simple network design, it is in many cases simply
not practically feasible to have an arbitrary number of sensors. The maxi-
mum number of sensors deployed in practice is often limited by hardware or
spatial constraints. Therefore in [62, 64] we have considered a more general
case which extends the results in [21]. Concretely, we have considered the
case where N is fixed or limited a-priori, and our aim is to optimally select
the set of rates in the network. We shall discuss these results in Section 2.2.

The optimal design of decision rules for the sensors arranged in tandem,
according to Figure 3, has also generated a significant amount of interest in
the past [27, 29, 33, 34, 65, 66]. Sensor $S_n, n = 2, \ldots, N$ in a tandem network
makes a decision according to its own observation x_n and the decision u_{n-1}
of its predecessor S_{n-1} using a decision function $\gamma_n : \mathcal{X}_n \times \mathcal{U}_{r_{n-1}} \to \mathcal{U}_{r_n}$,
i.e.,

$$
\gamma_n(x_n, u_{n-1}) = u_n,
\tag{7}
$$

where sensor S_N serves as the FC and therefore for its output message set
we have $\mathcal{U}_{r_N} = \mathcal{H}$. Sensor S_1 only uses its direct observation x_1 to make a
decision u_1 using a decision function $\gamma_1 : \mathcal{X}_1 \to \mathcal{U}_{r_1}$, i.e.,

$$
\gamma_1(x_1) = u_1.
\tag{8}
$$

As in the parallel topology, if there is no communication constraint on the
channels between the sensors, each sensor only gives its inputs as output
and the problem reduces to a centralized problem and the optimal fusion
decision γ_N can be found accordingly. Let us consider the case where the

channel between sensors S_n and S_{n+1} is an error-free but rate-constrained channel of rate r_n. Without loss of generality, we can assume that the output message $u_n, n = 1, \ldots, N - 1$ is from the set $\mathcal{U}_{r_n} = \{1, 2, \ldots, 2^{r_n}\}$, while the output message of S_N (the FC) is from the set $\mathcal{H} = \{H_0, \ldots, \mathcal{H}_{M-1}\}$ for an M-ary hypothesis testing problem. In this situation the detection problem is to design the decision rules of the sensor nodes and the FC by optimizing a performance measure, subject to the rate constraints of sensor-to-sensor channels, or equivalently:

$$
\begin{aligned}
&\text{Optimize: some performance measure,} \\
&\text{variables: } \gamma_1, \ldots, \gamma_N, \\
&\text{subject to: } r_1, \ldots, r_{N-1}.
\end{aligned}
\tag{9}
$$

The optimal design of the sensors in a tandem network was previously studied in [27, 29, 34] under the assumption that the observations at the sensors were, conditioned on the hypothesis, independent. This scenario has also been generalized in [33] to the case of conditionally dependent observations. While in [27, 29, 34] a binary hypothesis testing and binary messages between the sensors were considered, [33] relaxed these assumptions and considered general M-ary hypothesis testing, however with only M-valued messages (i.e., $\|\mathcal{U}_{r_n}\| = M$) for $M \geq 2$ and $n = 1, \ldots, N$.

Considering the optimal performance limits of tandem networks, it was shown in [16, 28] that for distributed networks with two sensors the optimal parallel network is outperformed by the optimal tandem network. However, for any network of more than two sensors, parallel networks perform better than serial networks, and for any given distributed detection problem with i.i.d. observations there exist a number of sensors at which the parallel network becomes better [16]. Moreover, the error probability in the case of a parallel topology with any logical decision functions, goes to zero very quickly as the number of sensors increases. This does however not hold in general for the tandem topology. In other words, as was shown in [65], the rate of error probability decay of the tandem network is always sub-exponential in the total number of sensors, while the error probability decay of a parallel network is exponential in the total number of sensors [67].

There are also some significant studies regarding the asymptotic performance of parallel and tandem networks, for instance [28, 65, 67–69]. In [28, 68] the problem of M-ary hypothesis testing in tandem networks was considered when the sensors were allowed to send M-valued messages. They have shown that in this situation a necessary and sufficient condition for the probability of error to asymptotically go to zero is that the log-likelihood ratio of the observation at each sensor, between any two arbitrary hypotheses, is unbounded in magnitude. We can then conclude that in the general case with potentially bounded log-likelihood ratios, having strictly more messages than the number of hypotheses is needed to have zero-limiting error

probability. In the case of a binary hypothesis testing (i.e., $M = 2$) and for the case of bounded log-likelihood ratios, Cover has proposed an algorithm in [68] with a four-valued message ($\|\mathcal{U}_{r_n}\| = 4$) which results in zero-limiting probability of error under each hypothesis. Koplowitz has generalized this idea in [69], where he has shown that, even if the log-likelihood ratios are bounded, $(M + 1)$-valued messages are sufficient for achieving zero-limiting probability of error in M-ary hypothesis testing.

Considering a tandem network of fixed size N, Papastavrou and Athans [28] proposed a simple but suboptimal scheme for the design of sensors. In their method, each sensor is optimized for locally minimal error probability at its output, instead of for globally optimal performance. For this particular scheme, they have also shown that a necessary and sufficient condition to achieve zero-limiting probability of error is that the log-likelihood ratio of the observation of each sensor be unbounded from both above and below. While optimizing the performance (e.g. minimizing the error probability) locally at the output of each sensor makes the algorithm relatively simple, it has a side effect that the messages are then constrained to be M-valued (i.e., $\|\mathcal{U}_{r_n}\| = M$) for the M-ary hypothesis testing problem. This is due to the fact that a one-to-one relation between the sensor output messages and the hypotheses is needed in the definition of the local error probabilities. Therefore, even though it is known that increasing the number of communication messages can improve the performance of a parallel network [70], the problem of designing the sensors in a tandem network for arbitrary-valued messages remains largely open [27]. Al-Ibrahim and AlHakeem [66] also have exemplified this point by allowing the first sensor in a tandem network of two sensors to communicate two-bit messages instead of one-bit messages. They have observed a significant improvement in the performance of the network for binary hypothesis testing. One way to view this result is as follows: multi-bit (soft) decisions are able to transmit more information to the FC for the final decision than a binary (hard) decision would. The difficulty is in figuring out how to best capture and quantize this additional information and this is one problem that we shall address in Section 2.1 of this thesis.

In the following section we will consider different choices of performance measure and discuss their properties.

1.4   Performance measures

In the Bayesian formulation, a decision at the FC is made in favor of a hypothesis based on given prior information, namely a-priori probability of hypotheses, and hypothesis-conditional probabilities. The optimal decision

rule is the Bayes' test that minimizes the Bayes' risk which is equal to

$$\mathcal{R} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} C_{ij} \pi_j P\left(\widehat{H}_i | H_j\right), \tag{10}$$

where C_{ij} represents the cost of deciding on the presence of $\widehat{H}_i$ while H_j is true and $P\left(\widehat{H}_i | H_j\right)$ represents the probability of deciding on the presence of $\widehat{H}_i$ while H_j is true [9].

When no cost is assigned for a correct decision and unit cost is assigned for a wrong decision, i.e.,

$$C_{ij} = \begin{cases} 1 & i \neq j, \\ 0 & i = j, \end{cases}$$

Bayes' risk simplifies to Bayes' error probability and Bayes' rule reduces to the *maximum a-posteriori* (MAP) decision rule [71]. According to the MAP rule, the FC based on its input z decides on hypothesis $\widehat{H}_i$ if [72]

$$\widehat{\pi}_i P_{Z|H}\left(z | \widehat{H}_i\right) = \max\left\{\pi_0 P_{Z|H}\left(z | H_0\right), \ldots, \pi_{M-1} P_{Z|H}\left(z | H_{M-1}\right)\right\}, \tag{11}$$

for a general M-ary hypothesis testing problem, where $P_{Z|H}\left(z|H_j\right)$ is the hypothesis-conditional probability of input $z \in \mathcal{Z}$ at the FC, and where $\widehat{\pi}_i$ is the a-priori probability of hypothesis $\widehat{H}_i$. The expected error probability when using the MAP rule at the FC is [73]

$$P_\mathrm{E} = 1 - \sum_{z \in \mathcal{Z}} \max\left\{\pi_0 P_{Z|H}\left(z | H_0\right), \ldots, \pi_{M-1} P_{Z|H}\left(z | H_{M-1}\right)\right\}. \tag{12}$$

In practice, Bayes' error probability is quite tricky to calculate and in the following we will discuss some of these difficulties in the design of sensor networks. We will start with the parallel network and find hypothesis-conditional probabilities and then discuss the calculation of hypothesis-conditional probabilities for the tandem network.

For a network of N sensors arranged in parallel or MAC, as in Figures 1 and 2 respectively, the hypothesis-conditional probability is

$$P_{Z|H}\left(z | H_j\right) = P_{\underline{U}|H}\left(u_1, \ldots, u_N | H_j\right), \tag{13}$$

where $\underline{U} \triangleq U_1, \ldots, U_N$, and where the complete input set at the FC is $\mathcal{Z} = \underline{\mathcal{U}} \triangleq \mathcal{U}_{r_1} \times \ldots \times \mathcal{U}_{r_N}$. When the observations X_n at the sensors, conditioned on the hypothesis, are independent, the hypothesis-conditional probability decouples as

$$P_{Z|H}\left(z | H_j\right) = P_{U_1|H}\left(u_1 | H_j\right) \ldots P_{U_N|H}\left(u_N | H_j\right). \tag{14}$$

This is due to the fact that each sensor output U_n is only a function of its corresponding input X_n. Since functions of independent random variables are also independent, the independence of sensors' inputs results in the independence of sensors' outputs.

Now consider a sensor S_n in the parallel network. Given a sensor decision rule γ_n and the true hypothesis $H_j \in \mathcal{H}$, the probability mass function associated with the message $U_n = \gamma_n(X_n) \in \mathcal{U}_{r_n}$ is calculated as

$$P_{U_n|H}\left(u_n|H_j\right) = \Pr\left(\gamma_n(X_n) = u_n|H_j\right)$$
$$= \int_{x \in \gamma_n^{-1}(u_n)} f_{X_n|H}\left(x|H_j\right) dx, \tag{15}$$

where $\gamma_n^{-1}(u_n)$ is the set of observations $x \in \mathcal{X}$ that satisfy $\gamma_n(x) = u_n$. This can be generalized to the joint probability mass function associated with the message set $\underline{U} = U_1, \ldots, U_N$ at the FC as

$$P_{\underline{U}|H}\left(u_1, \ldots, u_N|H_j\right) =$$
$$\int_{x_1 \in \gamma_1^{-1}(u_1)} \ldots \int_{x_N \in \gamma_N^{-1}(u_N)} f_{\underline{X}|H}\left(x_1, \ldots, x_N|H_j\right) dx_1 \ldots dx_N. \tag{16}$$

As (13) shows, in order to design the decision rules at the sensors, such that the MAP error probability in (12) is minimized, we need to find the joint probability mass functions which are found from the joint conditional distribution of the observations at the sensors. This makes the direct use of (12) as a performance measure for the design of sensor rules in a parallel network limited. However, using some mathematical reformulations in [74], we have shown that using the same complexity order as the existing methods (for example the proposed method in [17]) we can use (12) in a parallel network to design the sensors' decision rules, in a person-by-person framework.

For a network of N sensors arranged in tandem, as in Figure 3, the hypothesis-conditional probability is

$$P_{Z|H}\left(z|H_j\right) = P_{X_N, U_{N-1}|H}\left(x_N, u_{N-1}|H_j\right), \tag{17}$$

where for consistency we assume a discrete observation space $\mathcal{X}_n$ at (the sensors and therefore) the FC in a tandem network. Note that, although we restrict our attention to discrete observation spaces, $\mathcal{X}_n$ can be used to approximate observations in a continuous space using fine-grained binning. Longo *et al.* [17] used this idea by representing each bin, or interval, in the continuous observation space by an index x_n from the discrete set $\mathcal{X}_n$. The complete input set for the tandem network at the FC is then $\mathcal{Z} = \mathcal{X}_N \times \mathcal{U}_{r_{N-1}}$.

Using the same argument as in (14), for conditionally independent observations at the sensors, the hypothesis-conditional probability at the FC

decouples as

$$P_{Z|H}\left(z|H_j\right) = P_{X_N|H}\left(x_N|H_j\right) P_{U_{N-1}|H}\left(u_{N-1}|H_j\right). \tag{18}$$

Note that for conditionally independent observations at the sensors, (18) generalizes to a continuous observation space. To do this, we can replace probability mass function $P_{X_N|H}\left(x_N|H_j\right)$ with probability density function $f_{X_N|H}\left(x_N|H_j\right)$.

Now consider sensor S_n for $2 \leq n \leq N$ in the tandem network. Given its sensor rule γ_n and the true hypothesis $H_j \in \mathcal{H}$, the probability mass function associated with its message $U_n = \gamma_n(X_n, U_{n-1}) \in \mathcal{U}_{r_n}$ is calculated as

$$\begin{aligned}
P_{U_n|H}\left(u_n|H_j\right) &= \Pr\left(\gamma_n(X_n, U_{n-1}) = u_n|H_j\right) \\
&= \sum_{(x_n, u_{n-1}) \in \gamma_n^{-1}(u_n)} P_{X_n, U_{n-1}|H}\left(x_n, u_{n-1}|H_j\right),
\end{aligned} \tag{19}$$

where $\gamma_n^{-1}(u_n)$ is the set of tuples (x_n, u_{n-1}) that satisfy $\gamma_n(x_n, u_{n-1}) = u_n$. For sensor S_1, the probability mass function associated with its message $U_1 = \gamma_1(X_1) \in \mathcal{U}_{r_1}$ is calculated as

$$P_{U_1|H}\left(u_1|H_j\right) = \sum_{x_1 \in \gamma_1^{-1}(u_1)} P_{X_1|H}\left(x_1|H_j\right), \tag{20}$$

with the corresponding definition for $\gamma_1^{-1}(u_1)$. Equations (19) and (20) show that every FC output u_N depends on all the observations $x_1, \ldots, x_N$ through the sensor functions $\gamma_1, \ldots, \gamma_N$ in a complicated recursive way. This can also be seen from (7) and (8) for any FC decision $\widehat{h} \in \mathcal{H}$ as follows:

$$\begin{aligned}
\widehat{h} = u_N &= \gamma_N\left(x_N, u_{N-1}\right) \\
&= \gamma_N\left(x_N, \gamma_{N-1}\left(x_{N-1}, u_{N-2}\right)\right) \\
&\vdots \\
&= \gamma_N\left(x_N, \gamma_{N-1}\left(x_{N-1}, \gamma_{N-2}\left(x_{N-2}, \ldots, \gamma_2\left(x_2, \gamma_1\left(x_1\right)\right)\ldots\right)\right)\right).
\end{aligned} \tag{21}$$

These mathematical formulations show that finding the decision rules at the sensors which directly minimize the MAP error probability in a tandem network is even more challenging than in a parallel network for arbitrary network size. However, for small-sized networks (i.e., $N = 2$) and conditionally independent observations at the sensors, the MAP rule provides a computationally tractable tool for the design of such networks, and in [75], we have used this idea for the design of an arbitrary-sized tandem network

in a person-by-person framework. We have further generalized this idea in [76] for the design of sensor decision rules in an arbitrary tree topology.

Although the person-by-person methodology provides a computationally tractable tool for the design of sensor decision rules in a wireless sensor network, it does not necessarily lead to a globally optimal design and the problem of understanding the characteristics of an optimal network remains open. One way to tackle this problem in a binary hypothesis testing is to study the structure of optimal networks in terms of their error exponent. In what follows we will be discussing the Chernoff information and the Bhattacharyya distance as two performance metrics for the design of decision rules of sensors and the FC of a network [52, 63, 77].

Consider a binary hypothesis testing problem where sensors are arranged in parallel or MAC. The expected error probability when using the MAP rule at the FC is [see (12)]

$$
\begin{aligned}
P_{\mathrm{E}} &= 1 - \sum_{\underline{u}} \max \left\{ \pi_0 P_{\underline{U}|H} \left(\underline{u}|H_0 \right), \pi_1 P_{\underline{U}|H} \left(\underline{u}|H_1 \right) \right\} \\
&= \sum_{\underline{u}} p\left(\underline{u}\right) \left(1 - \max \left\{ P_{H|\underline{U}} \left(H_0|\underline{u} \right), P_{H|\underline{U}} \left(H_1|\underline{u} \right) \right\} \right) \\
&= \sum_{\underline{u}} p\left(\underline{u}\right) \left(\min \left\{ 1 - P_{H|\underline{U}} \left(H_0|\underline{u} \right), 1 - P_{H|\underline{U}} \left(H_1|\underline{u} \right) \right\} \right) \\
&= \sum_{\underline{u}} \min \left\{ \pi_0 P_{\underline{U}|H} \left(\underline{u}|H_0 \right), \pi_1 P_{\underline{U}|H} \left(\underline{u}|H_1 \right) \right\}.
\end{aligned}
\tag{22}
$$

Since the Bayes' error probability in (22) minimizes the average probability of error, bounding it tightly is crucial in hypothesis testing [78]. To upper bound the Bayes' error in (22), let us replace the min function by the smooth power function: namely for $a, b > 0$, we have

$$
\min\{a, b\} \leq a^{\alpha} b^{1-\alpha}, \forall \alpha \in (0, 1).
$$

Therefore we obtain the following Chernoff bound for the error probability [79]

$$
\begin{aligned}
P_{\mathrm{E}} &\leq \pi_0^{\alpha} \pi_1^{1-\alpha} \sum_{\underline{u}} \left(P_{\underline{U}|H} \left(\underline{u}|H_0 \right) \right)^{\alpha} \left(P_{\underline{U}|H} \left(\underline{u}|H_1 \right) \right)^{1-\alpha} \\
&\leq \sum_{\underline{u}} \left(P_{\underline{U}|H} \left(\underline{u}|H_0 \right) \right)^{\alpha} \left(P_{\underline{U}|H} \left(\underline{u}|H_1 \right) \right)^{1-\alpha},
\end{aligned}
\tag{23}
$$

which follows from $\pi_0, \pi_1 \leq 1$ and holds for any $\alpha \in (0, 1)$. Since (23) is true for any $\alpha \in (0, 1)$ we can take the minimum of it to achieve the Chernoff information bound as follows

$$
P_{\mathrm{E}} \leq e^{-\mathcal{C}_r(\underline{\gamma})},
\tag{24}
$$

where $\mathcal{C}_{\underline{r}}(\underline{\gamma})$ is the Chernoff information at the input of the FC for the given rate allocation $\underline{r} \triangleq r_1, \ldots, r_N$ and sensor rules $\underline{\gamma} \triangleq \gamma_1, \ldots, \gamma_N$,

$$\mathcal{C}_{\underline{r}}(\underline{\gamma}) \triangleq -\log \min_{\alpha \in (0,1)} \sum_{\underline{u}} \left(P_{\underline{U}|H}\left(\underline{u}|H_0\right) \right)^{\alpha} \left(P_{\underline{U}|H}\left(\underline{u}|H_1\right) \right)^{1-\alpha}. \tag{25}$$

The inequality in (24) suggests maximizing the Chernoff information in place of minimizing the error probability as a performance measure in the design of sensor networks, with the cost of solving an optimization problem. The Chernoff information was used as the performance measure in hypothesis testing problems, see [21, 80–82]. Lee and Sung [80] considered the performance of mismatched likelihood ratio detectors for binary hypothesis testing problems. Using large deviation theory, they have derived the maximum Bayesian error exponent for a mismatched detector and have shown that the maximum Bayesian error exponent is given by the Chernoff information. In [81], Fabeck and Mathar designed the decision rule at the sensors by locally optimizing the Chernoff information. In [82], Chamberland and Veeravalli studied the performance of power constrained wireless sensor networks in a parallel topology when the channels between the sensors and the FC are subject to additive noise, while in [21], they studied the structure of an optimal network of sensors arranged in a MAC configuration which is subject to a sum rate constraint.

While the Chernoff information provides a tight upper bound for the probability of error at the FC, for many applications and for hypothesis-conditional distributions finding the optimal parameter α can be mathematically intractable. In these situations a relatively looser upper bound can be used. By setting $\alpha = 0.5$ in (25) we obtain the Bhattacharyya distance [63] as

$$\mathcal{B}_{\underline{r}}(\underline{\gamma}) \triangleq -\log \sum_{\underline{u}} \sqrt{P_{\underline{U}|H}\left(\underline{u}|H_0\right) P_{\underline{U}|H}\left(\underline{u}|H_1\right)}. \tag{26}$$

It immediately follows that

$$\mathcal{B}_{\underline{r}}(\underline{\gamma}) \leq \mathcal{C}_{\underline{r}}(\underline{\gamma})$$

as the Bhattacharyya distance can be obtained from (25) with $\alpha = 0.5$ in place of optimization over α as noted above. The Bhattacharyya distance also provides an upper bound on the Bayesian error probability of the MAP detector according to [83]

$$P_{\mathrm{E}} \leq \sqrt{\pi_0 \pi_1} e^{-\mathcal{B}_{\underline{r}}(\underline{\gamma})} \tag{27}$$

which can be found from (23). As mentioned above, the main reason for using the Bhattacharyya distance in place of the Chernoff information is that it will increase the tractability of the problem by dropping the optimization over α. Furthermore, considering the Bhattacharyya distance instead

of the Chernoff information (in terms of mathematical tractability) has the following benefit: the Bhattacharyya distance of a network of sensors, arranged in parallel or MAC and with independent observations, is equal to the sum of the Bhattacharyya distances of individual sensors. Thus, a network which maximizes the Bhattacharyya distance at the FC is a network with individually optimized sensors.

It should however be noted that the Bhattacharyya distance has been frequently used in the past as a performance measure in the design of distributed detection systems [17, 52]. It will be used in this thesis for the problem of designing decision rules at the sensors and allocating rates to the sensors arranged in MAC configuration [62]. The Bhattacharyya distance will also be used as the performance measure for the design of sensors in a network of energy harvesting sensors [84, 85], as we shall discuss in the next section.

1.5   Energy constrained wireless sensor networks

In wireless sensor networks, a large number of sensors with small batteries and limited lifetimes are often used. Their limited lifetimes are a major limitation of using them [86]. In other words, the sensors work as long as their batteries last and this implies that the network itself also has a limited lifetime. To increase the lifetime of battery-powered sensor networks many solutions have been proposed [87–92], e.g., by choosing the best modulation strategy [89], or by exploiting power-saving modes (sleep/listen) periodically [91]. However, in all of these methods the aim is to find an energy usage strategy to maximize the lifetime of a network, and therefore the lifetime remains bounded and finite. An alternative way of dealing with this problem is to use energy harvesting devices at the sensor nodes. An energy harvesting device is capable of acquiring energy from nature or from man-made sources [93–95].

Energy harvesting technologies provide a promising future for wireless sensor networks, such as self sustainability and an effectively perpetual network lifetime which is not limited by the sensor battery's lifetime [95–97]. While acquiring energy from the environment makes it possible to deploy wireless sensor networks in situations which are impossible using conventional battery-powered sensors, it poses new challenges related to the management of the harvested energy. These new challenges are due to the randomness of the amount of energy available at a sensor, since the source of energy might not be available at all times that we may want to use the sensor nodes [96, 98–106].

Figure 5 shows a network of energy harvesting sensors, where sensors are arranged in parallel. Each sensor $S_n, n = 1, \ldots, N$ during time interval t makes an observation $x_{n,t}$ and sends a message $u_{n,t}$ towards the FC. In this configuration the output message of the sensor depends not only on its
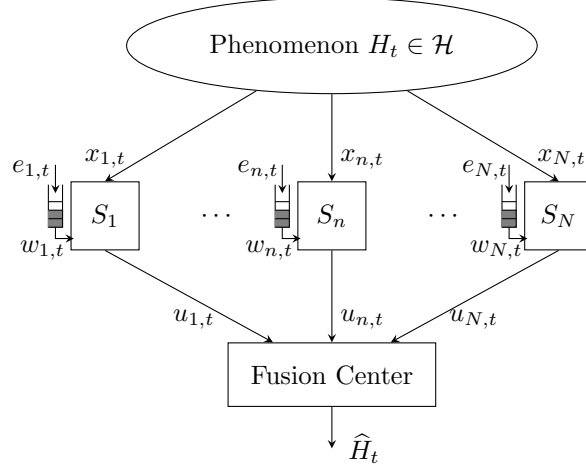
Figure 5: Decentralized hypothesis testing scheme in a parallel network of energy harvesting sensors.

observation, but also on its battery charge, i.e.,

$$\gamma_n(x_{n,t}, b_{n,t}) = u_{n,t} \,,$$

where $b_{n,t}$ denotes the battery charge of sensor S_n at transmission time t. A central problem in this configuration is how to design the sensors' decision rules in such a way that a performance measure is optimized. However, this problem is more challenging than the conventional problem of designing sensors' decision rules in a parallel network, due to the randomness of the battery charges. Note that in this situation, energy (denoted by $e_{n,t}$) arrives to the sensor according to a random process and in general the FC is not aware of the battery charges. To the best of our knowledge, the problem of decentralized hypothesis testing using energy harvesting sensors has not been considered before and in [84, 85] we have considered this problem. Concretely, in these works, we have formulated a decentralized detection problem with system costs due to the random behavior of the energy available at the sensors and we have shown how the problem formulation changes (compared to the unconstrained case) when we consider the energy features in the problem of designing decision rules at the sensors in the network.

2   Our contributions

In this section we discuss our contributions to the problem of decentralized hypothesis testing in wireless sensor networks. In the first part, we shall consider the problem of designing decision rules at the sensors and the FC in different configurations. We have studied this problem in [74–76] labeled as Papers A–C. In the second part, we shall consider the problem of rate allocation in wireless sensor networks. We have studied this problem in [62, 64] labeled as Papers D–E. Finally, we shall consider the problem of decentralized detection in energy harvesting sensor networks. We have studied this problem in [84, 85] labeled as Papers F–G. For each paper, we will briefly discuss the system model and assumptions and present the central results.

2.1   Problem of designing sensors' decision functions

Paper A: Bayesian design of decentralized hypothesis testing under communication constraints [74]

In this paper we have considered a binary hypothesis testing problem in the parallel topology as in Figure 1, where the peripheral nodes observe a common phenomenon and send their possibly correlated observations towards the FC via error-free but rate-constrained channels. The objective in this paper is to design the sensor nodes' decision rules $\gamma_n, n = 1, \ldots, N$ and the FC decision rule γ_{FC} such that the Bayesian error probability P_{E} at the FC is minimized.

This problem was studied by Longo *et al.* [17], who considered each sensor to be a scalar quantizer that satisfies the rate-constraints. They proposed a cyclic person-by-person optimization algorithm to design the quantizers. Longo *et al.* argued that the natural criterion of optimization—the power of a Neyman-Pearson test—made the design procedure intractable. They therefore proposed to instead maximize the Bhattacharyya distance (26) between the conditional probability distributions under the two hypotheses.

The main contribution of our work was to show that, contrary to previous claims, it is in fact possible, within a person-by-person framework, to minimize the probability of error directly, while having the same complexity order as in the previous work by Longo *et al.* To this end, we have proposed combining the idea of *fine-grained observation-binning* used in [17] with a method for updating the hypothesis-dependent mass functions of the joint index space (13).

The benefits of the proposed method are that: it takes into account the a-priori probability of the hypotheses in the design procedure, and it is applicable to general M-ary hypothesis testing problems and not only binary hypothesis tests. Furthermore, the proposed method shows better
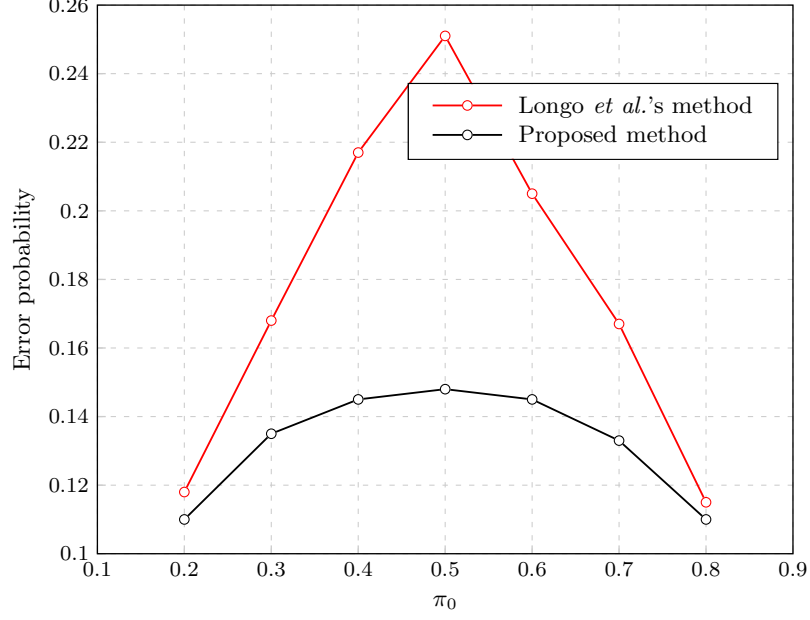
Figure 6: Error probability performance of the designed networks using different methods for different a-priori probabilities.

performance in terms of both the receiver operating characteristic (ROC) and the error probability. To exemplify these results, we have considered a network of $N = 2$ sensors with the following observation model:

$$
\begin{aligned}
H_0 &: \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \left[\begin{array}{c} n_1 \\ n_2 \end{array} \right], \\
H_1 &: \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \left[\begin{array}{c} a \\ -a \end{array} \right] + \left[\begin{array}{c} n_1 \\ n_2 \end{array} \right],
\end{aligned}
\tag{28}
$$

where $\mathbf{n} = \left[\begin{array}{c} n_1 \\ n_2 \end{array} \right]$ is a zero-mean Gaussian vector of covariance matrix

$$
\mathbf{\Sigma} = \left(\begin{array}{cc} \sigma^2 & r\sigma^2 \\ r\sigma^2 & \sigma^2 \end{array} \right),
$$

where $r = 0.9$ is the spatial correlation coefficient. As in [17], we have assumed that per channel signal-to-noise ratio $(\mathcal{E}/\sigma^2)$ is -5 dB and channel rates are the same, i.e., $r_1 = r_2$. We divided the interval $[-4, +4]$ (containing 0.9997 of the total probability mass for each DM) into 256 bins, and designed the sensors for various values of a-priori probabilities and channel rates

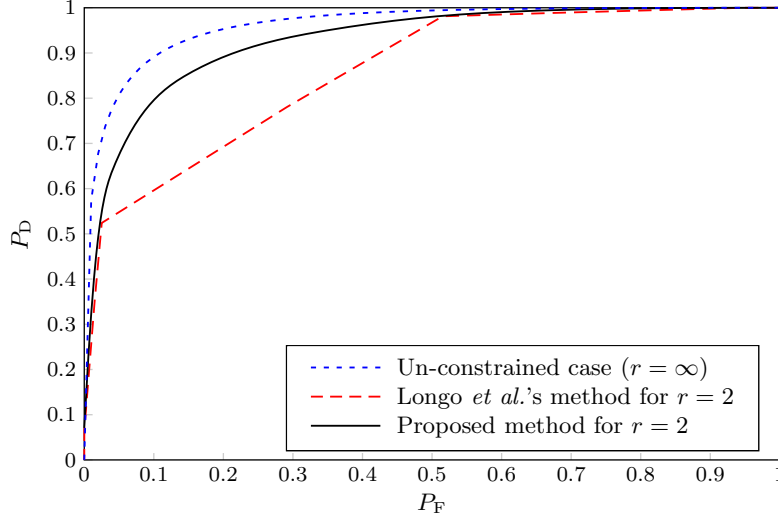Figure 7: The complete ROC curve achievable using the proposed design method, Longo *et al.*'s method, and the optimal centralized design.

(in bits/sample). Figure 6 shows the error probability performance of the sensor network designed using the proposed method and using Longo *et al.*'s method, and Figure 7 illustrates the achievable ROC curves for both design methods. Note that the close relation between the error probability and the ROC makes it possible to compare the results of the proposed method with those of Longo *et al.*'s in terms of the ROC curve as well. These figures imply that although the proposed design method is based on the Bayesian error probability, it also outperforms the previous method in terms of the ROC curve.

To see how the two methods work, let us consider the decision regions found using these methods in Figure 8, when $\pi_0 = 0.5$ and with the same parameters as before, i.e., $r = 0.9, r_1 = r_2 = 2$ and the same observation model as in (28) and for $a = 0.5$. In this figure, the decision boundaries for each sensor are shown by solid blue lines and the final decision by the FC (which uses the optimal MAP rule) in each region is shown by (red colored) zeros and ones, where zero and one correspond to H_0 and H_1, respectively. We further show sample values $p(x_1, x_2)$ by cyan and their conditional mean values by solid black circles. Also the corresponding decision regions using likelihood test [107] (for centralized design) are shown by dashed lines in both subfigures. As we can see, the decision regions in the proposed method are such that more likely observations are classified correctly, while less likely
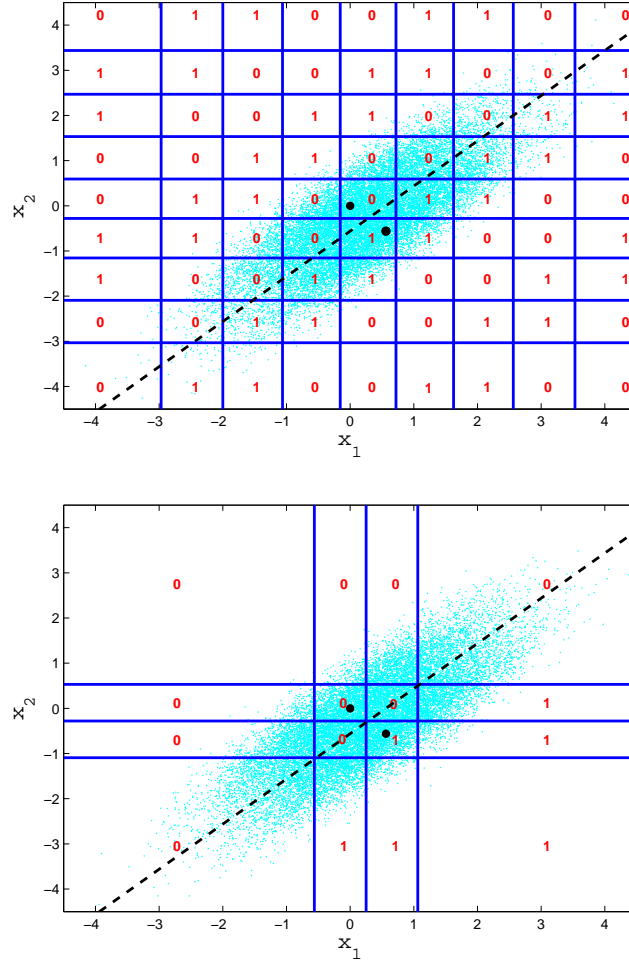
Figure 8: Decision regions resulting from our method (top) and Longo *et al.*'s method (bottom) for a sample data. Decision boundary using likelihood test for centralized design is also shown (dashed).

observations may not be classified correctly. In other words, the proposed method performs better in the detection of more probable observations, while it performs worse in the detection of less likely observations, which results in a better overall performance for the system.

Consider again the decision regions in Figure 8. We can see that the

designed sensors using *Longo et al.*'s method are monotone quantizers of rate 2: each axis is divided into 2^2 non-overlapping intervals, each interval corresponds to a sensor output message from the set $\{0, 1, 2, 3\}$, while the designed sensors using the proposed method are not monotone quantizers. Though it is known that monotone quantizers are often optimal decision functions at the sensors with conditionally independent observations, they are not proved to be in general optimal decision functions when the sensors make correlated observations. This can also be seen from our simulation results in Figure 8.

Paper B: Bayesian design of tandem networks for distributed detection with multi-bit sensor decisions [75]

In this paper we have considered the problem of decentralized hypothesis testing where several peripheral nodes are arranged in tandem as in Figure 3. We have assumed that the observations at the sensors, conditioned on the true hypothesis, are independent. Furthermore, the channels between every two successive sensors are error-free but rate-constrained. The objective in this paper is to design the sensor nodes' decision rules such that the error probability at the FC (the last sensor) is minimized.

This problem was previously considered in [27, 68] for a binary hypothesis test, where the sensors were able to send one-bit messages. Though these studies then were extended to M-ary hypothesis testing problems, the output of each sensor was constrained to be from an M-valued message set. In this paper we have proposed a cyclic numerical design algorithm for the design of sensors in a person-by-person methodology, where the number of communicated messages was not necessarily equal to the number of hypotheses.

Our main contribution is to introduce a numerical methodology for designing sensors' decision rules in a tandem network, where each sensor is capable of sending arbitrary-valued messages. We have designed the sensors in such a way that the final error probability of the network is minimized. We have proposed a modified person-by-person optimization in which each sensor's decision rule is jointly designed with the FC (the last sensor). In other words, we have designed each sensor under the assumption that the FC always employs the optimal MAP rule to its inputs.

First we have shown that the design of each sensor $S_n, n = 1, \ldots, N-1$ in this framework is equivalent to the design of sensor S in a network labeled as a *restricted network*, as shown in Figure 9, where S_N in both networks uses the MAP rule. This means that regardless of the network size, each sensor in a tandem network can be designed using the restricted model with a fixed computational burden.
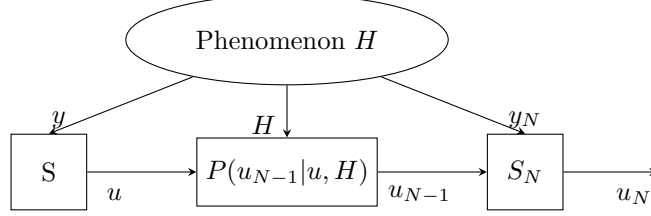
Figure 9: Restricted model used for the design of sensors in tandem networks.

In other words, in the proposed person-by-person methodology (where a sensor is designed together with the FC while all other sensors are kept fixed), the design of the sensor is equivalent to the design of sensor S in the restricted model. We have then found parameters of the restricted model according to the structure of the network: $P(u_{N-1}|u, H)$ was found according to the other sensors' (fixed) decision rules, y was found according to the sensor observation and the input from its predecessor, except for the first sensor in the network, while y_N had the same distribution as the FC observation.

Next, we have proposed a computationally efficient algorithm for the design of sensor S in the restricted model. We have further found the overall complexity of the proposed algorithm for the design of sensors in a tandem network.

In order to illustrate the benefits of the proposed method, we have compared the performance of designed networks using our proposed method with those of Cover [68] and Swaszek [27] for binary hypothesis testing. Figure 10 shows the performance of networks with different numbers of sensors designed using different methods. We assumed each real valued observation consists of an antipodal signal $\pm a$ in unit-variance additive white Gaussian noise, i.e.,

$$H_0 : x_i = -a + n_i \,,$$
$$H_1 : x_i = +a + n_i \,.$$

We also defined the per channel SNR for the binary hypothesis test as $\mathcal{E} \triangleq |a|^2$, and assumed that the hypotheses are equally likely ($\pi_0 = \pi_1 = 0.5$).

As noted before, the proposed method can be applied to general M-ary hypothesis testing problems. To see this, we considered a ternary hypothesis testing problem in which each real valued observation consists of a known signal $s_m, m = 0, 1, 2$ in unit variance additive white Gaussian noise. We have assumed an equal distance signal set in the interval $[-a, a]$, i.e., the
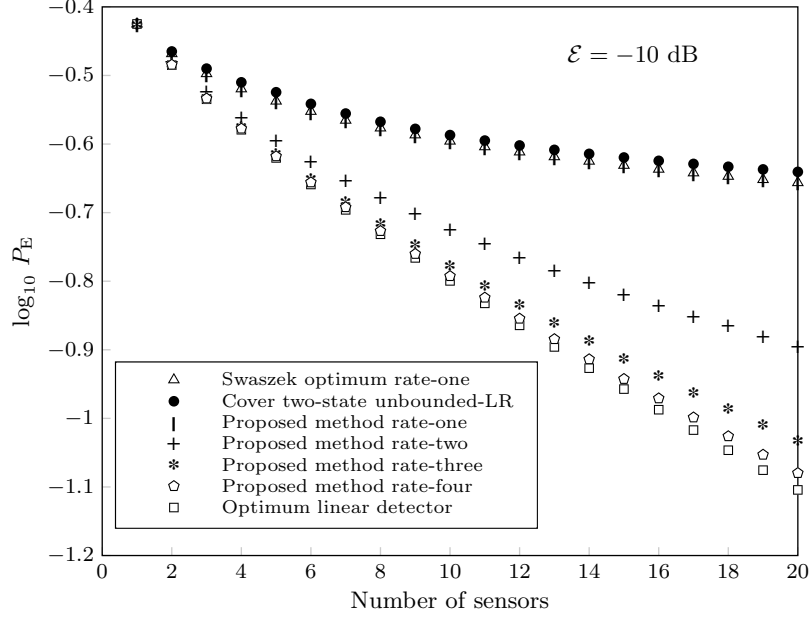
Figure 10: Error probability performance of tandem networks with different channel rates for a binary hypothesis testing problem as a function of number of sensors. Error probability of an unconstrained tandem network and existing methods for rate-one channels are also shown.

test signal set was $\{-a, 0, a\}$. The observation model at each sensor was

$$H_m : x_i = s_m + n_i, \quad m = 0, 1, 2.$$

Similar to the binary hypothesis test, we assumed equally likely hypotheses and defined the SNR as $\mathcal{E} \triangleq |a|^2$. Figure 11 shows the performance of networks with different numbers of sensors designed using the proposed method for different channel rates. As we can see from Figure 10 and Figure 11, networks with multi-bit (soft) message sensors outperform those of binary (hard) message sensors and by increasing the channel rates (length of transmitted messages) the performance of the networks improves.
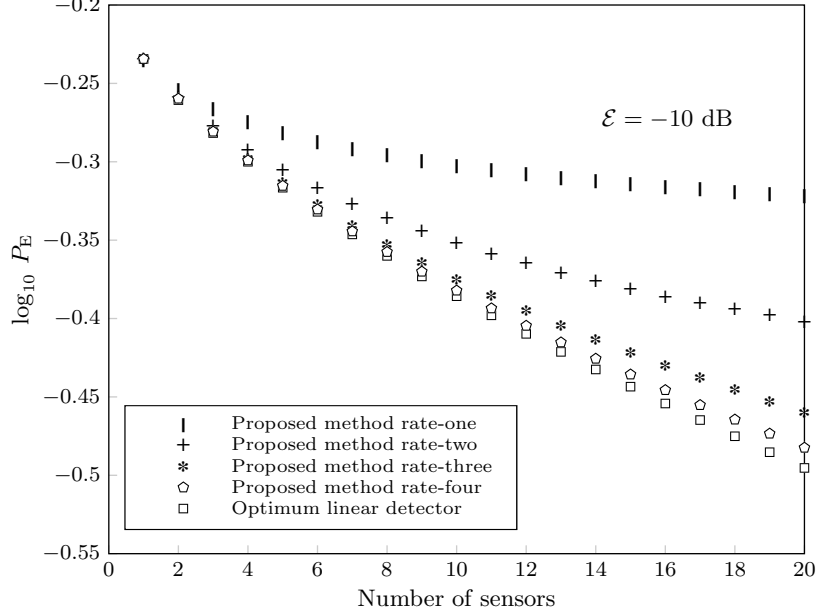
Figure 11: Error probability performance of tandem networks with different channel rates for different numbers of sensors with an unconstrained tandem network, for a ternary hypothesis testing problem.

Paper C: A general method for the design of tree networks under communication constraints [76]

This paper mainly generalizes the results in [75]. In this paper, we have considered a distributed detection problem where several nodes are arranged in an arbitrary tree topology. Communication channels between the sensors are again rate-constrained but error-free and observations at the sensors are conditionally independent.

We have proposed a cyclic design procedure using the expected minimum error probability by adopting a person-by-person methodology for the design of sensors' decision rules in the network. Concretely, we design the decision rule of each sensor jointly together with the FC while all the other sensors in the network are kept fixed. In order to obtain a tractable solution during the design of a sensor, say S_n, all other sensors were modeled using a Markov chain. We have further shown that the design of sensor S_n jointly with the FC is analogous to the design of a special case of a network with only two nodes, which is again referred to as the restricted model. The restricted model for the design of sensors in a tree network is shown in Figure 12.
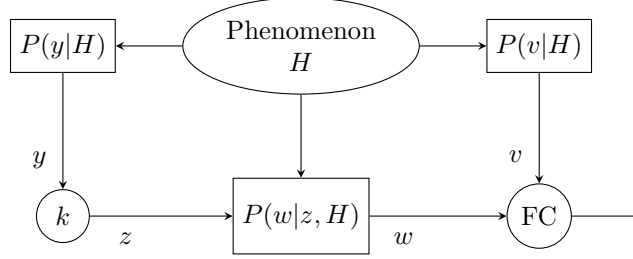
Figure 12: Restricted model for the design of nodes in tree topology.

Next, we have shown how, according to the structure of the problem and Markovian properties of sensors in a network, parameters of the restricted model can be found in a computationally efficient way.

As an example consider the tree network shown in Figure 13. Let us assume only sensors $S_1, \ldots, S_4$ make observations (labeled as leaves) and sensors S_5, S_6 are relays that summarize their received messages from their neighboring nodes and send their messages towards the FC. We also assume that the observation model at each leaf consists of an antipodal signal $\pm a$ in additive unit-variance white Gaussian noise as

$$H_0 : x_i = -a + n_i \,,$$
$$H_1 : x_i = +a + n_i \,.$$

We define the per channel SNR for a binary hypothesis test again as $\mathcal{E} \triangleq |a|^2$, and we assume that the hypotheses are equally likely ($\pi_0 = \pi_1 = 0.5$). We further assume leaf-to-relay and relay-to-FC channels have the same rate r. Using the proposed algorithm, we have designed the tree network for different rates r. In Figure 14, error probability performance of the

Figure 13: An example of a tree network.

Figure 14: Error probability performance of a designed tree network shown in Figure 13 for different channel rates.

designed tree network for different rates is shown. As expected, increasing channel rates results in better error probability performance.

2.2   Problem of rate allocation in wireless sensor networks
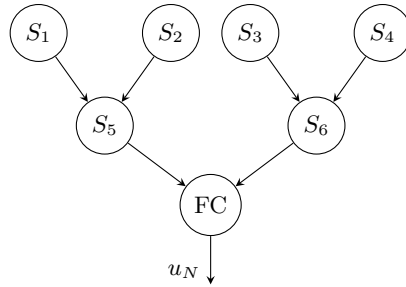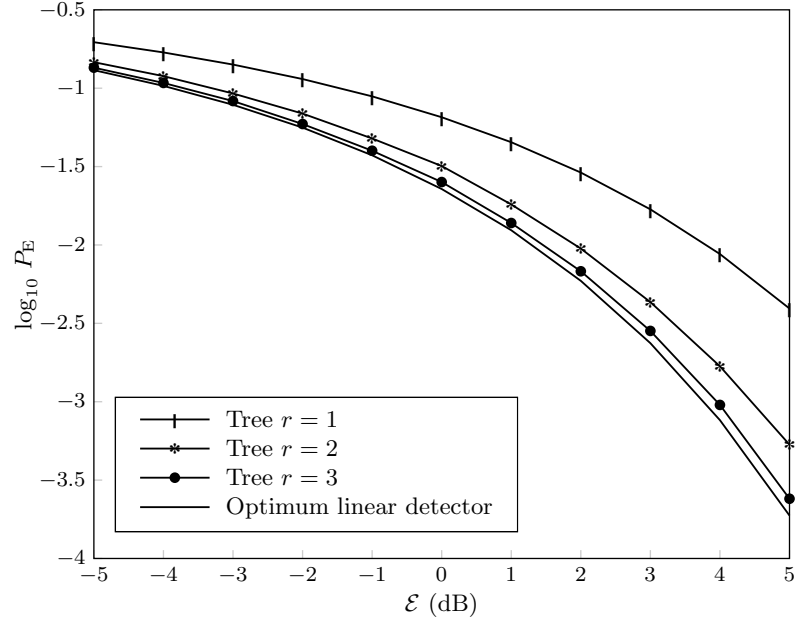
Paper D: Rate allocation for decentralized detection in wireless sensor networks [64]

In this paper we have considered the problem of rate allocation to sensors arranged as in Figure 2, in which the sensors make noisy observations of a binary phenomenon and send a quantized information towards a FC to make the final decision. The sensors send their messages through a MAC channel which is modeled as a sum-rate constrained channel, with capacity R.

Using the Chernoff information (25), Chamberland and Veeravalli [21] studied the structure of an optimal network model in terms of the optimal number of sensors and their rate allocation. They proved that, if for a given observation model there exists a (one-bit) quantization rule which leads to the transfer of at least half of the Chernoff information contained in each raw observation, an optimal strategy is to employ $N = R$ sensors each quantizing its observation to a one-bit message. Although their result greatly simplifies the design of sensors in a MAC network, it suffers from the constraining assumption of having an unlimited number of active sensors. In fact, due to cost and space constraints, it may not be feasible to have $N = R$ sensors in a network.

In this paper we have addressed the problem of finding an optimal rate allocation $\underline{r}$ when the total number of active sensors N is fixed a-priori. We assumed $R = mN$, where m is a positive integer. Using the Chernoff information at the FC as a performance measure, we have found conditions under which uniform rate allocation to the sensors in the network is an optimal rate allocation. These conditions are captured in Theorem 1 of the paper and is restated as follows:

Theorem: Given a sensor design method, if for a single sensor S_n the resulting Chernoff information $\mathcal{C}_{r_n}$ is a discrete concave function of rate r_n, a uniform rate allocation across sensors is optimal.

As stated above, the optimality of the uniform rate allocation in the network relies on the concavity of the Chernoff information of the sensor design method. In order to show the benefits of the results, we have explored this point numerically for a couple of sensor design methods. First, we have considered Benitz and Bucklew's [54] method for the design of a sensor rule (or quantization rule) in detection with i.i.d. observations. When the observation model at any sensor consists of an antipodal signal $\pm a$ in an additive unit-variance white Gaussian noise, or

$$H_0 : x_i = -a + n_i \,,$$
$$H_1 : x_i = +a + n_i \,,$$
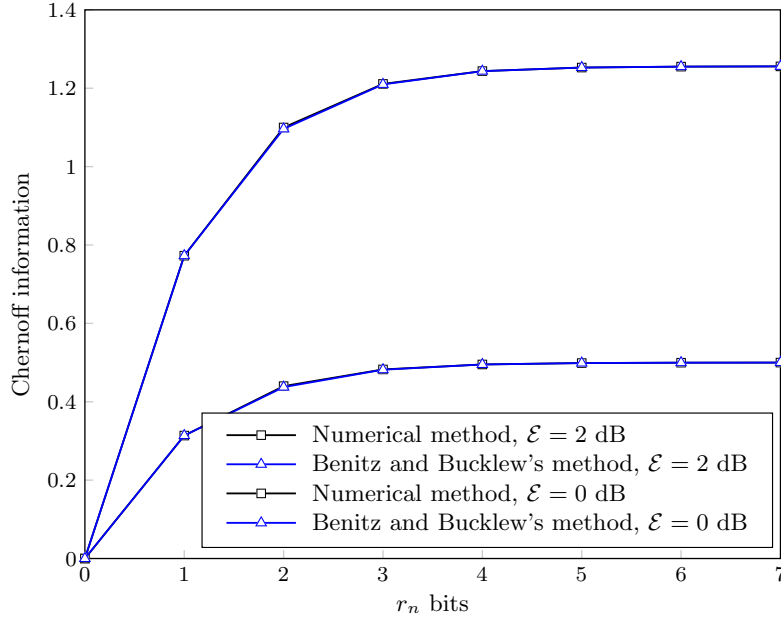
Figure 15: Chernoff information of a single sensor designed using Benitz and Bucklew's method and using the numerical method for different SNRs, $\mathcal{E}$.

we have proved, at high rates, the concavity condition holds. Also, using simulation results we have shown that the Chernoff information of a sensor designed using Benitz and Bucklew's method is a concave function of its rate. Then, we have concluded that for a network of sensors whose decision rules are designed using Benitz and Bucklew's method, an optimal rate allocation to the sensors is a uniform rate allocation.

Next, we have proposed a numerical method for the design of a sensor through a numerical optimization. In the proposed method, for a rate-r quantizer, the real interval was divided into 2^r non-overlapping intervals randomly, where each interval corresponds to an output message. Then, in an iterative fashion, the intervals are modified (in a person-by-person methodology) to maximize the Chernoff information, until a stopping criterion was satisfied. Our simulation results show that the Chernoff information resulting from the proposed numerical method is also a concave function of sensor rate. Figure 15 illustrates resulting Chernoff information curves of sensors whose decision rules are designed using the aforementioned methods for different SNRs ($\mathcal{E} = |a|^2$). As we can see from this figure, the Chernoff information resulting from both methods is a discrete concave function of
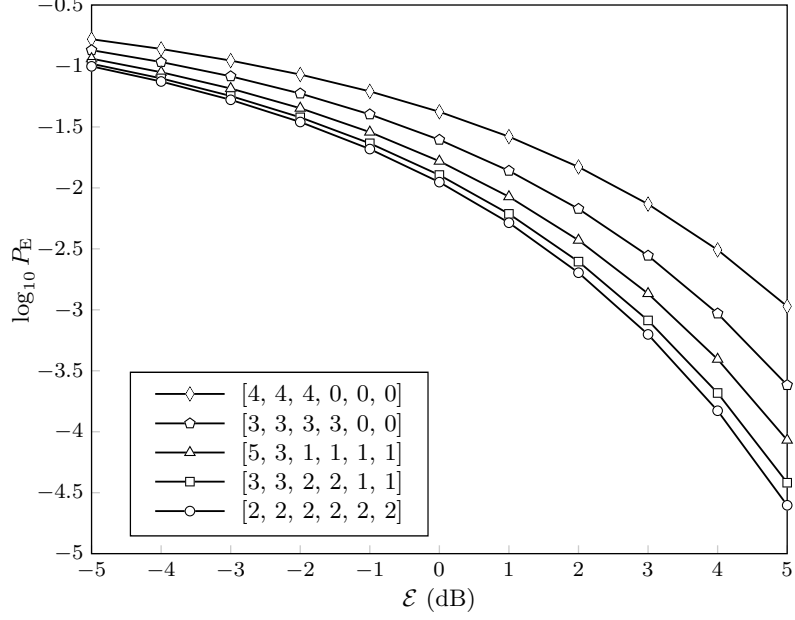
Figure 16: Error probability performance of a network of $N = 6$ sensors for different rate allocation schemes.

rate.

According to concavity of the Chernoff information, we can conclude that for a network of N sensors, where the sensors share a common MAC channel with sum rate capacity R (where R divides N), an optimal rate allocation to the sensors is a uniform rate allocation. To exemplify this, we have considered a network of $N = 6$ sensors and sum rate capacity $R = 12$ and compared their performance in terms of error probability at the FC (note that the FC applies the MAP rule to make the final decision). It can be observed from Figure 16 that the uniform rate allocation results in the best error probability performance, which is consistent with the results obtained using the Chernoff information.

In this paper, we have found sufficient conditions for the optimality of uniform rate allocation in a decentralized detection network. We have numerically verified that these conditions hold true for some sensor decision rule design methods. However, it is hard to stringently prove the required concavity property for Chernoff information. This difficulty arises mainly because of the optimization problem over parameter α in the definition of the Chernoff information. In Paper E, using the Bhattacharyya distance, we have generalized the results.

Paper E: Optimality of rate balancing in wireless sensor networks [62]

This paper generalizes the results in Paper D, in which we have considered the problem of rate allocation for decentralized detection in a network of sensors arranged as in Figure 2. In this configuration, the sensors share a MAC channel to send their observations towards the FC. The MAC channel is modeled as a sum rate constrained channel of capacity R bits per unit time.

This problem was previously studied in [21] by Chamberland and Veeravalli, where they argued that when an unbounded number of sensors with i.i.d. observations compete for rates under sum rate constraints at the input of the FC, it is often optimal to use as many sensors as possible and let each sensor communicate with the FC over a one bit link. They proposed the Chernoff information at the input of the FC as a measure of optimality given the intractability of the Bayesian probability of error. They proved that if there exists a one bit sensor rule with a Chernoff information of the sensor output that is at least half of the Chernoff information of the original observation, having as many one bit sensors as possible is optimal. They also proved that such a sensor rule exists when the observations are drawn from particular Gaussian and exponential observation models.

Our contribution is, in the same vein, to study when equal rate allocation or rate balancing is an optimal solution for a fixed number of sensors operating in a network under a common sum rate constraint. Concretely, we have driven a sufficient condition for when rate balancing is an optimal strategy in the sense that one can, without loss of optimality, assume that rates of any two sensors differ by at most one bit.

One problem with extending the results in Paper D to a general case (when R does not divide N) is that parameter α which optimizes the Chernoff information of a rate r_1 sensor is not necessarily equal to that of a sensor of rate r_2 for $r_1 \neq r_2$. In order to circumvent this difficulty we have instead of using the Chernoff information, used the Bhattacharyya distance, as the performance measure of the network. The Bhattacharyya distance has another benefit over the Chernoff information which is captured in the following lemma:

Lemma: The Bhattacharyya distance of a network of sensors, arranged as in Figure 2 and with independent observations, is equal to the sum of the Bhattacharyya distances of individual sensors, i.e.,

$$\mathcal{B}_{\underline{r}}(\underline{\gamma}) = \sum_{n=1}^{N} \mathcal{B}_{r_n}(\gamma_n),$$

where $\mathcal{B}_{r_n}(\gamma_n)$ is the Bhattacharyya distance of a sensor of rate r_n and decision rule γ_n.

This lemma implies that Bhattacharyya distances of individual sensors completely describe the Bhattacharyya distance of the network, and a network which optimizes the Bhattacharyya distance is a network with individually optimized sensors. Then by defining $\mathcal{B}^\star_{r_n}$ as the maximum Bhattacharyya distance of a sensor of rate r_n, we have found the following theorem for an optimal rate allocation:

Theorem: If, for a given observation distribution at the sensors, $\mathcal{B}^\star_{r_n}$ is a discrete concave function in the rate r_n, then rate balancing is an optimal rate allocation.

Although this theorem provides sufficient conditions for the optimality of rate balancing, it however is difficult to analytically characterize $\mathcal{B}^\star_{r_n}$ in general. To circumvent this difficulty and obtain a mathematically tractable criterion, we have found a sufficient condition under which $\mathcal{B}^\star_{r_n}$ is a concave function of rate r_n, and this sufficient condition is described in the following remark:

Remark: Conditioned on an observation X being in any given interval of the real line (or an interval of an optimum monotone quantizer), if there is a one bit quantization of X with a Bhattacharyya distance more than half of the Bhattacharyya distance of X itself, then having rate balanced sensors is optimal.

This is in agreement with Chamberland and Veeravalli's result. The conditioning on X being in an interval of an optimum monotone quantizer, is in part what generalizes the result to higher rates. However, verifying this condition is considerably harder than verifying the condition of [21] as it needs to be established for all possible optimal intervals for any arbitrary rate. Nevertheless, we proceed to discuss the Laplacian and the Gaussian observation models where the conditions of the remark can be established in practice. Then we have shown that rate balancing is an optimal strategy when the observations at the sensors are distributed as Gaussian or Laplacian.

Our next contribution was then to show how the performance of different rate allocation schemes can be partially compared using *majorization theory* [108] and also the concavity properties of the Bhattacharyya distance. Concretely, we have shown that the optimal Bhattacharyya distance of a network of sensors is Schur-concave and consequently, if a rate allocation vector is majorized by another rate allocation vector, it has higher Bhattacharyya distance. This is in agreement with our previous result in the sense that the rate allocation vector of a balanced rate network is majorized by any other rate allocation vector. Therefore a balanced rate network outperforms any other network with the same size N and the same sum rate
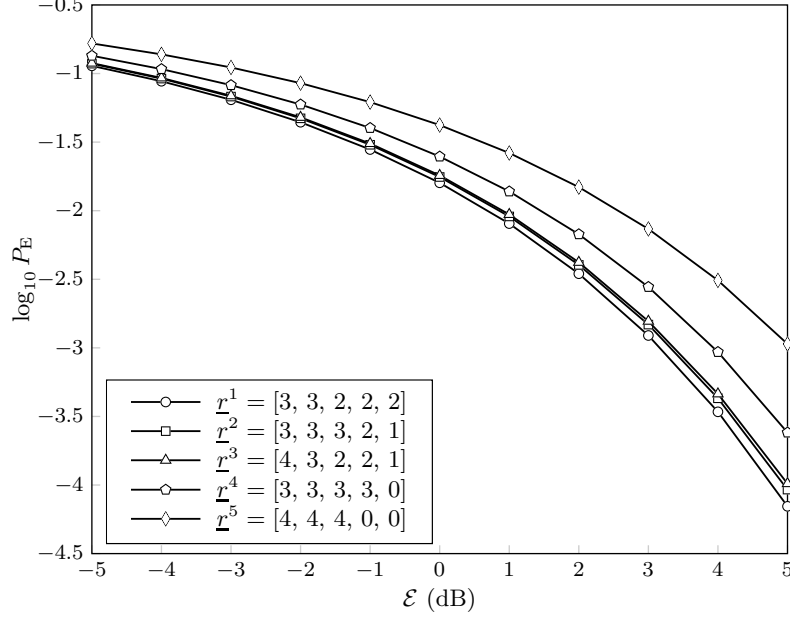
Figure 17: Error probability performance of designed sensor networks with different rate allocation schemes and Gaussian observations as a function of channels SNR $\mathcal{E}$, for $N = 5$ sensors and $R = 12$ bits per unit time.

constraint R.

Figure 17 illustrates error probability performance of designed networks of $N = 5$ sensors with different rate allocations, where sum rate capacity of the networks is $R = 12$ bits per unit time. The observation model at any sensor consists of an antipodal signal $\pm a$ in an additive unit-variance white Gaussian noise, or

$$H_0 : x_i = -a + n_i \,,$$
$$H_1 : x_i = +a + n_i \,.$$

We observe from this figure that the balanced rate allocation (denoted by $\underline{r}^1$) results in the best error probability performance, and the second best performance is for rate allocation denoted by $\underline{r}^2$ which is majorized by other rate allocation schemes.

Figure 18 shows the evolution of error probability performance of different networks with the relation $R = 2N$ with different rate allocation schemes, as a function of the number of sensors N. In the first scheme, denoted by $[2, 2, \ldots, 2, 2]$, all sensors have the same rate 2, in the second
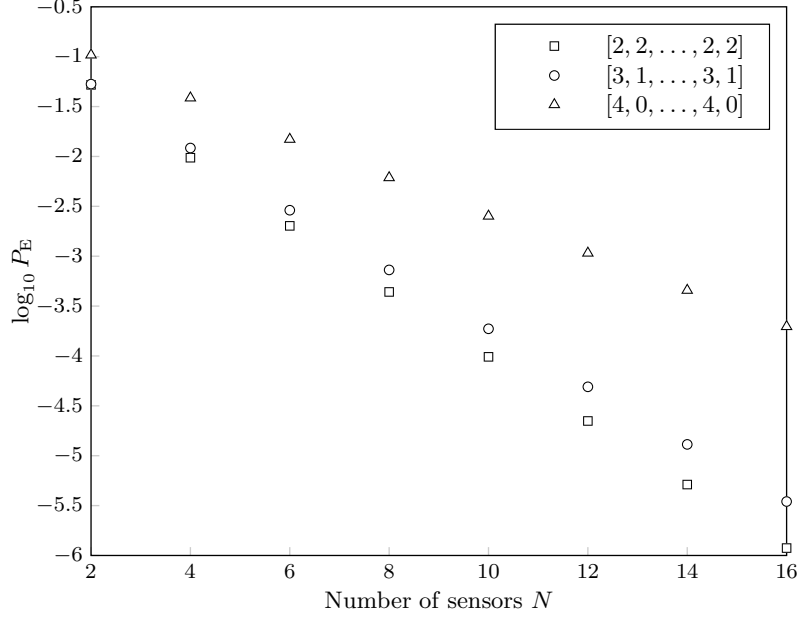
Figure 18: Evolution of error probability performance of designed sensor networks of $R = 2N$ bits per unit time with different rate allocation schemes and Gaussian observations with SNR $\mathcal{E} = 0$ dB, as a function of number of sensors N.

scheme, denoted by $[3, 1, \ldots, 3, 1]$, half of the sensors have rate 3 while the other half have rate 1, finally in the third scheme, denoted by $[4, 0, \ldots, 4, 0]$, only half of the sensors are active with rate 4. Note that all the three schemes satisfy the rate constraint $R = 2N$. As we observe from the presented results in Figure 18, the uniform rate allocation not only has the best error probability performance for any N, it also has the best error exponent, i.e., decay rate as a function of total number of sensors; this was predicted by superior Bhattacharyya distance previously.

2.3 Problem of decentralized detection in energy harvesting sensor networks

Paper F: Decentralized detection in energy harvesting wireless sensor networks [85]

In this paper and Paper G, we have considered decentralized detection networks which use energy harvesting peripheral nodes. The sensors are arranged in parallel as in Figure 5 and make noisy observations of a time varying phenomenon. Each sensor then sends a message towards a FC about the present phenomenon and the FC according to aggregate received messages makes the final decision about the present hypothesis at each time instance t. We have assumed that each sensor is an energy harvesting device equipped with a battery and is capable of harvesting all the energy it needs to communicate from its environment.

Our contribution in this paper is to formulate and analyze a decentralized binary hypothesis testing problem with energy harvesting sensors that are allowed to form a long-term energy usage policy. Our analysis is based on a queuing-theoretic model for the battery and we have assumed that the battery has infinite capacity (storage). This problem is structurally similar to the binary decentralized detection problem over a parallel network where each sensor is capable of communicating with the FC using a one bit link. We have considered the optimization of the Bhattacharyya distance between the two hypotheses at the input of the FC. We have shown how the Bhattacharyya distance jointly depends on the sensor transmission rule and the battery depletion probability. We have also found a closed-form expression for the steady state depletion probability of an infinite capacity battery sensor as

$$p_0 = \begin{cases} 0 & p_e \geq q \,, \\ 1 - \frac{p_e}{q} & \text{Otherwise} \,, \end{cases} \tag{29}$$

where p_e is related to energy arrival features and q is a function of sensor decision rules. Concretely, q is the probability of sending a positive message in an on-off keying (OOK) strategy. In order to find the expression above for the depletion probability of the sensor, we modeled the sensor battery using a birth-death process [109]. We have shown the usefulness of this result by showing how to optimize the sensor transmission rule.

Figure 19 shows the error probability performance of a network of $N = 4$ energy harvesting sensors with two types of decision rules at the sensors: traditional decision rules optimized for unconstrained sensors, and adapted decision rules for energy constrained sensors. According to this figure, using adapted decision rules at the sensors results in better a performance in terms of the error probability. Note that each observation at the sensors is either from a Rayleigh distribution of unit scale parameter or from a Rician
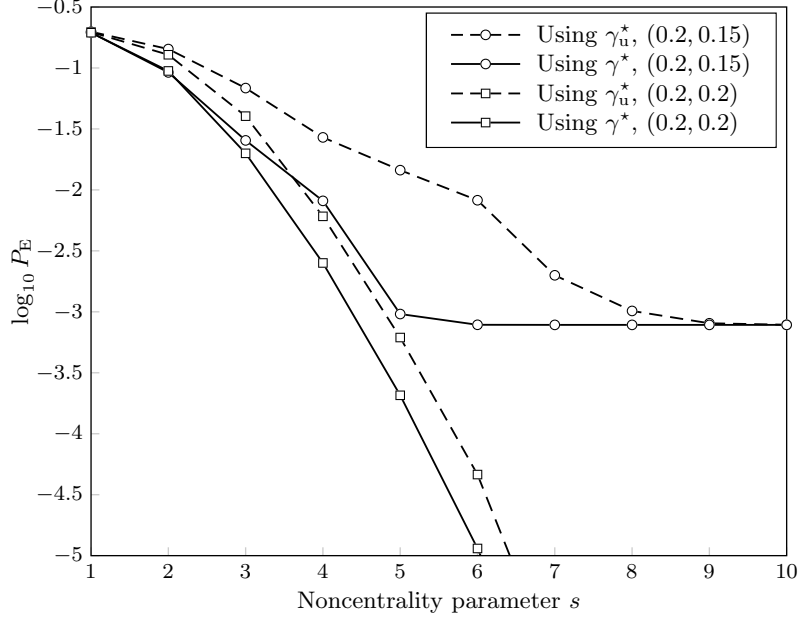
Figure 19: Error probability performance of a network of $N = 4$ energy harvesting sensors while using a traditional decision rule $\gamma_{\mathrm{u}}^{\star}$ and an adapted decision rule $\gamma^{\star}$, for different (π_1, p_e).

distribution of unit scale parameter and noncentrality parameter s, i.e.,

$$
\begin{aligned}
f_{X|H_t}(x|0) &= xe^{-\frac{x^2}{2}} \\
f_{X|H_t}(x|1) &= xe^{-\frac{x^2+s^2}{2}} I_0(xs),
\end{aligned}
\tag{30}
$$

where $I_0(z)$ is the modified Bessel function of the first kind with order zero. This observation model corresponds to an energy harvesting sensor applied to detect the presence of a known signal in Gaussian noise by received power; a relevant case for low complexity sensors in wireless sensor networks.

Paper G: Decentralized hypothesis testing in energy harvesting wireless sensor networks [84]

In this paper we have considered decentralized detection networks which use energy harvesting peripheral nodes. The sensors are arranged in parallel as in Figure 5 and make noisy observations of a time varying phenomenon H_t. Each sensor then sends a message towards a FC about the present phenomenon and the FC according to aggregate received messages makes the final decision about the present hypothesis at each time instance t. We

Figure 20: Binary asymmetric channels between sensors and the FC in Figure 5.
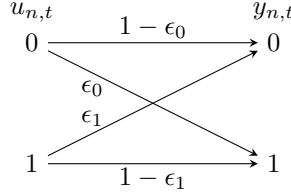
have assumed that each sensor is an energy harvesting device equipped with a battery and is capable of harvesting all the energy it needs to communicate from its environment.

This paper generalizes our results in Paper F in the sense that in contrast to the previous work, it considers the case where sensor-to-FC channels are not error-free, and it generalizes the results in Paper F to limited-capacity batteries. In other words, in this work we have assumed the battery is capable of saving at most K energy packets, for a positive integer K.

As in previous studies in energy harvesting networks [96, 98–106], we assumed that the energy arrives in packets and at each time interval the sensors are capable of harvesting at most one packet of energy. Also we assumed only sending a positive message costs a packet of energy while a negative message is conveyed through non-transmission with no cost in energy.

We have considered erroneous communication channels between sensors and the FC. We assumed each sensor-to-FC link is a binary asymmetric channel (BAC) shown in Figure 20. This is also a relevant model for fading channels when the FC uses an energy detector to detect its input $y_{n,t}$.

Our contribution in this paper is to formulate the problem using the Bhattacharyya distance at the FC and to propose a numerical method for the design of decision rules (likelihood ratio test) at the sensors. We have shown how the Bhattacharyya distance depends on different parameters in the network, e.g., depletion probability of the battery and communication channel parameters. Concretely, we have found a closed-form expression for the depletion probability at a sensor of battery size K which generalizes (29).

Further, we have found an upper bound for the Bhattacharyya distance resulting from a single sensor at the FC. In the following theorem we restate this result:

Theorem: Consider a K-slot-battery energy harvesting sensor S. Assume that the probability of harvesting energy at each time interval is p_e, and
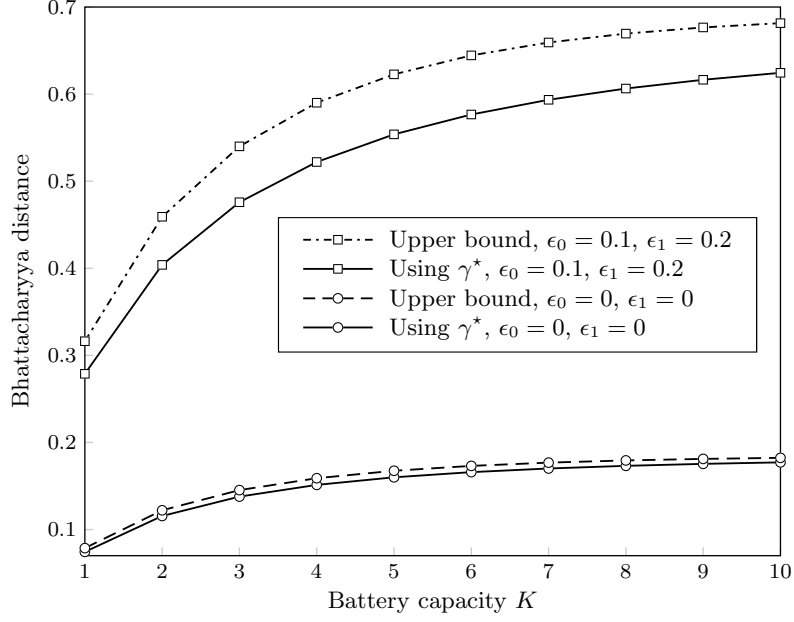
Figure 21: Bhattacharyya distance of an energy harvesting sensor as a function of battery capacity K, for noisy and noiseless communication channels and for $\pi_1 = 0.2$, $p_e = 0.15$, and corresponding upper bounds.

the a-priori probability of hypothesis $H_t = 1$ is π_1 and the sensor-to-FC channels are BAC channels as in Figure 20. The BD of this sensor at the input of the FC can not exceed

$$\overline{\mathcal{B}} \triangleq -\log\left[\sqrt{\epsilon_0\left(1 - \epsilon_1 - \overline{p}_0\delta\right)} + \sqrt{\left(1 - \epsilon_0\right)\left(\epsilon_1 + \overline{p}_0\delta\right)}\right],$$

where

$$\overline{p}_0 = \left[1 + \frac{1}{1 - \pi_1}\sum_{k=1}^{K}\left(\frac{p_e(1 - \pi_1)}{\pi_1(1 - p_e)}\right)^k\right]^{-1},$$

and

$$\delta = 1 - \epsilon_0 - \epsilon_1.$$

Note that according to (27), upper bounded Bhattacharrya distance can be translated to lower bounded error probability, which means the error probability is not zero-approaching while the corresponding Bhattacharyya distance is upper bounded.
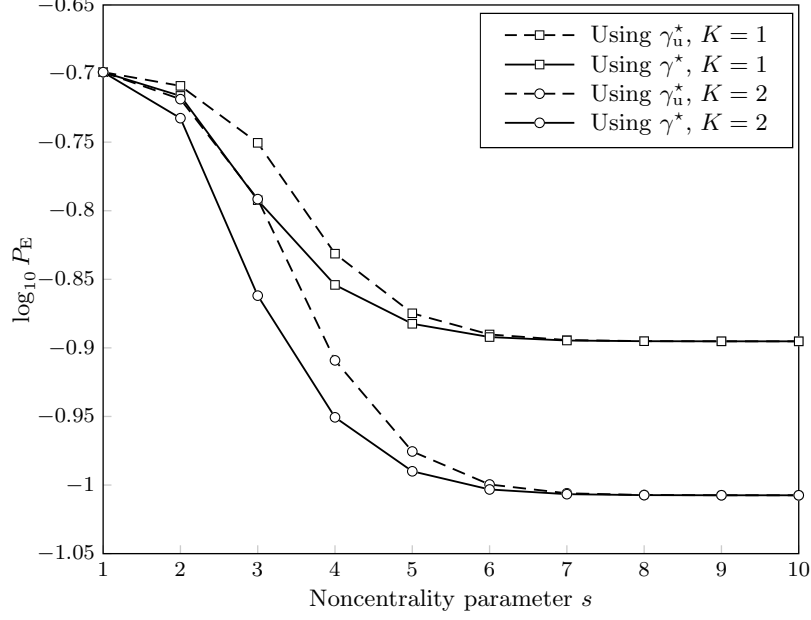
Figure 22: Error probability performance of networks with $N = 4$ energy harvesting sensors with different battery capacities, when $\pi_1 = 0.2$ and $p_e = 0.15$, and noisy communication channels with $\epsilon_0 = 0.1$ and $\epsilon_1 = 0.2$.

Figure 21 shows the Bhattacharyya distance of an energy harvesting sensor as a function of battery size K, and corresponding upper bounds according to the theorem above, where the observation model at the sensor is according to (30), for $s = 5$.

In order to show the benefit of our results, in Figure 21 we show the error probability performance of a network of $N = 4$ energy harvesting sensors, designed using the proposed formulation ($\gamma^\star$) and the traditional formulation ($\gamma_{\mathrm{u}}^\star$). We observe from this figure that the proposed formulation results in better error probability performance. This result is parallel with our previous results based on the Bhattacharyya distance.

3   Conclusions

This dissertation aims to push the frontier of knowledge relating to de-
centralized detection when the network is subject to some communication
and/or energy constraints. Specifically, it formulates decentralized detection
problems and presents numerical algorithms and analytical results relating
to the optimal design of decision rules at the sensors in a network.

This dissertation mainly examines the Bayesian design of decision rules
at the sensors, by minimizing the Bayes' misclassification error (Paper A,
Paper B, and Paper C). When direct use of the error probability is not
tractable, it invokes dissimilarity measures like the Chernoff information
(Paper D) and the Bhattacharyya distance (Paper E, Paper F, and Paper G)
to study the structure of an optimal decentralized detection network.

It was effectively established that applying the PBP methodology for the
design of decision rules at sensors, arranged in an arbitrary topology and
making conditionally independent observations, is analogous to the design
of decision rules of sensors in an equivalent network of two sensors. Though
this result leads to a computationally efficient way of designing the sensor's
decision rules in a network, it is however not completely known how this
result can be generalized to the case where the sensors make correlated
observations. To tackle the problem of designing decision rules at the sensors
with correlated observations, the dissertation proposes a PBP optimization
method for the design of decision rules of sensors arranged in a parallel
topology and making correlated observations, by directly minimizing the
probability of making erroneous decisions at the FC.

An analytical proof was given that shows the optimality of rate balanc-
ing in decentralized detection networks. By maximizing the Bhattacharyya
distance for a network of sensors, arranged in MAC and making i.i.d. obser-
vations, the dissertation proves the optimality of having equal rate sensors.
This extends the previous significant asymptotic result on the optimality
of allocating single bit rates to the sensors. An optimal rate allocation
strategy for the case where sensors have non-identical observation models
remains to be found, but using a similar approach (minimizing the error ex-
ponent in place of the error probability) seems likely to be more challenging
in this case. One reason might be (as we have seen in Paper A) because
of sub-optimality of monotone quantizers for correlated observations. This
problem has not yet been addressed: even the structure of an asymptoti-
cally optimal network is not known. Furthermore, when the observations
at the sensors are correlated, these results do not hold anymore and a more
in-depth analysis for finding an optimal rate allocation strategy is needed.
Another possible extension could be extending these results to composite
hypothesis testing problems.

Further, this dissertation has formulated a decentralized detection prob-
lem with system costs due to the random behavior of energy available at the

sensors. Concretely, decentralized detection, in which the sensors are energy harvesting devices and are able to acquire all the energy they need from the environment, was considered. This formulation is based on the Bhattacharyya distance and generalizes the previous formulations. The analysis in this dissertation provides the first steps in the area of decentralized detection in the presence of energy harvesting sensors and numerous in-depth analyses of these type of networks are needed. In our analysis, we assumed an independent observation model and uncorrelated energy arrival processes at different sensors. Further study could take into account possible correlations between observations and the energy arrival processes. In addition to that, a more in-depth study of such models could consider more sophisticated decision-making rules at the sensors, i.e., energy-dependent-threshold-tests at the sensors are examples of such decision rules. Another possible extension to our work is to study the structure of optimal rate allocation to the network of energy harvesting sensors arranged as in MAC. Finally, in this dissertation, decentralized detection in energy harvesting sensor networks, where the sensors are arranged in parallel, was considered, while study of energy harvesting networks for arbitrary topologies remains largely open.

References

[1] C. S. Raghavendra, K. M. Sivalingam, and T. Znati, *Wireless sensor networks.* Springer, 2006.

[2] B. Krishnamachari, *Networking wireless sensors.* Cambridge University Press, 2005.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, 2002.

[4] V. Veeravalli and P. Varshney, "Distributed inference in wireless sensor networks," *Phil. Trans. A, Math. Phys. Eng. Sci.*, vol. 370, no. 1958, pp. 100–117, 2012.

[5] J.-F. Chamberland and V. Veeravalli, "Wireless sensors in distributed detection applications," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 16–25, 2007.

[6] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[7] R. Tenney and N. R. Sandell, "Detection with distributed sensors," in *Proc. 19th IEEE Conf. Decision and Control including the Symposium on Adaptive Processes*, vol. 19, 1980, pp. 433–437.

[8] L. K. Ekchian and R. R. Tenney, "Detection networks," in *Proc. IEEE Conf. Decision and Control*, 1982, pp. 686–691.

[9] P. Varshney, *Distributed Detection and Data Fusion.* Springer-Verlag New York, Inc., 1996.

[10] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 1, pp. 98–101, Jan 1986.

[11] Z. Chair and P. Varshney, "Distributed bayesian hypothesis testing with distributed data fusion," *IEEE Trans. Systems, Man and Cybernetics*, vol. 18, no. 5, pp. 695–699, 1988.

[12] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors: Part I–Fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, 1997.

[13] R. Blum, S. Kassam, and H. Poor, "Distributed detection with multiple sensors II. Advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, 1997.

[14] B. Chen, L. Tong, and P. Varshney, "Channel aware distributed detection in wireless sensor networks," *IEEE Signal Process. Mag.*, 2006.

[15] H. Chen, B. Chen, and P. Varshney, "A new framework for distributed detection with conditionally dependent observations," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1409–1419, 2012.

[16] J. N. Tsitsiklis, "Decentralized detection," *Adv. Statist. Signal Process.*, vol. 2, no. 2, pp. 297–344, 1993.

[17] M. Longo, T. Lookabaugh, and R. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 241–255, Mar. 1990.

[18] W. A. Hashlamoun and P. Varshney, "An approach to the design of distributed bayesian detection structures," *IEEE Trans. Systems, Man, Cybern.*, vol. 21, no. 5, pp. 1206–1211, 1991.

[19] G. Polychronopoulos and J. N. Tsitsiklis, "Explicit solutions for some simple decentralized detection problems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, no. 2, pp. 282–292, Mar 1990.

[20] P. Willett and D. Warren, "The suboptimality of randomized tests in distributed and quantized detection systems," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 355–361, March 1992.

[21] J.-F. Chamberland and V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.

[22] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 822–833, 2007.

[23] C. R. Berger, M. Guerriero, S. Zhou, and P. Willett, "PAC vs. MAC for decentralized detection using noncoherent modulation," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3562–3575, 2009.

[24] F. Li, J. S. Evans, and S. Dey, "Decision fusion over noncoherent fading multiaccess channels," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4367–4380, 2011.

[25] D. Ciuonzo, G. Romano, and P. S. Rossi, "Optimality of received energy in decision fusion over rayleigh fading diversity mac with non-identical sensors," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 22–27, 2013.

[26] S. A. Aldosari and J. M. Moura, "Detection in decentralized sensor networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, vol. 2, 2004, pp. ii–277.

[27] P. Swaszek, "On the performance of serial networks in distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 254–260, 1993.

[28] J. Papastavrou and M. Athans, "Distributed detection by a large team of sensors in tandem," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 639–653, 1992.

[29] R. Viswanathan, S. C. A. Thomopoulos, and R. Tumuluri, "Optimal serial distributed decision fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 366–376, 1988.

[30] I. Bahceci, G. Al-Regib, and Y. Altunbasak, "Serial distributed detection for wireless sensor networks," in *Proc. Int. Symp. Inf. Theory*, Sep. 2005, pp. 830–834.

[31] L. Yin, Y. Wang, and D.-W. Yue, "Serial distributed detection performance analysis in wireless sensor networks under noisy channel," in *Proc. 5th Int. Conf. Wireless Commun., Networking and Mobile Computing, WiCom 09.*, 2009, pp. 1–4.

[32] G. Fabeck and R. Mathar, "Optimization of linear wireless sensor networks for serial distributed detection applications," in *Proc. 71th Int. IEEE Conf. Vehicular Technology*, May 2010, pp. 1–5.

[33] P. Yang, B. Chen, H. Chen, and P. Varshney, "Tandem distributed detection with conditionally dependent observations," in *Proc. 15th Int. Conf. Inf. Fusion (FUSION)*, 2012, pp. 1808–1813.

[34] Z.-B. Tang, K. Pattipati, and D. Kleinman, "Optimization of detection networks. I. Tandem structures," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 5, pp. 1044–1059, 1991.

[35] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *The Review of Economic Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.

[36] L. Smith and P. Sørensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, 2000.

[37] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis, "On learning with finite memory," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6859–6872, 2013.

[38] V. Krishnamurthy and H. V. Poor, "A tutorial on interactive sensing in social networks," *IEEE Trans. Comput. Social Systems*, vol. 1, no. 1, pp. 3–21, 2014.

[39] W. P. Tay, "Whose opinion to follow in multihypothesis social learning? a large deviations perspective," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 2, pp. 344–359, March 2015.

[40] A. Lalitha, A. Sarwate, and T. Javidi, "Social learning and distributed hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2014, pp. 551–555.

[41] V. Krishnamurthy, O. N. Gharehshiran, and M. Hamdi, "Interactive sensing and decision making in social networks," *arXiv preprint arXiv:1405.1129*, 2014.

[42] J. Ho, W. P. Tay, and T. Q. Quek, "Robust detection and social learning in tandem networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, 2014, pp. 5457–5461.

[43] W. P. Tay, J. Tsitsiklis, and M. Win, "Data fusion trees for detection: Does architecture matter?" *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4155–4168, Sep. 2008.

[44] S. Alhakeem and P. Varshney, "A unified approach to the design of decentralized detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 1, pp. 9–20, 1995.

[45] V. Saligrama, M. Alanyali, and O. Savas, "Distributed detection in sensor networks with packet losses and finite capacity links," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4118–4132, 2006.

[46] S. Kar and J. M. Moura, "Distributed consensus algorithms in sensor networks: Quantized data and random link failures," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1383–1400, 2010.

[47] S. S. Ram, V. V. Veeravalli, and A. Nedic, "Distributed and recursive parameter estimation in parametrized linear state-space models," *IEEE Trans. Autom. Control*, vol. 55, no. 2, pp. 488–492, 2010.

[48] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "Bayesian detection in bounded height tree networks," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4042–4051, 2009.

[49] J. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Trans. Autom. Control*, vol. 30, no. 5, pp. 440–446, 1985.

[50] B. Aazhang and H. Poor, "On optimum and nearly optimum data quantization for signal detection," *IEEE Trans. Commun.*, vol. 32, no. 7, pp. 745–751, 1984.

[51] H. V. Poor, "Fine quantization in signal detection and estimation," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 960–972, 1988.

[52] H. Poor and J. B. Thomas, "Applications of ali-silvey distance measures in the design generalized quantizers for binary decision systems," *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 893–900, 1977.

[53] S. Kassam, "Optimum quantization for signal detection," *IEEE Trans. Commun.*, vol. 25, no. 5, pp. 479–484, 1977.

[54] G. R. Benitz and J. A. Bucklew, "Asymptotically optimal quantizers for detection of iid data," *IEEE Trans. Inf. Theory*, vol. 35, no. 2, pp. 316–325, 1989.

[55] R. Gupta and A. O. Hero, "High-rate vector quantization for detection," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1951–1969, 2003.

[56] R. R. Tenney and N. R. Sandell Jr, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, pp. 501–510, 1981.

[57] W. W. Irving and J. N. Tsitsiklis, "Some properties of optimal thresholds in decentralized detection," *IEEE Trans. Autom. Control*, vol. 39, no. 4, pp. 835–838, Apr 1994.

[58] J. N. Tsitsiklis, "Extremal properties of likelihood-ratio quantizers," *IEEE Trans. Commun.*, vol. 41, no. 4, pp. 550–558, 1993.

[59] D. Warren and P. Willett, "Optimum quantization for detector fusion: some proofs, examples, and pathology," *J. Franklin Inst.*, vol. 336, no. 2, pp. 323–359, 1999.

[60] B. Chen and P. K. Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 693–699, Feb 2005.

[61] B. Liu and B. Chen, "Channel-optimized quantizers for decentralized detection in sensor networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3349–3358, July 2006.

[62] A. Tarighati and J. Jaldén, "Optimality of rate balancing in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3735–3749, July 2016.

[63] T. Kailath, "The divergence and bhattacharyya distance measures in signal selection," *IEEE Trans. Commun.*, vol. 15, no. 1, pp. 52–60, Feb. 1967.

[64] A. Tarighati and J. Jaldén, "Rate allocation for decentralized detection in wireless sensor networks," in *Proc. 16th IEEE Int. Workshop Signal Process. Advances in Wireless Commun. (SPAWC)*, June 2015, pp. 341–345.

[65] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "On the sub-exponential decay of detection error probabilities in long tandems," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4767–4771, 2008.

[66] M. Al-Ibrahim and S. AlHakeem, "Optimization of a serial distributed detection system with 2 bits communication constraint," *Int. J. Syst. Sci.*, vol. 32, no. 9, pp. 1169–1175, 2001.

[67] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Contr., Signals, Syst.*, vol. 1, no. 2, pp. 167–182, 1988.

[68] T. M. Cover, "Hypothesis testing with finite statistics," *The Annals of Mathematical Statistics*, vol. 40, no. 3, pp. 828–835, 1969.

[69] J. Koplowitz, "Necessary and sufficient memory size for m-hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 44–46, 1975.

[70] C. Lee and J. Chao, "Optimum local decision space partitioning for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 4, pp. 536–544, 1989.

[71] F. Nielsen, "Generalized bhattacharyya and chernoff upper bounds on bayes error using quasi-arithmetic means," *Pattern Recognition Letters*, vol. 42, pp. 25–34, 2014.

[72] A. Lapidoth, *A Foundation in Digital Communication.* Cambridge University Press, 2009.

[73] M. Feder and N. Merhav, "Relations between entropy and error probability," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 259–266, 1994.

[74] A. Tarighati and J. Jaldén, "Bayesian design of decentralized hypothesis testing under communication constraints," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, May 2014, pp. 7624–7628.

[75] ——, "Bayesian design of tandem networks for distributed detection with multi-bit sensor decisions," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1821–1831, 2015.

[76] ——, "A general method for the design of tree networks under communication constraints," in *Proc. 17th Int. Conf. inf. Fusion (FUSION)*, July 2014, pp. 1–7.

[77] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Stat.*, pp. 493–507, 1952.

[78] F. Nielsen, "Chernoff information of exponential families," *arXiv preprint arXiv:1102.2684*, 2011.

[79] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, 2012.

[80] Y. Lee and Y. Sung, "Generalized chernoff information for mismatched bayesian detection and its application to energy detection," *IEEE Signal Process. Letters*, vol. 19, no. 11, pp. 753–756, Nov. 2012.

[81] G. Fabeck and R. Mathar, "Chernoff information-based optimization of sensor networks for distributed detection," in *Proc. IEEE Int. Symp. Signal Process. and Inf. Tech. (ISSPIT)*, Dec. 2009, pp. 606–611.

[82] J.-F. Chamberland and V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1007–1015, Aug. 2004.

[83] D. Boekee and J. van der Lubbe, "Some aspects of error bounds in feature selection," *Pattern Recognition*, vol. 11, no. 5–6, pp. 353–360, 1979.

[84] A. Tarighati, J. Gross, and J. Jaldén, "Decentralized hypothesis testing in energy harvesting wireless sensor networks," *to be submitted to IEEE Trans. Signal Process.*

[85] ——, "Decentralized detection in energy harvesting wireless sensor networks," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Aug. 2016.

[86] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1326–1336, April 2010.

[87] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 51–60, 2002.

[88] S. J. Baek, G. De Veciana, and X. Su, "Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation," *IEEE J. sel. Areas Commun.*, vol. 22, no. 6, pp. 1130–1140, 2004.

[89] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, 2005.

[90] P. Nuggehalli, V. Srinivasan, and R. R. Rao, "Energy efficient transmission scheduling for delay constrained wireless networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 531–539, 2006.

[91] A. Sinha and A. Chandrakasan, "Dynamic power management in wireless sensor networks," *IEEE Design & Test of Computers*, vol. 18, no. 2, pp. 62–74, 2001.

[92] S. Singh, M. Woo, and C. S. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proc. 4th ACM/IEEE Int. Conf. Mobile Comput. Network.*, 1998, pp. 181–190.

[93] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embedded Computing Systems*, vol. 6, no. 4, p. 32, 2007.

[94] D. Niyato, E. Hossain, M. M. Rashid, and V. K. Bhargava, "Wireless sensor networks with energy harvesting technologies: a game-theoretic approach to optimal energy management," *IEEE Wireless Commun.*, vol. 14, no. 4, pp. 90–96, 2007.

[95] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover, and K. Huang, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas Commun.*, 2015.

[96] D. Gunduz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Commun. Magazine*, vol. 52, no. 1, pp. 210–216, 2014.

[97] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Communications Surveys & Tutorials,*, vol. 13, no. 3, pp. 443–461, 2011.

[98] N. Michelusi, K. Stamatiou, and M. Zorzi, "On optimal transmission policies for energy harvesting devices," in *Proc. Inf. Theory and Applications Workshop (ITA)*, 2012, pp. 249–254.

[99] N. Michelusi and M. Zorzi, "Optimal adaptive random multiaccess in energy harvesting wireless sensor networks," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1355–1372, 2015.

[100] N. Michelusi, K. Stamatiou, and M. Zorzi, "Transmission policies for energy harvesting sensors with time-correlated energy supply," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2988–3001, 2013.

[101] R. Valentini and M. Levorato, "Optimal aging-aware channel access control for wireless networks with energy harvesting," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2016, pp. 2754–2758.

[102] A. Biason and M. Zorzi, "Joint online transmission and energy transfer policies for energy harvesting devices with finite batteries," in *Proc. 21th European Wireless Conference European Wireless*, May 2015, pp. 1–7.

[103] R. Valentini, M. Levorato, and F. Santucci, "Aging aware random channel access for battery-powered wireless networks," *IEEE Wireless Commun. Letters*, vol. 5, no. 2, pp. 176–179, 2016.

[104] A. Biason and M. Zorzi, "Transmission policies for an energy harvesting device with a data queue," in *Proc. Int. Conf. Computing, Networking and Commun. (ICNC)*, 2015, pp. 189–195.

[105] K. Tutuncuoglu, O. Ozel, A. Yener, and S. Ulukus, "Binary energy harvesting channel with finite energy storage," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2013, pp. 1591–1595.

[106] ——, "Improved capacity bounds for the binary energy harvesting channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2014, pp. 976–980.

[107] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I.* Wiley-Interscience, 2001.

[108] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications.* Springer, 2010.

[109] D. Gross, J. F. Shortie, J. M. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory.* John Wiley & Sons, Inc., 2008.

Part II

Included papers